



George Karlin-Neumann  
Francisco Bizouarn *Editors*

# Digital PCR

## Methods and Protocols

# METHODS IN MOLECULAR BIOLOGY

*Series Editor*

John M. Walker

School of Life and Medical Sciences  
University of Hertfordshire  
Hatfield, Hertfordshire, AL10 9AB, UK

For further volumes:  
<http://www.springer.com/series/7651>

# Digital PCR

## Methods and Protocols

Edited by

**George Karlin-Neumann and Francisco Bizouarn**

*Digital Biology Center  
Bio-Rad Laboratories, Pleasanton, CA, USA*



*Editors*

George Karlin-Neumann  
Digital Biology Center  
Bio-Rad Laboratories  
Pleasanton, CA, USA

Francisco Bizouarn  
Digital Biology Center  
Bio-Rad Laboratories  
Pleasanton, CA, USA

ISSN 1064-3745

ISSN 1940-6029 (electronic)

Methods in Molecular Biology

ISBN 978-1-4939-7776-5

ISBN 978-1-4939-7778-9 (eBook)

<https://doi.org/10.1007/978-1-4939-7778-9>

Library of Congress Control Number: 2018937639

© Springer Science+Business Media, LLC, part of Springer Nature 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Humana Press imprint is published by the registered company Springer Science+Business Media, LLC part of Springer Nature.

The registered company address is: 233 Spring Street, New York, NY 10013, U.S.A.

---

## Preface

The ability to detect and manipulate a single, specific DNA or RNA molecule via the polymerase chain reaction (PCR) has revolutionized our understanding of biology and the possibilities for influencing it. This remarkable technique has been successively honed over the past two decades to improve upon its ability to detect and quantify targets of interest. Digital PCR is both one of the oldest and newest implementations of PCR and with the recent technical innovation of scalable, droplet-based digital PCR, brings to the practice the greater ease of use, precision, specificity and reproducibility needed to further advance many areas of investigation and clinical practice. This volume will provide guidance and illustrate how it is helping to make deeper inroads in:

- Infectious disease;
- Evolution of cancer and response to treatment;
- Genome structural variation and associated phenotypes;
- Prevalence of somatic mosaicism;
- Genome editing and cell therapy;
- Non-invasive blood monitoring of fetus, organ or tumor status;
- Environmental monitoring; and
- Food testing for pathogens and GMO's.

For these and other diverse applications, it details optimal experimental design for achieving the user's needs for precision and sensitivity, assay design considerations for various target types and sample types, insights into data analysis and interpretation, and it reveals other related benefits of sample partitioning such as target size determination and linkage measurements for haplotyping. This volume should prove useful for a wide range of specialists including geneticists, molecular biologists, virologists, immunologists, oncologists, pathologists—those involved in basic and translational research, as well as clinicians and diagnosticians—and those interested in applied and environmental sciences.

*Pleasanton, CA, USA*

*George Karlin-Neumann  
Francisco Bizouarn*

---

# Contents

Preface .....	v
Contributors .....	xi

## PART I BACKGROUND CONCEPTS

1 Entering the Pantheon of 21 <sup>st</sup> Century Molecular Biology Tools: A Perspective on Digital PCR .....	3
<i>George Karlin-Neumann and Francisco Bizouarn</i>	
2 Basic Concepts and Validation of Digital PCR Measurements .....	11
<i>Leonardo Pinheiro and Kerry R. Emslie</i>	
3 Fundamentals of Counting Statistics in Digital PCR: I Just Measured Two Target Copies—What Does It Mean? .....	25
<i>Svilen Tzonev</i>	
4 Control Materials and Digital PCR Methods for Evaluation of Circulating Cell-Free DNA Extractions from Plasma .....	45
<i>Alexandra S. Whale, Ana Fernandez-Gonzalez, Alice Gutteridge, and Alison S. Devonshire</i>	

## PART II ABSOLUTE QUANTIFICATION

5 Multiplex Droplet Digital PCR Protocols for Quantification of GM Maize Events .....	69
<i>David Dobnik, Bjørn Spilsberg, Alexandra Bogožalec Košir, Dejan Štebih, Dany Morisset, Arne Holst-Jensen, and Jana Žel</i>	
6 Using Droplet Digital PCR to Detect Coinfection of Human Herpesviruses 6A and 6B (HHV-6A and HHV-6B) in Clinical Samples .....	99
<i>Ashley Vellucci, Emily C. Leibovitch, and Steven Jacobson</i>	
7 Biomarkers in Cerebrospinal Fluid: Analysis of Cell-Free Circulating Mitochondrial DNA by Digital PCR .....	111
<i>Petar Podlesniy and Ramon Trullas</i>	
8 Testing of General and Human-Associated Fecal Contamination in Waters .....	127
<i>Yiping Cao, Meredith R. Raith, and John F. Griffith</i>	

## PART III COPY NUMBER VARIATION

9 Analyzing Copy Number Variation with Droplet Digital PCR .....	143
<i>Avery Davis Bell, Christina L. Usher, and Steven A. McCarroll</i>	
10 Assessing HER2 Amplification in Plasma cfDNA .....	161
<i>Isaac Garcia-Murillas and Nicholas C. Turner</i>	

- 11 Detection and Quantification of Mosaic Genomic DNA Variation in Primary Somatic Tissues Using ddPCR: Analysis of Mosaic Transposable-Element Insertions, Copy-Number Variants, and Single-Nucleotide Variants ..... 173  
*Bo Zhou, Michael S. Haney, Xiaowei Zhu, Reenal Pattni, Alexej Abyzov, and Alexander E. Urban*

#### PART IV RARE MUTATION AND RARE ALLELE DETECTION

- 12 Monitoring of Response and Resistance in Plasma of EGFR-Mutant Lung Cancer Using Droplet Digital PCR ..... 193  
*Tanan Kuang, Allison O'Connell, Adrian G. Sacher, Nora Feeney, Ryan S. Alden, Geoffrey R. Oxnard, and Cloud P. Paweletz*
- 13 Detection of Cancer DNA in Early Stage and Metastatic Breast Cancer Patients ..... 209  
*Arielle J. Medford, Riaz N. Gillani, and Ben Ho Park*
- 14 Droplet Digital PCR for Minimal Residual Disease Detection in Mature Lymphoproliferative Disorders ..... 229  
*Daniela Drandi, Simone Ferrero, and Marco Ladetto*
- 15 Quantitation of JAK2 V617F Allele Burden by Using the QuantStudio™ 3D Digital PCR System ..... 257  
*Elena Kinz and Axel Muendlein*
- 16 Novel Multiplexing Strategies for Quantification of Rare Alleles Using ddPCR ..... 275  
*Miguel Alcaide and Ryan D. Morin*
- 17 Identification and Use of Personalized Genomic Markers for Monitoring Circulating Tumor DNA ..... 303  
*Yilun Chen, Anthony M. George, Eleonor Olsson, and Lao H. Saal*
- 18 Single Color Multiplexed ddPCR Copy Number Measurements and Single Nucleotide Variant Genotyping ..... 323  
*Christina M. Wood-Bouwens and Hanlee P. Ji*
- 19 A Universal Droplet Digital PCR Approach for Monitoring of Graft Health After Transplantation Using a Preselected SNP Set ..... 335  
*Julia Beck, Michael Oellerich, and Ekkehard Schütz*
- 20 Detection and Quantification of HDR and NHEJ Induced by Genome Editing at Endogenous Gene Loci Using Droplet Digital PCR ..... 349  
*Yuichiro Miyaoka, Steven J. Mayerl, Amanda H. Chan, and Bruce R. Conklin*
- 21 DNA Methylation Analysis Using Droplet Digital PCR ..... 363  
*Ming Yu, Tai J. Heinzerling, and William M. Grady*

#### PART V GENE EXPRESSION AND RNA QUANTIFICATION

- 22 Simultaneous Quantification of Multiple Alternatively Spliced mRNA Transcripts Using Droplet Digital PCR ..... 387  
*Bing Sun and Yun-Ling Zheng*

23	Using Droplet Digital PCR to Analyze Allele-Specific RNA Expression . . . . .	401
	<i>Nolan Kamitaki, Christina L. Usher, and Steven A. McCarroll</i>	
24	Very Low Abundance Single-Cell Transcript Quantification with 5-Plex ddPCR <sup>TM</sup> Assays . . . . .	423
	<i>George Karlin-Neumann, Bin Zhang, and Claudia Litterst</i>	
25	Quantification of Circulating MicroRNAs by Droplet Digital PCR . . . . .	445
	<i>Manuela Ferracin and Massimo Negrini</i>	
26	Droplet Digital PCR for Absolute Quantification of Extracellular MicroRNAs in Plasma and Serum: Quantification of the Cancer Biomarker hsa-miR-141 . . . . .	459
	<i>Maria D. Giraldez, John R. Chevillet, and Muneesh Tewari</i>	

## PART VI OTHER USES OF PARTITIONING

27	Droplet Digital <sup>TM</sup> PCR Next-Generation Sequencing Library QC Assay . . . . .	477
	<i>Nicholas J. Heredia</i>	
28	Phasing DNA Markers Using Digital PCR . . . . .	489
	<i>John Regan and George Karlin-Neumann</i>	
29	ddTRAP: A Method for Sensitive and Precise Quantification of Telomerase Activity . . . . .	513
	<i>Andrew T. Ludlow, Dawne Shelton, Woodring E. Wright, and Jerry W. Shay</i>	
30	Highly Efficient and Reliable DNA Aptamer Selection Using the Partitioning Capabilities of ddPCR: The Hi-Fi SELEX Method . . . . .	531
	<i>Aaron Ang, Eric Ouellet, Karen C. Cheung, and Charles Haynes</i>	
	<i>Index</i> . . . . .	555

---

## Contributors

ALEXEJ ABYZOV • *Department of Health Sciences Research, Center for Individualized Medicine, Mayo Clinic, Rochester, MN, USA*

MIGUEL ALCAIDE • *Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, BC, Canada*

RYAN S. ALDEN • *Lowe Center for Thoracic Oncology, Dana-Farber Cancer Institute, Boston, MA, USA; Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA*

AARON ANG • *Michael Smith Laboratories, Department of Chemical and Biological Engineering, University of British Columbia, Vancouver, BC, Canada; Centre for Blood Research, University of British Columbia, Vancouver, BC, Canada*

JULIA BECK • *Chronix Biomedical GmbH, Göttingen, Germany*

AVERY DAVIS BELL • *Department of Genetics, Harvard Medical School, Boston, MA, USA; Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA; Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA*

FRANCISCO BIZOUARN • *Digital Biology Center, Bio-Rad Laboratories, Pleasanton, CA, USA*

YIPING CAO • *Source Molecular Corporation, Miami, FL, USA; Southern California Coastal Water Research Project Authority, Costa Mesa, CA, USA*

AMANDA H. CHAN • *Gladstone Institute of Cardiovascular Disease, San Francisco, CA, USA*

YILUN CHEN • *Translational Oncogenomics Unit, Division of Oncology and Pathology, Department of Clinical Sciences, Lund University, Lund, Sweden*

KAREN C. CHEUNG • *Centre for Blood Research, University of British Columbia, Vancouver, BC, Canada; Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada*

JOHN R. CHEVILLET • *Institute for Systems Biology, Seattle, WA, USA*

BRUCE R. CONKLIN • *Gladstone Institute of Cardiovascular Disease, San Francisco, CA, USA; Department of Medicine, University of California San Francisco, San Francisco, CA, USA; Department of Cellular and Molecular Pharmacology, University of California San Francisco, San Francisco, CA, USA*

ALISON S. DEVONSHIRE • *Molecular and Cell Biology Team, LGC, Teddington, UK*

DAVID DOBNIK • *Department of Biotechnology and Systems Biology, National Institute of Biology, Ljubljana, Slovenia*

DANIELA DRANDI • *Department of Molecular Biotechnologies and Health Sciences, Hematology Division, University of Torino, Torino, Italy*

KERRY R. EMLIE • *National Measurement Institute, Lindfield, NSW, Australia*

NORA FEENEY • *Belfer Center for Applied Cancer Science, Dana-Farber Cancer Institute, Boston, MA, USA*

ANA FERNANDEZ-GONZALEZ • *Molecular and Cell Biology Team, LGC, Teddington, UK*

MANUELA FERRACIN • *Department of Experimental, Diagnostic and Specialty Medicine—DIMES, University of Bologna, Bologna, Italy*

SIMONE FERRERO • *Department of Molecular Biotechnologies and Health Sciences, Hematology Division, University of Torino, Torino, Italy*

- ISAAC GARCIA-MURILLAS • *Breast Cancer Now Research Centre, The Institute of Cancer Research, London, UK*
- ANTHONY M. GEORGE • *Translational Oncogenomics Unit, Division of Oncology and Pathology, Department of Clinical Sciences, Lund University, Lund, Sweden*
- RIAZ N. GILLANI • *Johns Hopkins School of Medicine, Baltimore, MD, USA; Boston Children's Hospital, Boston, MA, USA*
- MARIA D. GIRALDEZ • *Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA*
- WILLIAM M. GRADY • *Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; Department of Medicine, University of Washington School of Medicine, Seattle, WA, USA*
- JOHN F. GRIFFITH • *Southern California Costal Water Research Project Authority, Costa Mesa, CA, USA*
- ALICE GUTTERIDGE • *University College London Cancer Institute, London, UK*
- MICHAEL S. HANEY • *Department of Psychiatry and Behavioral Sciences, Stanford Center for Genomics and Personalized Medicine, Stanford University School of Medicine, Palo Alto, CA, USA; Program on Genetics of Brain Function, Department of Genetics, Stanford Center for Genomics and Personalized Medicine, Palo Alto, CA, USA*
- CHARLES HAYNES • *Michael Smith Laboratories, Department of Chemical and Biological Engineering, University of British Columbia, Vancouver, BC, Canada; Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada*
- TAI J. HEINZERLING • *Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA*
- NICHOLAS J. HEREDIA • *Digital Biology Center, Bio-Rad Laboratories, Pleasanton, CA, USA*
- ARNE HOLST-JENSEN • *Section of Virology, Norwegian Veterinary Institute, Oslo, Norway*
- STEVEN JACOBSON • *Neuroimmunology Branch, Viral Immunology Section, National Institute of Neurological Disorders and Stroke (NINDS), National Institutes of Health, Bethesda, MD, USA*
- HANLEE P. JI • *Division of Oncology, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA; Stanford Genome Technology Center, Stanford University, Palo Alto, CA, USA*
- NOLAN KAMITAKI • *Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA; Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA; Department of Genetics, Harvard Medical School, Boston, MA, USA*
- GEORGE KARLIN-NEUMANN • *Digital Biology Center, Bio-Rad Laboratories, Pleasanton, CA, USA*
- ELENA KINZ • *Vorarlberg Institute for Vascular Investigation and Treatment (VIVIT), Feldkirch, Austria*
- ALEXANDRA BOGOŽALEC KOŠIR • *Department of Biotechnology and Systems Biology, National Institute of Biology, Ljubljana, Slovenia; Jožef Stefan International Postgraduate School, Ljubljana, Slovenia*
- YANAN KUANG • *Belfer Center for Applied Cancer Science, Dana-Farber Cancer Institute, Boston, MA, USA*
- MARCO LADETTO • *Division of Hematology, A.O. SS Antonio e Biagio e Cesare Arrigo, Alessandria, Italy*

- EMILY C. LEIBOVITCH • *Neuroimmunology Branch, Viral Immunology Section, National Institute of Neurological Disorders and Stroke (NINDS), National Institutes of Health, Bethesda, MD, USA; Institute for Biomedical Sciences, School of Medicine and Health Sciences, Washington, DC, USA*
- CLAUDIA LITTERST • *Digital Biology Center, Bio-Rad Laboratories, Pleasanton, CA, USA*
- ANDREW T. LUDLOW • *Department of Cell Biology, UT Southwestern Medical Center, Dallas, TX, USA*
- STEVEN J. MAYERL • *Gladstone Institute of Cardiovascular Disease, San Francisco, CA, USA*
- STEVEN A. MCCARROLL • *Department of Genetics, Harvard Medical School, Boston, MA, USA; Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA; Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA*
- ARIELLE J. MEDFORD • *Johns Hopkins School of Medicine, Baltimore, MD, USA; Massachusetts General Hospital, Boston, MA, USA*
- YUICHIRO MIYAOKA • *Gladstone Institute of Cardiovascular Disease, San Francisco, CA, USA; Regenerative Medicine Project, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan*
- RYAN D. MORIN • *Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, BC, Canada*
- DANY MORISSET • *Department of Biotechnology and Systems Biology, National Institute of Biology, Ljubljana, Slovenia*
- AXEL MUENDLEIN • *Vorarlberg Institute for Vascular Investigation and Treatment (VIVIT), Feldkirch, Austria*
- MASSIMO NEGRINI • *Laboratory for Technologies of Advanced Therapies (LTТА), Department of Morphology, Surgery and Experimental Medicine, University of Ferrara, Ferrara, Italy*
- ALLISON O'CONNELL • *Belfer Center for Applied Cancer Science, Dana-Farber Cancer Institute, Boston, MA, USA*
- MICHAEL OELLERICH • *Institute for Clinical Pharmacology, University Medical Center Göttingen, Göttingen, Germany*
- ELEONOR OLSSON • *Translational Oncogenomics Unit, Division of Oncology and Pathology, Department of Clinical Sciences, Lund University, Lund, Sweden*
- ERIC OUELLET • *Michael Smith Laboratories, Department of Chemical and Biological Engineering, University of British Columbia, Vancouver, BC, Canada; Centre for Blood Research, University of British Columbia, Vancouver, BC, Canada*
- GEOFFREY R. OXNARD • *Lowe Center for Thoracic Oncology, Dana-Farber Cancer Institute, Boston, MA, USA; Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA*
- BEN HO PARK • *Department of Oncology, Johns Hopkins University, Baltimore, MD, USA; Sidney Kimmel Comprehensive Cancer Institute, Johns Hopkins University, Baltimore, MD, USA*
- REENAL PATTNI • *Department of Psychiatry and Behavioral Sciences, Stanford Center for Genomics and Personalized Medicine, Stanford University School of Medicine, Palo Alto, CA, USA; Program on Genetics of Brain Function, Department of Genetics, Stanford Center for Genomics and Personalized Medicine, Palo Alto, CA, USA*
- CLOUD P. PAWELETZ • *Belfer Center for Applied Cancer Science, Dana-Farber Cancer Institute, Boston, MA, USA*
- LEONARDO PINHEIRO • *National Measurement Institute, Lindfield, NSW, Australia*

- PETAR PODLENIY • *Neurobiology Unit, Institut d'Investigacions Biomèdiques de Barcelona, Consejo Superior de Investigaciones Científicas (CSIC), Barcelona, Spain; Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Barcelona, Spain*
- MEREDITH R. RAITH • *Southern California Coastal Water Research Project Authority, Costa Mesa, CA, USA*
- JOHN REGAN • *Digital Biology Center, Bio-Rad Laboratories, Pleasanton, CA, USA*
- LAO H. SAAL • *Translational Oncogenomics Unit, Division of Oncology and Pathology, Department of Clinical Sciences, Lund University, Lund, Sweden*
- ADRIAN G. SACHER • *Lowe Center for Thoracic Oncology, Dana-Farber Cancer Institute, Boston, MA, USA; Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA*
- EKKEHARD SCHÜTZ • *Chronix Biomedical GmbH, Göttingen, Germany*
- JERRY W. SHAY • *Department of Cell Biology, UT Southwestern Medical Center, Dallas, TX, USA*
- DAWNE SHELTON • *Digital Biology Center, Bio-Rad Laboratories, Pleasanton, CA, USA*
- BJØRN SPILSBERG • *Section of Virology, Norwegian Veterinary Institute, Oslo, Norway*
- DEJAN ŠTEBIH • *Department of Biotechnology and Systems Biology, National Institute of Biology, Ljubljana, Slovenia*
- BING SUN • *Georgetown University, Washington, DC, USA*
- MUNEESH TEWARI • *Comprehensive Cancer Center, University of Michigan, Ann Arbor, MI, USA; Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA; Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI, USA; Biointerfaces Institute, University of Michigan, Ann Arbor, MI, USA; Center for Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA*
- RAMON TRULLAS • *Neurobiology Unit, Institut d'Investigacions Biomèdiques de Barcelona, Consejo Superior de Investigaciones Científicas (CSIC), Barcelona, Spain; Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Barcelona, Spain; Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain*
- NICHOLAS C. TURNER • *Breast Cancer Now Research Centre, The Institute of Cancer Research, London, UK; Breast Unit, The Royal Marsden Hospital, London, UK*
- SVILEN TZONEV • *Digital Biology Center, Bio-Rad Laboratories, Pleasanton, CA, USA*
- ALEXANDER E. URBAN • *Department of Psychiatry and Behavioral Sciences, Stanford Center for Genomics and Personalized Medicine, Stanford University School of Medicine, Palo Alto, CA, USA; Program on Genetics of Brain Function, Department of Genetics, Stanford Center for Genomics and Personalized Medicine, Palo Alto, CA, USA*
- CHRISTINA L. USHER • *Department of Genetics, Harvard Medical School, Boston, MA, USA; Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA; Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA*
- ASHLEY VELLUCCI • *Neuroimmunology Branch, Viral Immunology Section, National Institute of Neurological Disorders and Stroke (NINDS), National Institutes of Health, Bethesda, MD, USA*
- ALEXANDRA S. WHALE • *Molecular and Cell Biology Team, LGC, Teddington, UK*

- CHRISTINA M. WOOD-BOUWENS • *Division of Oncology, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA; Stanford Genome Technology Center, Stanford University, Palo Alto, CA, USA*
- WOODRING E. WRIGHT • *Department of Cell Biology, UT Southwestern Medical Center, Dallas, TX, USA*
- MING YU • *Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA*
- JANA ŽEL • *Department of Biotechnology and Systems Biology, National Institute of Biology, Ljubljana, Slovenia*
- BIN ZHANG • *Digital Biology Center, Bio-Rad Laboratories, Pleasanton, CA, USA*
- YUN-LING ZHENG • *Georgetown University, Washington, DC, USA; Cancer Prevention and Control Program, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, DC, USA*
- BO ZHOU • *Department of Psychiatry and Behavioral Sciences, Stanford Center for Genomics and Personalized Medicine, Stanford University School of Medicine, Palo Alto, CA, USA; Program on Genetics of Brain Function, Department of Genetics, Stanford Center for Genomics and Personalized Medicine, Palo Alto, CA, USA*
- XIAOWEI ZHU • *Department of Psychiatry and Behavioral Sciences, Stanford Center for Genomics and Personalized Medicine, Stanford University School of Medicine, Palo Alto, CA, USA; Program on Genetics of Brain Function, Department of Genetics, Stanford Center for Genomics and Personalized Medicine, Palo Alto, CA, USA*

# **Part I**

## **Background Concepts**



# Chapter 1

## Entering the Pantheon of 21<sup>st</sup> Century Molecular Biology Tools: A Perspective on Digital PCR

George Karlin-Neumann and Francisco Bizouarn

### Abstract

After several decades of relatively modest use, in the last several years digital PCR (dPCR) has grown to become the new gold standard for nucleic acid quantification. This coincides with the commercial availability of scalable, affordable, and reproducible droplet-based dPCR platforms in the past five years and has led to its rapid dissemination into diverse research fields and testing applications. Among these, it has been adopted most vigorously into clinical oncology where it is beginning to be used for plasma genotyping in cancer patients undergoing treatment. Additionally, innovation across the scientific community has extended the benefits of reaction partitioning beyond DNA and RNA quantification alone, and demonstrated its usefulness in evaluating DNA size and integrity, the physical linkage of colocalized markers, levels of enzyme activity and specific cation concentrations in a sample, and more. As dPCR technology gains in popularity and breadth, its power and simplicity can often be taken for granted; thus, the reader is reminded that due diligence must be exercised in order to make claims not only of precision but also of accuracy in their measurements.

**Key words** Digital PCR, dPCR, Droplet digital PCR, ddPCR, qPCR, Real-time PCR, Reproducibility, Precision, Accuracy, Rare mutation detection, Copy number variation, DNA, RNA, Absolute quantification, Control materials, Partitioning, Linkage

---

### 1 Introduction

After a 25 year gestation period, digital PCR (dPCR) has finally emerged as an indispensable tool in the research world and is headed toward a pivotal role in the clinic [1–7]. It joins a host of other technologies developed over the past 4 decades—such as Sanger sequencing, DNA and protein gel electrophoresis, real-time PCR, DNA microarrays, and next-generation sequencing (NGS)—which continue to help drive forward much of the progress of current molecular biology research and testing. Each new technology brings to the table particular strengths and limitations.

What had not been adequately served before the explosive adoption of digital PCR in the last several years was the need for more precise, specific and reproducible quantification of DNA and

RNA from diverse biological sample types. Although its earlier maturing cousin, real-time quantitative PCR (qPCR), has also enabled the sensitive detection and quantification of nucleic acids, it is more suited to relative than to absolute quantification, does not readily detect rare single-nucleotide mutations, and is more sensitive to factors affecting amplification efficiency such as assay design and PCR inhibitors. These factors affecting its robustness and interlab reproducibility have combined to slow its advance into clinical use [8–10]. Conversely, the endpoint reactions occurring in digital PCR partitions (be they prefabricated microchambers or de novo formed droplets) enable the robust amplification and absolute digital quantification of target sequences without the need for a standard curve—critical to measurements in environmental, clinical, and other sample types [11–14]; permit the highly specific detection of rare mutations in the midst of a large background of wildtype sequence—crucial to the evaluation of heterogeneous tumor samples and liquid biopsy specimens [1, 15–17], viruses [18–20], and somatic mosaicism [21]; and provide the high precision needed to distinguish copy number alleles dispersed throughout the human genome—key to the understanding of their role in genome structure and disease [22, 23].

For these and other reasons, digital PCR has become the new gold standard for nucleic acid quantification to which other methods compare themselves [24–27]. The methods in this volume have been selected to illustrate its application to a range of experimental questions organized by commonality of purpose. After an introductory part on “Background Concepts” in digital PCR (more on that below), examples are given in Part II of its adoption for “Absolute Quantification” of GMO contamination in foodstocks (Dobnik et al., Chapter 5); of viral variants of herpesvirus in different clinical sample types (Vellucci et al., Chapter 6); of mitochondrial DNA levels in the CSF as a biomarker for dementia patients (Podlesniy and Trullas, Chapter 7); and in water-quality testing for fecal contamination (Cao et al., Chapter 8).

Limitations in the achievable precision of gene amplification measurements have hampered investigations of their role in various diseases and in cancer [28–30]. Part III offers protocols for studying “Copy Number Variation.” Davis et al. (Chapter 9) explain how assays can be designed and evaluated for the study of copy number alleles and their potential association with phenotypes of interest. Combining the challenges of liquid biopsy in breast cancer recurrence with evaluation of copy number variation, Garcia-Murillas and Turner (Chapter 10) describe how to assess whether amplification of Her2 is likely causal to disease progression. Zhou and colleagues (Chapter 11) illustrate the power of dPCR for studying somatic mosaicism of amplified genomic sequences, SNPs, and transposable elements in primary human tissue and iPSCs, a

phenomenon that was little recognized before the availability of high precision dPCR technology.

Part IV (“Rare Mutation/Allele Detection”) reflects the intensity and inventiveness with which digital PCR has been applied to the study of rare mutations or foreign alleles accessible in the blood, in both plasma cell-free DNA (cfDNA) and in cells. Kuang et al. (Chapter 12) illustrate the use of dPCR for detection and monitoring of both sensitizing and resistance mutations in liquid biopsy of non-small-cell lung cancer, while Medford and colleagues (Chapter 13) describe a method for preamplification of cfDNA from early-stage breast cancer patients to assess whether residual disease is still present. To assess minimal residual disease in multiple lymphoid malignancies, Drandi and her team (Chapter 14) present methods for both the identification of patient-specific rearrangements and their conversion to droplet digital™ PCR (ddPCR™) assays for serial blood monitoring. The workflow for use of the QuantStudio™ 3D dPCR system is described by Kinz and Muenklein (Chapter 15) for measuring JAK2 V617F allele burden, prevalent in myeloproliferative diseases. Alcaide and Morin (Chapter 16) tackle the common challenge in liquid biopsy of having limited cfDNA from a patient sample and wanting maximal sensitivity for multiple mutations. They describe a suite of multiplex assay design strategies for achieving this where they minimize the number of probes used and hence cost required. In some design strategies they screen for different mutations but do not distinguish among them, in others, they do discriminate between the specific mutation(s) that may be present. Chen and coworkers (Chapter 17) detail an approach to achieve early detection of occult metastases in breast cancer by monitoring patient-specific rearrangements with ddPCR fusion assays in serial bloods. They derive these trunk mutations from NGS sequencing of tumor tissue and convert them to digital PCR assays for highly sensitive monitoring of potential recurrence.

In keeping with the desire for flexibility, speed, and low cost for detection and monitoring of CNVs and SNPs, Wood and Ji (Chapter 18) demonstrate how to multiplex with the probe-less EvaGreen dye-binding chemistry in droplet digital PCR. This method requires only the use of specially designed allele-specific primers and entirely avoids the cost and time needed for probe synthesis. In a non-cancer application, Beck and colleagues (Chapter 19) illustrate the use of the liquid biopsy paradigm for the monitoring of transplant health. Treating the transplanted organ as a “genome transplant,” they identify which of a preselected set of high minor allele-frequency SNPs can be used to quantify turnover of the donor tissue within the recipient, and thus its state of health. In yet another variation of digital PCR for the detection and quantification of rare alleles, Miyaoka and his coworkers (Chapter 20) describe the use of ddPCR for monitoring

of gene editing experiments where they are able to simultaneously quantify the frequency of both homologous (HDR) and nonhomologous (NHEJ) repair events. Part IV concludes with an example of epigenetic analysis where Yu et al. describe MethylLight ddPCR (Chapter 21), the adaptation of bisulfite methylation analysis to dPCR to enable the study of “field cancerization” in solid tumors.

Gene expression and RNA quantification studies (Part V) have also benefited from the increased precision and absolute quantification of dPCR. In Chapter 22, Sun and Zheng address the need for better understanding of the role of alternative splicing in the regulation of telomerase activity and function. They illustrate a general method for simultaneously measuring the levels of the four major splice variants of the human telomerase gene in a single droplet digital PCR reaction. Kamitaki and colleagues also leverage the high precision of dPCR in Chapter 23 to enable measurement of allele-specific expression differences between two alleles in an individual differing by only a single SNP. These subtle SNP effects on expression may explain the role of these variants on phenotypes with which they have been associated in genome-wide association studies. In Chapter 24, Karlin-Neumann et al. illustrate the application of ddPCR to the study of single cell expression, especially useful where quantifying very low abundance transcripts is important. To increase the power of the method, they present a strategy for probe-multiplexing (or “radial” multiplexing) which enables them to quantify five different targets within a single PCR reaction, or by splitting the cDNA from each cell into halves, as many as nine or ten genes per cell rapidly and with absolute quantification. The two remaining chapters in this part provide guidance in quantifying miRNAs from serum or plasma using either EvaGreen probeless chemistry (Ferracin and Negrini, Chapter 25) or probe-based Taqman chemistry (Giraldez et al., Chapter 26).

Although absolute quantification of DNA and RNA is the predominant use of digital PCR systems, Part VI (“Other Uses of Partitioning”) reveals some less anticipated uses for which partitioning can be deployed. Besides the quantification of nucleic acids, partitioning can also provide information on DNA fragment size as illustrated by Heredia et al. (Chapter 27) for characterization of NGS library preparations. Regan and colleague (Chapter 28) demonstrate how the physical linkage of two genetic markers can be tested by whether or not assay signals for each are colocalized in partitions more often than would be expected by chance alone. They also describe a simple method for preparation of large DNA satisfactory for testing linkage over as much as 150 or even 200Kb. The detection and quantification of telomerase activity by Ludlow et al. (Chapter 29) reveals a more general paradigm for translating non-DNA measurements (e.g., enzyme activities that can act on nucleic acid templates) into a DNA readout which can be quantified via digital PCR. Finally, the act of partitioning a reaction into many

smaller reactions can greatly improve the signal-to-noise ratio over bulk reactions and has been innovatively leveraged by Ouillet and colleagues (Chapter 30) to develop a highly efficient droplet-based SELEX method (Hi-Fi SELEX) which accomplishes in 3 or 4 rounds of selection over several weeks what normally requires ~8 rounds and several months, and achieves the identification of high affinity aptamers.

Although not included in this volume due to their more recent publication, additional clever applications have been developed by the world's dPCR user community. Among these is the simultaneous quantification of a protein and its corresponding mRNA level in single mammalian cells jointly using a digital proximity ligation assay and an RT-ddPCR reaction on a crude cell lysate [31]. Perhaps even more unexpected, Cheng et al. [32] have been able to quantify picomolar concentrations of Ag(I) and Hg(II) through their dose-dependent effects in permitting extension of mismatched primers annealed to their nearly complementary DNA templates! Again, the presence and amount of an initial non-DNA analyte is translated into a quantifiable DNA signal. So it is highly probable that even more unexpected applications will develop from the fundamental ability to partition samples.

Returning to the opening Part I on “Background Concepts,” it is important to remember that despite the seductive appeal of rendering absolute measurements of the numbers of target copies in a sample, it is still necessary to be critical of our experimental methods before drawing conclusions from their results, whether for research or testing purposes. The quantitative nature of digital PCR should not make us complacent about the non-uniform structure of genomic targets, i.e., there are regions and contexts of genomic sequences that can mask their accessibility to PCR reactions resulting in greatly under-quantified measurements, presumably due to secondary structure. This may be mitigated by alternative assay designs, reducing the size of the genomic target fragments or modified reaction conditions. That said, the onus is on the experimenter to validate the accuracy of their results and not to confuse precision and reproducibility with truth.

Similarly, the beauty and power of digital PCR should not allow us to forget that there are other hurdles that lie between our experimental results and the questions we seek to answer. As an example, Whale and colleagues (Chapter 4) describe in Part I the use of sample extraction controls that are important when trying to calculate back from a dPCR measurement to the number of tumor DNA sequences present in a milliliter of blood. Similarly, Pinheiro and Emslie (Chapter 2) explain the concept of digital PCR and what factors influence the final measurement obtained so that we can perform these experiments with appropriate vigilance and with suitable run controls. Lastly, in applications where detection of rare species is common, it is crucial to understand the limit of detection

of our assay and whether or not we are confident we have detected a species of interest. Tzonev (Chapter 3) provides us a comprehensive overview of test performance descriptors such as sensitivity and specificity, sources of variability and uncertainty in molecular counting and how a measured false positive rate influences the limit of detection achievable.

The active ferment of dPCR investigations in thousands of labs throughout the world is also evidenced by the increasing number of dPCR system implementations that are being devised in academic labs (e.g., SlipChip, ref. 33) and also making their way to commercial availability. At the time we began assembling this collection of methods, we aimed to include representative peer-reviewed procedures using all of the commercially available dPCR systems. At that time, this comprised Fluidigm's Biomark, Bio-Rad's QX100/200, RainDance's RainDrop, and Thermo-Fisher's QuantStudio 3D systems, although the majority of published studies were heavily weighted toward the QX100/200 platform and thus most of the contributions in this volume were based on it. Though we were unable to obtain contributions from users of either the RainDance or Fluidigm systems, in part due to the smaller user base of the former and the diminished use of the latter, we trust that the overall strategies described in these chapters will be generally applicable to various dPCR platforms present and future. In that vein, since commencement of this project, three new systems have recently become commercially available from Formulatrix (the Constellation™) and JN Medsys (the Clarity™), both with fixed chambers, and from Stillia (the Naica™, with droplets). Suffice it to say, dPCR research and testing is growing at an ever-expanding rate with an increasing variety of implementations, and promises to become an enduring and incisive tool in the arsenal of the molecular biologist.

## References

- Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC, Kitano TK, Hodel MR, Petersen JF, Wyatt PW, Steenblock ER, Shah PH, Bousse LJ, Troup CB, Mellen JC, Wittmann DK, Erndt NG, Cauley TH, Koehler RT, So AP, Dube S, Rose KA, Montesclaros L, Wang S, Stumbo DP, Hodges SP, Romine S, Milanovich FP, White HE, Regan JF, Karlin-Neumann GA, Hindson CM, Saxonov S, Colston BW (2011) High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. *Anal Chem* 83(22):8604–8610. <https://doi.org/10.1021/ac202028g>
- Baker M (2012) Digital PCR hits its stride. *Nat Methods* 9(6):541–544. <https://doi.org/10.1038/nmeth.2027>
- Miliaras N (2014) Digital PCR comes of age. *Genetic Engineering & Biotechnology News* 34(4):14–16. <https://doi.org/10.1089/gen.34.04.08>
- Perkel J (2014) The digital PCR revolution. *Science* 344(6180):212–214. <https://doi.org/10.1126/science.344.6180.212>
- Pinheiro & Emslie, Chapter 2, this MIMB volume
- Sacher AG, Pawletz C, Dahlberg SE, Alden RS, O'Connell A, Feeney N, Mach SL, Jänne PA, Oxnard GR (2016) Prospective validation of rapid plasma genotyping for the detection of EGFR and KRAS mutations in advanced lung cancer. *JAMA Oncol* 2(8):1014–1022. <https://doi.org/10.1001/jamaoncol.2016.0173>

7. Mellert H, Foreman T, Jackson L, Maar D, Thurston S, Koch K, Weaver A, Cooper S, Dupuis N, Sathyaranayana UG, Greer J, Hahn W, Shelton D, Stonemetz P, Pestano GA (2017) Development and clinical utility of a blood-based test service for the rapid identification of actionable mutations in non-small cell lung carcinoma. *J Mol Diagn* 19(3):404–416. <https://doi.org/10.1016/j.jmoldx.2016.11.004>
8. Whale AS, Huggett JF, Cowen S, Speirs V, Shaw J, Ellison S, Foy CA, Scott DJ (2012) Comparison of microfluidic digital PCR and conventional quantitative PCR for measuring copy number variation. *Nucleic Acids Res* 40(11):e82. <https://doi.org/10.1093/nar/gks203>
9. Hindson CM, Chevillet JR, Briggs HA, Galli- chotte EN, Ruf IK, Hindson BJ, Vessella RL, Tewari M (2013) Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat Methods* 10(10):1003–1005. <https://doi.org/10.1038/nmeth.2633>
10. Whale AS, Devonshire AS, Karlin-Neumann G, Regan J, Javier L, Cowen S, Fernandez- Gonzalez A, Jones G, Redshaw N, Beck J, Berger A, Combaret V, Kjersgaard ND, Davis L, Fina F, Forshew F, Andersen RF, Galbiati S, Hernández AG, Haynes CA, Janku F, Lacave R, Lee J, Mistry V, Pender A, Pradines A, Proudhon C, Saal L, Stieglitz E, Ulrich B, Foy CA, Parkes H, Tzonev S, Huggett JF (2017) International interlaboratory digital pcr study demonstrating high reproducibility for the measurement of a rare sequence variant. *Anal Chem* 89(3):1724–1733. <https://doi.org/10.1021/acs.analchem.6b03980>
11. Cao Y, Raith MR, Griffith JF (2015) Droplet digital PCR for simultaneous quantification of general and human-associated fecal indicators for water quality assessment. *Water Res* 70:337–349. <https://doi.org/10.1016/j.watres.2014.12.008>
12. Nadauld L, Regan JF, Miotke L, Pai RK, Longacre TA, Kwok SS, Saxonov S, Ford JM, Ji HP (2012) Quantitative and sensitive detection of cancer genome amplifications from formalin fixed paraffin embedded tumors with droplet digital PCR. *Transl Med (Sunnyvale)* 2(2):1–12. <https://doi.org/10.4172/2161-1025.1000107>
13. Belgrader P, Tanner SC, Regan JF, Koehler R, Hindson BJ, Brown AS (2013) Droplet digital PCR measurement of HER2 copy number alteration in formalin-fixed paraffin-embedded breast carcinoma tissue. *Clin Chem* 59(6):991–994. <https://doi.org/10.1373/clinchem.2012.197855>
14. Dingle TC, Sedlak RH, Cook L, Jerome KR (2013) Tolerance of droplet-digital PCR vs real-time quantitative PCR to inhibitory substances. *Clin Chem* 59(11):1670–1672. <https://doi.org/10.1373/clinchem.2013.211045>
15. Oxnard GR, Paweletz CP, Kuang Y, Mach SL, O'Connell A, Messineo MM, Luke JJ, Butaney M, Kirschmeier P, Jackman DM, Jänne PA (2014) Noninvasive detection of response and resistance in EGFR-mutant lung cancer using quantitative next-generation genotyping of cell-free plasma DNA. *Clin Cancer Res* 20(6):1698–1705. <https://doi.org/10.1158/1078-0432.CCR-13-2482>
16. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ, Cheang M, Osin P, Nerurkar A, Kozarewa GJA, Dowsett M, Reis-Filho JS, Smith IE, Turner NC (2015) Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Sci Transl Med* 7(302):302ra133. <https://doi.org/10.1126/scitranslmed.aab0021>
17. Janku F, Huang HJ, Fujii T, Shelton DN, Madwani K, Fu S, Tsimberidou AM, Piha-Paul SA, Wheler JJ, Zinner RG, Naing A, Hong DS, Karp DD, Cabriolo G, Kopetz ES, Subbiah V, Luthra R, Kee BK, Eng C, Morris VK, Karlin-Neumann GA, Meric-Bernstam F (2017) Multiplex KRASG12/G13 mutation testing of unamplified cell-free DNA from the plasma of patients with advanced cancers using droplet digital polymerase chain reaction. *Ann Oncol* 28(3):642–650. <https://doi.org/10.1093/annonc/mdw670>
18. Strain MC, Lada SM, Luong T, Rought SE, Gianella S, Terry VH, Spina CA, Woelk CH, Richman DD (2013) Highly precise measurement of HIV DNA by droplet digital PCR. *PLoS One* 8(4):e55943. <https://doi.org/10.1371/journal.pone.0055943>
19. Persaud D, Gay H, Ziemiak C, Chen YH, Piatak M Jr, Chun TW, Strain M, Richman D, Luzuriaga K (2013) Absence of detectable HIV-1 viremia after treatment cessation in an infant. *N Engl J Med* 369(19):1828–1835. <https://doi.org/10.1056/NEJMoa1302976>
20. Leibovitch EC, Brunetto GS, Caruso B, Fenton K, Ohayon J, Reich DS, Jacobson S (2014) Coinfection of human herpesviruses 6A (HHV-6A) and HHV-6B as demonstrated by novel digital droplet PCR assay. *PLoS One* 9(3):e92328. <https://doi.org/10.1371/journal.pone.0092328>

21. Abyzov A, Mariani J, Palejev D, Zhang Y, Haney MS, Tomasini L, Ferrandino AF, Rosenberg Belmaker LA, Szekely A, Wilson M, Kocabas A, Calixto NE, Grigorenko EL, Huttner A, Chawarska K, Weissman S, Urban AE, Gerstein M, Vaccarino FM (2012) Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. *Nature* 492(7429):438–442. <https://doi.org/10.1038/nature11629>
22. Handsaker RE, Van Doren V, Berman JR, Genovese G, Kashin S, Boettger LM, McCarroll SA (2015) Large multiallelic copy number variations in humans. *Nat Genet* 47 (3):296–303. <https://doi.org/10.1038/ng.3200>
23. Davis JM, Searles VB, Anderson N, Keeney J, Dumas L, Sikela JM (2014) DUF1220 dosage is linearly associated with increasing severity of the three primary symptoms of autism. *PLoS Genet* 10(3):e1004241. <https://doi.org/10.1371/journal.pgen.1004241>
24. Guttery DS, Page K, Hills A, Woodley L, Marchese SD, Rghebi B, Hastings RK, Luo J, Pringle JH, Stebbing J, Coombes RC, Ali S, Shaw JA (2015) Noninvasive detection of activating estrogen receptor 1 (ESR1) mutations in estrogen receptor-positive metastatic breast cancer. *Clin Chem* 61(7):974–982. <https://doi.org/10.1373/clinchem.2015.238717>
25. Thress KS, Brant R, Carr TH, Dearden S, Jenkins S, Brown H, Hammett T, Cantarini M, Barrett JC (2015) EGFR mutation detection in ctDNA from NSCLC patient plasma: A cross-platform comparison of leading technologies to support the clinical development of AZD9291. *Lung Cancer* 90 (3):509–515. <https://doi.org/10.1016/j.lungcan.2015.10.004>
26. Gu J, Zang W, Liu B, Li L, Huang L, Li S, Rao G, Yu Y, Zhou Y (2017) Evaluation of digital PCR for detecting low-level EGFR mutations in advanced lung adenocarcinoma patients: a cross-platform comparison study. *Oncotarget*. <https://doi.org/10.18632/oncotarget.18866>. [Epub ahead of print]
27. Low H, Chan SJ, Soo GH, Ling B, Tan EL (2017) Clarity™ digital PCR system: a novel platform for absolute quantification of nucleic acids. *Anal Bioanal Chem* 409(7):1869–1875. <https://doi.org/10.1007/s00216-016-0131-7>
28. Aldhous MC, Abu Bakar S, Prescott NJ, Palla R, Soo K, Mansfield JC, Mathew CG, Satsangi J, Armour JA (2010) Measurement methods and accuracy in copy number variation: failure to replicate associations of beta-defensin copy number with Crohn's disease. *Hum Mol Genet* 19(24):4930–4938. <https://doi.org/10.1093/hmg/ddq411>
29. Liu S, Yao L, Ding D, Zhu H (2010) CCL3L1 copy number variation and susceptibility to HIV-1 infection: a meta-analysis. *PLoS One* 5 (12):e15778. <https://doi.org/10.1371/journal.pone.0015778>
30. Usher CL, Handsaker RE, Esko T, Tuke MA, Weedon MN, Hastie AR, Cao H, Moon JE, Kashin S, Fuchsberger C, Metspalu A, Pato CN, Pato MT, McCarthy MI, Boehnke M, Altshuler DM, Frayling TM, Hirschhorn JN, McCarroll SA (2015) Structural forms of the human amylase locus and their relationships to SNPs, haplotypes and obesity. *Nat Genet* 47 (8):921–925. <https://doi.org/10.1038/ng.3340>
31. Albayrak C, Jordi CA, Zechner C, Lin J, Bichsel CA, Khammash M, Tay S (2016) Digital quantification of proteins and mRNA in single mammalian cells. *Mol Cell* 61(6):914–924. <https://doi.org/10.1016/j.molcel.2016.02.030>
32. Cheng N, Zhu P, Xu Y, Huang K, Luo Y, Yang Z, Xu W (2016) High-sensitivity assay for Hg (II) and Ag (I) ion detection: A new class of droplet digital PCR logic gates for an intelligent DNA calculator. *Biosens Bioelectron* 84:1–6. <https://doi.org/10.1016/j.bios.2016.04.084>
33. Shen F (2017) SlipChip device for digital nucleic acid amplification. *Methods Mol Biol* 1547:123–132. [https://doi.org/10.1007/978-1-4939-6734-6\\_10](https://doi.org/10.1007/978-1-4939-6734-6_10)



# Chapter 2

## Basic Concepts and Validation of Digital PCR Measurements

Leonardo Pinheiro and Kerry R. Emslie

### Abstract

Use of digital polymerase chain reaction (dPCR) technology is rapidly growing and diversifying into a range of areas in life science. The release of dPCR commercial systems has facilitated access, leading to recognition of the potential advantages compared to previous quantitative PCR technologies, and the scope for novel applications. The capability of dPCR to deliver unprecedented levels of precision, accuracy, and resolution in quantification of nucleic acids has triggered a strong interest by academia and the life sciences industry in use of this technology as a molecular diagnostic tool. However, the performance of dPCR, as for a “classical” PCR assay, essentially still relies on enzyme-based amplification of nucleic acid using specific reagents and instrumentation. This chapter describes basic concepts, key properties, and important factors to consider for the verification and validation of dPCR measurements.

**Key words** Digital PCR, Nucleic acids quantification, Validation

---

### 1 Introduction

Use of dPCR technology is undoubtedly on the rise considering the rapid increase in number of peer-reviewed publications and publicity over the past few years. A Pubmed search (<http://www.ncbi.nlm.nih.gov/pubmed>) of the words “digital PCR” identified about 20 publications in 2012, while for the period between January 2013 and October 2015 there were more than 300 publications. The main drivers behind increased use of dPCR have been development of commercial systems offering reduced costs and more flexible workflow, and recognition by the broader scientific community of the potential advantages of the technology and the wide scope of its application.

Recent developments in application of dPCR technology have been far reaching and diverse, including detection of various pathogens [1–4], monitoring food safety [5, 6] and water quality [7, 8], and studying microbial ecology [9]. However, due to the high level of precision afforded by dPCR and consequent ability to detect small differences in amount of nucleic acid molecules, it has mostly

been used in clinical research including development and detection of biomarkers [10–12], detection of genome amplification status and genetic variants in cancer patients [13–17], genetic screening [18], genome editing [19, 20], detection of single nucleotide changes and copy number alterations in stem cells [21, 22], and monitoring graft-derived DNA and chimeras in transplant recipients [23, 24]. As a result, leading research groups and the life sciences industry are developing dPCR technology based tests with the aim of utilizing these tests for molecular diagnosis. For translation of a dPCR-based test from research laboratory to clinical laboratory setting, the test must satisfy requirements of regulatory authorities which include validation and/or verification studies [25].

For over 20 years, real-time quantitative PCR (qPCR) [26, 27] has been the main technology used for quantitative measurement of specific targets within nucleic acid samples. However, despite thousands of publications reporting on clinical applications for qPCR and a number of PCR tests cleared by regulatory authorities, we are yet to see widespread use of routine quantitative nucleic acids measurement in clinical molecular diagnostics. One reason for this limited uptake of qPCR in a diagnostic setting may be the difficulty encountered in reproducing qPCR results reliably. qPCR is an analog measurement based on monitoring exponential increase in fluorescence signal after each PCR cycle. The point at which the fluorescence signal crosses an intensity threshold is called the quantification or threshold cycle (C<sub>q</sub>). Target nucleic acid concentration in a sample is determined using a calibration curve which fits the logarithm of nucleic acid concentration for a dilution series of calibrant to the C<sub>q</sub> value for each calibrant dilution. To ensure qPCR data is reproducible over time and between laboratories, the calibrant should demonstrate appropriate stability and homogeneity and be traceable back to an international or higher-order DNA reference material. However, these DNA reference materials are only available for a limited number of target sequences. In addition, accurate qPCR measurements rely on good amplification efficiency for both calibrant and sample, with minimal effects from sample matrix and inhibitory substances. In contrast, dPCR does not require a calibration curve and its linear digital signal read out has the potential to provide more precise data than that obtained from the exponential amplification signal of qPCR. These key factors give dPCR a technical advantage over qPCR.

The dPCR workflow involves random distribution of PCR mix containing target nucleic acid, primers, probe, and mastermix across a large number of uniformly sized partitions, such that some partitions contain no nucleic acid template and others contain one or more template copies. Following PCR amplification partitions containing target DNA can be distinguished from partitions

containing no target DNA through their fluorescence signal. The target nucleic acid copy number per partition is derived based on the number of positive partitions and the total number of partitions using a Poisson model.

As for qPCR, dPCR essentially relies on enzyme-based amplification of nucleic acid using specific reagents and instrumentation. In the absence of assay optimization and validation, either reduced assay performance or complete failure of an assay is possible. For dPCR to realize its full potential as a quantitative molecular diagnostic tool, carefully designed validation studies are necessary and sufficient understanding of sources of variation and bias that can lead to incorrect results or poor reproducibility of data is needed. This chapter aims to describe basic concepts, key properties of dPCR and to highlight important factors to consider for verification and validation of dPCR measurements.

---

## 2 Development of dPCR

### 2.1 Limiting Dilution—Key to the Development of dPCR

Limiting dilution consists of diluting a sample to such a degree that, when divided among several individual assays, the analyte of interest will be present in some but not all of the assays. The principle behind limiting dilution originated from work developed during the first-world war for the enumeration of bacterial growth for infection management purpose [28]. Limiting dilution followed by PCR was first used to quantify the number of human immunodeficiency proviral molecules [29]. Soon after, a similar methodology was described for quantification of leukaemic cells [30]. The first study describing the general applicability of limiting dilution for quantification of DNA target molecules was published in 1992 [31]. These initial studies used a small number of standard PCR replicates (less than 30) to generate results using the limiting dilution principle. At that time, development and commercialization of qPCR technology [26, 27] offered a practical and easy option of PCR based nucleic acid quantification and discouraged further development of limiting dilution to quantify nucleic acids. In 1999, limiting dilution was applied across a much larger number of replicates using a commercially available 384-well plate format to quantify disease-associated mutations in colorectal cancer patients. The authors then coined the term “digital PCR” [32]. However, the method was still too cumbersome to compete with qPCR and was not widely adopted at the time. By 2002, advances in microfluidics engineering allowed development of microfluidic chips amenable to rapid, simple, and scalable fabrication [33]. This turned dPCR into a relatively simple and practical method of accurate nucleic acid quantification and the first publications using commercial microfluidic devices describing the potential of dPCR technology started to appear [34–37].

## 2.2 Commercial dPCR Systems

In 2006, Fluidigm Corporation launched a commercial dPCR platform (Biomark™) based on the company's trademark integrated fluidics technology using array chips comprising 12 panels of 765 partitions. In 2009, it released a second chip comprising 48 panels of 770 partitions which increased throughput and decreased costs of using the technology. Also in 2009, Life Technologies launched the OpenArray® system which utilizes small plates perforated with over 3000 micron-size through-holes that retain PCR mix through surface tension. In 2013, they introduced high density nanofluidic chips carrying up to 20,000 data points with simplified workflow and reduced cost. In 2014, Formulatrix launched the Constellation system which utilizes a proprietary 96-well microfluidic system with each well containing 496 chambers where individual PCR assays take place.

The launch of droplet-based dPCR commercial systems led to more widespread uptake of dPCR technology. The workflow of droplet dPCR involves generation of a water-in-oil emulsion in which aqueous droplets comprise the PCR mix. The fine-tuned titration of PCR and oil chemistry can produce thousands or millions of nanoliter or picoliter size uniform droplets which can withstand fast flow through microfluidic channels and thermal cycling without bursting or coalescing. By the end of 2011 Bio-Rad Laboratories released the QX100 droplet dPCR system which utilizes probe based chemistry for detection. In the following year the QX200 droplet dPCR system was launched which utilizes both probe based and intercalating dye chemistries for detection. Both systems generate approximately 20,000 nanoliter-sized droplets from each PCR mix at a lower cost compared to early array chip-based systems. Also in 2012, RainDance Technologies launched the RainDrop® system, a dPCR platform which generates millions of picoliter-sized droplets from each PCR mix.

The various dPCR platforms can differ in workflow, partition number, partition volume and sample throughput. In addition, the proportion of the PCR mix which is analyzed by dPCR varies between platforms. In some systems, a portion of the PCR mix is retained in the microfluidics and not partitioned. Partitions may also be excluded from data analysis if they fail to meet predefined specifications. These properties may have particular advantages and disadvantages depending on the application. Before making a decision on purchase of a dPCR instrument, consideration should be given to the required sensitivity, precision, dynamic range and throughput for the intended application, as well as per sample operating costs.

---

### 3 Basic Concepts of dPCR

#### 3.1 The Workflow

The workflow of dPCR is generally fairly simple. In chip-based systems, PCR mix is evenly distributed across prefabricated partitions normally with the aid of loading instruments. The loaded chip is placed onto a thermal cycler for PCR amplification. After thermal cycling, partitions containing target molecules (positive partitions) are distinguished from partitions with no target molecules (negative partitions) by imaging fluorescence intensity for each partition in the chip using a camera. Based on a normalized fluorescence intensity signal, a threshold is applied to assign partitions as either positive or negative. Some chip-based systems also collect real-time data, thus providing a Cq value for each partition. In droplet-based systems, partitions are generated dynamically by the user: PCR mix and oil are typically pipetted into wells of a cartridge and then the cartridge is transferred to specifically designed instruments in which a water-in-oil emulsion is generated. The droplet emulsion is transferred to either a well in a 96 well plate or a tube in a strip of tubes and is then placed into a thermal cycler for PCR amplification. After thermal cycling, the plate or tubes are transferred to a droplet reading instrument where droplets are aspirated and fast-flow streamed past a fluorescence detector, in a manner similar to a flow cytometer, for measurement of fluorescence in each droplet. Based on the fluorescence signal, a threshold is applied to assign droplets to either the positive or negative droplets population.

#### 3.2 Estimating Copy Number

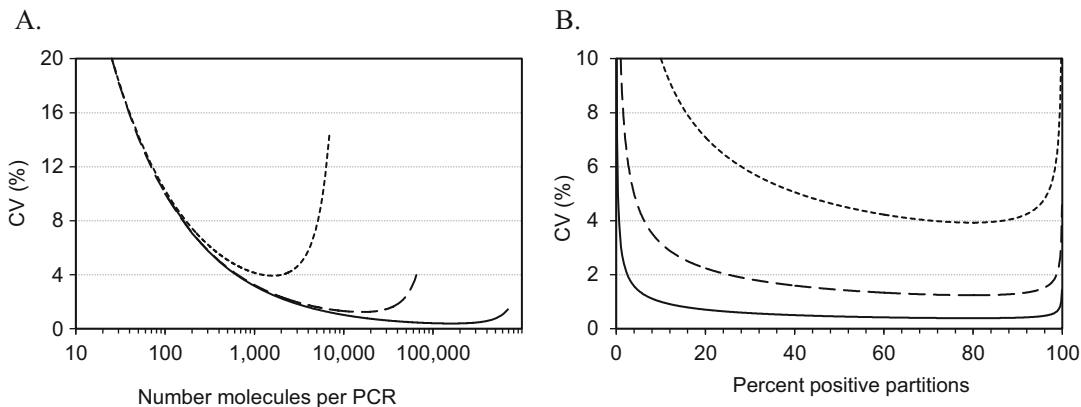
Poisson statistics, which take into account the probability of any given partition containing zero, one or more than one molecule, are used to estimate the average number of molecules per partition [36]. If the average partition volume,  $V_p$ , is known, Poisson statistics can also be used to estimate the concentration of target nucleic acid molecules in the PCR mix using Eq. 1.

$$[\text{Target DNA}] = -\ln \left( 1 - \frac{N_p}{N_T} \right) \frac{1}{V_p} \text{ copy number per unit volume} \quad (1)$$

where  $[\text{Target DNA}]$  is the number of target DNA molecules per unit volume of PCR mix,  $N_p$  is the number of positive partitions, and  $N_T$  is the total number of partitions. Each component in the equation needs to be measured accurately to avoid compromising the accuracy of the nucleic acid concentration estimate.

#### 3.3 Dynamic Range and Precision

The theoretical dynamic range of dPCR is limited by the number of partitions analyzed (Fig. 1). Because the dPCR partitioning process follows a binomial distribution, the dynamic range extends beyond the number of partitions analyzed [36] although precision is compromised at both low and high ends of the dynamic range



**Fig. 1** Variability arising from subsampling and partitioning of sample in relation to (a) the number of target molecules and (b) the percent of positive partitions for digital PCR assays containing 1000 (.....), 10,000 (- - -), and 100,000 (—) partitions. CV, coefficient of variation

[37, 38]. The uniformity of partition size can also impact on the dynamic range and precision of dPCR. Monodisperse droplets are required for the highest precision. Methods that use polydisperse partition sizes will have lower precision since larger droplets will have a higher likelihood of containing multiple targets than smaller droplets in the same sample. This, in turn, will impact on the Poisson statistics calculation since this calculation is based on the assumption that the partitions are of equivalent volume. Assuming uniform partition size a typical 20,000 partition dPCR assay provides a theoretical dynamic range over orders of magnitude and up to  $10^5$  molecules [38]. One commercially available dPCR instrument is capable of producing millions of partitions, thus extending the dynamic range to over six orders of magnitude [39]. Assuming sufficient biological sample is available, a large dynamic range can be useful for detecting rare target molecules in a background of nontarget molecules since nontarget molecules will be distributed across a large number of partitions, thus reducing the possibility that they will interfere with detection of rare target molecules [40].

The theoretical precision of dPCR can be estimated using Poisson statistics and improves with increasing partition number, provided that analysis is undertaken within the optimal window of the dynamic range (Fig. 1). The level of precision and consequent resolving power offered by dPCR is typically better than can be achieved using qPCR [41] or other technologies such as fluorescence in situ hybridization [42], and is useful for applications such as copy number variation measurements. However, theoretical precision is not always reflected in experimental precision observed from analyzing replicates since the Poisson model is only one component in a range of factors that will combine to reflect observed experimental precision. In general, if the number of

partitions is less than 1000, the Poisson model copy number estimate will have a coefficient of variation of at least 4% and can be one of the main factors contributing to the observed imprecision, whereas if the number of partitions is more than 10,000 the coefficient of variation can be less than 1.5%, so other factors may become the dominant sources of imprecision (Fig. 1).

### **3.4 Partitioning and Inhibitory Effects**

dPCR has a lower susceptibility to inhibition than qPCR for inhibitory substances commonly found in clinical [43, 44], food [5], and environmental [7] samples. In conventional qPCR, inhibitors can reduce amplification efficiency and increase the Cq value [45] with a resultant impact on accurate quantification of the target molecule. In contrast, moderate levels of inhibitor in dPCR can often be tolerated without compromising dPCR accuracy, since accuracy relies on the ability to distinguish negative and positive partitions after end-point PCR, and this is often possible even when efficiency is slightly reduced. Also the possibility of inhibitory effects due to cross-reactivity or “cross talk” observed multiplex qPCR assays is much reduced by the partitioning process in dPCR. However, if present at high enough concentrations, inhibitory substances can still reduce amplification efficiency to the point where some partitions are misclassified, resulting in an inaccurate estimate of the original template concentration. As the susceptibility of different PCR assays to inhibition is variable and the mechanism of interference can differ between inhibitory substances [7, 46], it is important to empirically assess potential inhibitory effects by carefully designed control experiments prior to establishing dPCR methods.

### **3.5 Partitioning and Molecule Segregation**

Random distribution of target nucleic acids molecules across partitions is necessary for accurate quantification since this is a fundamental assumption of the Poisson statistics model. The number of target molecules in the sample will be underestimated following nonrandom partitioning as, for instance, when target sequences are linked or in tandem since linked copies will colocate in the same partition. Restriction digest treatment of samples prior to partitioning can be used to separate linked target sequences [47]. For that, prior knowledge of restriction enzymes that do not cut the target sequence is necessary and an evaluation of the effect of the restriction enzyme and buffer on subsequent dPCR assay performance should be undertaken.

Conversely, potential overestimation of copy number can occur if a portion of the DNA molecules are single-stranded. The presence of single strands of nucleic acid during partitioning can result in up to a twofold increase in concentration estimate as the single strand conformation can produce up to twice as many independently segregated amplifiable templates [48].

### **3.6 The Partition Volume Effect**

To calculate copy number concentration using dPCR (see Eq. 1), the average copy number per partition is divided by partition volume. Three properties of partition volume can affect accuracy of dPCR measurements: accuracy of the manufacturer-specified partition volume, variability in average partition volume between replicate dPCR measurements (interreaction variability) and variability in partition volume within a single dPCR replicate (intrareaction variability). The partition volume in both currently available chip and droplet-based dPCR technologies can vary slightly from manufacturer's stated volume [37, 38] and, if uncorrected, will bias copy number concentration estimates [49]. Both interreaction and intrareaction variability in partition volume will affect precision and, to a lesser extent, accuracy of dPCR concentration estimates [50]. Variability in partition volume will contribute to observed technical precision and is one factor which will influence how closely aligned technical precision is with the theoretical precision estimated using Poisson statistics.

The effect of partition volume variation is minimized for copy number ratio measurements between two assays conducted in a duplex dPCR format. However, partition volume variability is important for clinical applications requiring absolute quantification of nucleic acids [11, 51].

## **4 Validation of dPCR Assays**

### **4.1 Validation, Verification, and Measurement Uncertainty Definitions**

The term verification is defined as “provision of objective evidence that a given item fulfills specified requirements” while validation is defined as “verification, where the specified requirements are adequate for an intended use” [52]. In the context of dPCR measurements, verification confirms that specified performance properties of the measurement system (e.g., a dPCR assay targeting a specific gene sequence) are achieved, whereas validation establishes that the dPCR assay is suitable for the intended application.

Approval from regulatory authority generally requires a test to be validated for a specific purpose using specified equipment, reagents, and controls. If any specifications are altered or modified, it may be necessary to revalidate the modified measuring system to ensure it is still adequate for its intended use or “fit-for-purpose.” Validated performance characteristics may need to be verified at regular intervals to demonstrate that specified performance properties continue to be met, for example with different batches of reagents.

During analytical validation, it is important to identify and quantify sources of analytical variation and potential bias. The combined impact of these sources is best captured as an estimate of measurement uncertainty. Measurement uncertainty is defined as “non-negative parameter characterizing the dispersion of the

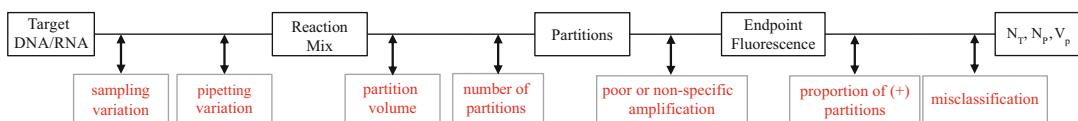
quantity values being attributed to a measurand, based on the information used” [51]. More simply it is defined as the “an estimate of the range of values within which the true value lies”. The process of estimating measurement uncertainty for quantitative DNA measurements produced by qPCR has been reviewed [53] and many of the principles described in this review also apply to dPCR measurements.

The following section briefly discusses important factors to consider when assessing the performance of dPCR assays.

#### 4.2 Considerations for Validation of dPCR Assays

To systematically identify factors that may affect dPCR assays performance, each step in the method should be assessed (Fig. 2). A dPCR method normally involves preparing the sample (i.e., extracting nucleic acid), setting up PCR mix, partitioning the PCR mix, amplifying target by thermal cycling, counting the number of partitions, classifying partitions into positive and negative populations based on end point fluorescence signal and estimating target sequence concentration in the reaction using Poisson statistics and the known partition volume (see Eq. 1). Sample preparation and reaction set up are similar to those required for other quantitative PCR methods. Many factors which can introduce variation in these two steps are likely to affect both dPCR and qPCR. Control materials such as a positive extraction control or a spike-in control can be used to assess extraction efficiency for absolute quantification. However, care should be taken in selection of the positive control as inhibitors can have different effects depending on the assays. For dPCR of genomic DNA, physical methods such as ultrasonication [24] and shredding [54] to fragment DNA have been successfully used to facilitate partitioning and efficient amplification of target sequences, as well as the use of aptly chosen restriction endonuclease digestion. From this point onward in the method, the partitioning process and use of binary end point fluorescence signal detection for quantification distinguishes dPCR from other PCR methods and introduces different potential sources of bias and variation to the end result.

The preparation of dPCR mix involves pipetting and mixing the sample containing nucleic acid into a master mix typically comprising buffer, enzyme, primers, and probe. Pipetting imprecision introduced at this stage are inevitable, but can be reduced by



**Fig. 2** Diagram representing different steps involved in the digital PCR workflow. Potential factors capable of introducing variability to the end measurement result are highlighted by arrows between the different steps in the workflow. Adapted from Jacobs et al. [49]

using calibrated pipettes and larger volumes, and for very high accuracy measurements, gravimetric dilutions using a calibrated balance [38]. Sample heterogeneity and, in the case of very low template DNA concentrations, stochastic effects will also impact on precision of replicates.

Primers and probe sequences that are highly selective to the target nucleic acid sequence are critical for performance and reproducibility of dPCR measurements. Since it is usually impractical to recover dPCR amplicons for confirmatory analysis, specificity of a set of oligonucleotides primers and probe should initially be checked by qPCR using control materials, if available, and/or nucleotide sequencing of amplicons produced from qPCR assays.

Accuracy is a primary performance characteristic for validation of any quantitative analytical method. Accurate quantification of nucleic acid molecules by dPCR relies on correct classification of partitions into negative and positive populations since the proportion of positive partitions is an integral component of the equation used to determine copy number concentration (Eq. 1). A study using mathematical simulation to assess variance components in dPCR concluded that incorrect classification of partitions is one of the largest contributors to inaccuracy of nucleic acid quantification results [50]. Therefore, a key objective in optimization of a dPCR assay is to maximize fluorescence amplitude difference between negative and positive partitions, and to minimize the number of partitions with intermediate fluorescence intensity. In practice, this can be achieved by optimizing efficiency of the dPCR assay.

While dPCR is more tolerant of suboptimal amplification efficiency than qPCR, the risk of misclassifying partitions, specifically false negative classification, is reduced with a sensitive, efficient dPCR assay. Ideally, reference materials at known copy number concentration should be used for assay optimization, to optimize the number of cycles and to determine the appropriate fluorescence threshold. In the absence of suitable reference materials, two or more assays spanning across the target sequence region can be used to cross-check that optimal conditions have been reached and the assay is not being compromised by DNA secondary structure effects. Thermal cycling conditions should be optimized using an annealing temperature gradient to determine conditions for maximum separation between fluorescence levels of negative and positive partitions. This can be followed by a denaturation temperature gradient, particularly in the case of high GC content templates. Optimization of primer and probe concentrations can be undertaken using similar approaches used for qPCR. Optimization of primer concentration is also required for dPCR assays using intercalating dye chemistry. If optimal qPCR conditions have previously been established, these should be verified using dPCR, since

conditions optimized using qPCR are not always optimal for dPCR. In some cases, redesign of assays may be required.

After optimization, partitions containing at least one target molecule are expected to generate a positive fluorescence signal, while partitions without target molecule are expected to generate a negative signal. However, in practice, both false positive and false negative signal can be produced from dPCR assays. The threshold setting will determine false positive and false negative rates. A small proportion of false positives will have a large impact on accuracy of assays with low number of target molecules, while a small proportion of false negatives will have a large impact on accuracy of assays containing high numbers of target molecules.

Limit of detection (LoD) and limit of blank (LoB) are also two important performance parameters to be established for validation of dPCR measurements. In brief, LoD and LoB are, respectively, defined as “the lowest amount of analyte that can detected” and “the highest measurement result that is likely to be observed for a blank sample” with respective probability stated [55]. The general approach to determine LoB involves determination of the distribution of values from multiple measurements of blank samples and for LoD involves a series of measurements of samples containing very low levels of analyte. The 95th percentile of distribution of the blank signal typically determines the LoB. The LoD will be greater than the LoB [55]. Applied to dPCR, LoD could be defined as the smallest number of target nucleic acid molecules that is statistically distinguishable from the background or negative control. One approach to estimating LoD and LoB for a dPCR assay has been experimentally demonstrated in a study to determine lower limits of detection for cancer-related mutations [56].

Verification and monitoring of instrument performance is important for implementation of dPCR as a routine diagnostic. Instrument related factors can introduce bias to results even when assay protocol is followed correctly. For example, accurate and reproducible temperature parameters during amplification are critical for reproducible results. Inaccurate thermal cycling temperatures, due to poor uniformity of temperature across the 96 wells of a thermocycler heating block, can result in no dPCR amplification or low amplification efficiency in wells where optimal temperatures are not reached.

---

## 5 Concluding Remarks

dPCR has increasingly been proved as a “ground-breaking” technology for nucleic acid quantification. Recent developments in commercial dPCR systems made the technology accessible and affordable, and as an immediate consequence the number of users and diversity of application has progressed at staggering pace. As

evident by the published literature from both academia and the industry, the main focus of interest in dPCR technology is in clinical diagnostics applications. However, the translation of a dPCR based test from research laboratory to clinical laboratory setting will require the test to satisfy validation requirements from regulatory authorities.

For dPCR to realize its full potential as a quantitative molecular diagnostic tool, carefully designed validation studies, and sufficient understanding of sources of variation, bias, and associated uncertainty that can lead to incorrect results or poor reproducibility of dPCR measurements data are necessary.

## Acknowledgments

We thank Kate Griffiths and Somanath Bhat for their inputs while reviewing the manuscript.

## References

- Strain MC, Lada SM, Luong T et al (2013) Highly precise measurement of HIV DNA by droplet digital PCR. PLoS One 8(4):e55943. <https://doi.org/10.1371/journal.pone.0055943> [doi] PONE-D-12-26722 [pii]
- Yang R, Paparini A, Monis P et al (2014) Comparison of next-generation droplet digital PCR (ddPCR) with quantitative PCR (qPCR) for enumeration of cryptosporidium oocysts in faecal samples. Int J Parasitol 44 (14):1105–1113. <https://doi.org/10.1016/j.ijpara.2014.08.004>
- Pond MJ, Nori AV, Patel S et al (2015) Performance evaluation of automated urine microscopy as a rapid, non-invasive approach for the diagnosis of non-gonococcal urethritis. Sex Transm Infect 91(3):165–170. <https://doi.org/10.1136/sextrans-2014-051761>
- Devonshire AS, Honeyborne I, Gutteridge A et al (2015) Highly reproducible absolute quantification of mycobacterium tuberculosis complex by digital PCR. Anal Chem 87 (7):3706–3713. <https://doi.org/10.1021/ac5041617>
- Morisset D, Stebih D, Milavec M et al (2013) Quantitative analysis of food and feed samples with droplet digital PCR. PLoS One 8(5): e62583. <https://doi.org/10.1371/journal.pone.0062583>
- Floren C, Wiedemann I, Brenig B et al (2015) Species identification and quantification in meat and meat products using droplet digital PCR (ddPCR). Food Chem 173:1054–1058. <https://doi.org/10.1016/j.foodchem.2014.10.138>
- Racki N, Dreо T, Gutierrez-Aguirre I et al (2014) Reverse transcriptase droplet digital PCR shows high resilience to PCR inhibitors from plant, soil and water samples. Plant Methods 10(1):42. <https://doi.org/10.1186/s13007-014-0042-6>
- Cao Y, Raith MR, Griffith JF (2015) Droplet digital PCR for simultaneous quantification of general and human-associated fecal indicators for water quality assessment. Water Res 70:337–349. <https://doi.org/10.1016/j.watres.2014.12.008>
- Kim TG, Jeong SY, Cho KS (2014) Comparison of droplet digital PCR and quantitative real-time PCR for examining population dynamics of bacteria in soil. Appl Microbiol Biotechnol 98(13):6105–6113. <https://doi.org/10.1007/s00253-014-5794-4>
- Day E, Dear PH, McCaughan F (2013) Digital PCR strategies in the development and analysis of molecular biomarkers for personalized medicine. Methods (San Diego, Calif) 59 (1):101–107. <https://doi.org/10.1016/jymeth.2012.08.001>
- Jennings LJ, George D, Czech J et al (2014) Detection and quantification of BCR-ABL1 fusion transcripts by droplet digital PCR. J Mol Diagn 16(2):174–179. <https://doi.org/10.1089/hgtb.2013.131> 10.1016/j.jmoldx.2013.10.007

12. Regan JF, Kamitaki N, Legler T et al (2015) A rapid molecular approach for chromosomal phasing. *PLoS One* 10(3):e0118270. <https://doi.org/10.1371/journal.pone.0118270>
13. Nadauld L, Regan JF, Miotke L et al (2012) Quantitative and sensitive detection of cancer genome amplifications from formalin fixed paraffin embedded tumors with droplet digital PCR. *Transl Med (Sunnyvale, Calif)* 2(2). <https://doi.org/10.4172/2161-1025.1000107>
14. Sanmamed MF, Fernandez-Landazuri S, Rodriguez C et al (2015) Quantitative cell-free circulating BRAFV600E mutation analysis by use of droplet digital PCR in the follow-up of patients with melanoma being treated with BRAF inhibitors. *Clin Chem* 61(1):297–304. <https://doi.org/10.1373/clinchem.2014.230235>
15. Heredia NJ, Belgrader P, Wang S et al (2013) Droplet digital PCR quantitation of HER2 expression in FFPE breast cancer samples. *Methods* 59(1):S20–S23. <https://doi.org/10.1016/j.ymeth.2012.09.012>
16. Garcia-Murillas I, Lambros M, Turner NC (2013) Determination of HER2 amplification status on tumour DNA by digital PCR. *PLoS One* 8(12):e83409. <https://doi.org/10.1371/journal.pone.0083409>
17. Beaver JA, Jelovac D, Balukrishna S et al (2014) Detection of cancer DNA in plasma of patients with early-stage breast cancer. *Clin Cancer Res* 20(10):2643–2650. <https://doi.org/10.1158/1078-0432.ccr-13-2933>
18. Barrett AN, McDonnell TC, Chan KC et al (2012) Digital PCR analysis of maternal plasma for noninvasive detection of sickle cell anemia. *Clin Chem* 58(6):1026–1032. <https://doi.org/10.1373/clinchem.2011.178939>
19. Weber ND, Stone D, Sedlak RH et al (2014) AAV-mediated delivery of zinc finger nucleases targeting hepatitis B virus inhibits active replication. *PLoS One* 9(5):e97579. <https://doi.org/10.1371/journal.pone.0097579>
20. Miyaoka Y, Chan AH, Judge LM et al (2014) Isolation of single-base genome-edited human iPS cells without antibiotic selection. *Nat Methods* 11(3):291–293. <https://doi.org/10.1038/nmeth.2840>
21. Cai J, Miao X, Li Y et al (2014) Whole-genome sequencing identifies genetic variances in culture-expanded human mesenchymal stem cells. *Stem Cell Reports* 3(2):227–233. <https://doi.org/10.1016/j.stemcr.2014.05.019>
22. Gao S, Zheng C, Chang G et al (2015) Unique features of mutations revealed by sequentially reprogrammed induced pluripotent stem cells. *Nat Commun* 6:6318. <https://doi.org/10.1038/ncomms7318>
23. Beck J, Bierau S, Balzer S et al (2013) Digital droplet PCR for rapid quantification of donor DNA in the circulation of transplant recipients as a potential universal biomarker of graft injury. *Clin Chem* 59(12):1732–1741. <https://doi.org/10.1373/clinchem.2013.210328>
24. George D, Czech J, John B et al (2013) Detection and quantification of chimerism by droplet digital PCR. *Chimerism* 4(3):102–108. <https://doi.org/10.4161/chim.25400>
25. Halling KC, Schrijver I, Persons DL (2012) Test verification and validation for molecular diagnostic assays. *Arch Pathol Lab Med* 136(1):11–13. <https://doi.org/10.5858/arpa.2011-0212-ED>
26. Higuchi R, Dollinger G, Walsh PS et al (1992) Simultaneous amplification and detection of specific DNA sequences. *Biotechnology* 10(4):413–417
27. Heid CA, Stevens J, Livak KJ et al (1996) Real time quantitative PCR. *Genome Res* 6(10):986–994
28. McCrady MH (1915) The numerical interpretation of fermentation-tube results. *J Infect Dis* 17(1):183–212. <https://doi.org/10.2307/30083495>
29. Simmonds P, Balfé P, Peutherer JF et al (1990) Human immunodeficiency virus-infected individuals contain provirus in small numbers of peripheral mononuclear cells and at low copy numbers. *J Virol* 64(2):864–872
30. Brisco MJ, Condon J, Sykes PJ et al (1991) Detection and quantitation of neoplastic cells in acute lymphoblastic leukaemia, by use of the polymerase chain reaction. *Br J Haematol* 79(2):211–217
31. Sykes PJ, Neoh SH, Brisco MJ et al (1992) Quantitation of targets for PCR by use of limiting dilution. *BioTechniques* 13(3):444–449
32. Vogelstein B, Kinzler KW (1999) Digital PCR. *Proc Natl Acad Sci U S A* 96(16):9236–9241
33. Thorsen T, Maerkli SJ, Quake SR (2002) Microfluidic large-scale integration. *Science* 298(5593):580–584. <https://doi.org/10.1126/science.1076996>
34. Ottesen EA, Hong JW, Quake SR et al (2006) Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. *Science* 314(5804):1464–1467. <https://doi.org/10.1126/science.1131370>. 314/5804/1464 [pii]
35. Warren L, Bryder D, Weissman IL et al (2006) Transcription factor profiling in individual

- hematopoietic progenitors by digital RT-PCR. *Proc Natl Acad Sci U S A* 103(47):17807–17812
36. Dube S, Qin J, Ramakrishnan R (2008) Mathematical analysis of copy number variation in a DNA sample using digital PCR on a nanofluidic device. *PLoS One* 3(8):e2876
37. Bhat S, Herrmann J, Armishaw P et al (2009) Single molecule detection in nanofluidic digital array enables accurate measurement of DNA copy number. *Anal Bioanal Chem* 394(2):457–467
38. Pinheiro LB, Coleman VA, Hindson CM et al (2012) Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Anal Chem* 84:1003–1011. <https://doi.org/10.1021/ac202578x>
39. Zhong Q, Bhattacharya S, Kotsopoulos S et al (2011) Multiplex digital PCR: breaking the one target per color barrier of quantitative PCR. *Lab Chip* 11(13):2167–2174. <https://doi.org/10.1039/c1lc20126c>
40. Pekin D, Skhiri Y, Baret JC et al (2011) Quantitative and sensitive detection of rare mutations using droplet-based microfluidics. *Lab Chip* 11(13):2156–2166. <https://doi.org/10.1039/c1lc20128j> [doi]
41. Whale AS, Huggett JF, Cowen S et al (2012) Comparison of microfluidic digital PCR and conventional quantitative PCR for measuring copy number variation. *Nucleic Acids Res* 40(11):e82. <https://doi.org/10.1093/nar/gks203>
42. Belgrader P, Tanner SC, Regan JF et al (2013) Droplet digital PCR measurement of HER2 copy number alteration in formalin-fixed paraffin-embedded breast carcinoma tissue. *Clin Chem* 59(6):991–994. <https://doi.org/10.1373/clinchem.2012.197855>
43. Nixon G, Garson JA, Grant P et al (2014) Comparative study of sensitivity, linearity, and resistance to inhibition of digital and nondigital polymerase chain reaction and loop mediated isothermal amplification assays for quantification of human cytomegalovirus. *Anal Chem* 86(9):4387–4394. <https://doi.org/10.1021/ac500208w>
44. Sedlak RH, Kuypers J, Jerome KR (2014) A multiplexed droplet digital PCR assay performs better than qPCR on inhibition prone samples. *Diagn Microbiol Infect Dis* 80(4):285–286. <https://doi.org/10.1016/j.diagmicrobio.2014.09.004>
45. Huggett JF, Novak T, Garson JA et al (2008) Differential susceptibility of PCR reactions to inhibitors: an important and unrecognised phenomenon. *BMC Res Notes* 1:70
46. Dingle TC, Sedlak RH, Cook L et al (2013) Tolerance of droplet-digital PCR vs real-time quantitative PCR to inhibitory substances. *Clin Chem* 59(11):1670–1672. <https://doi.org/10.1373/clinchem.2013.211045>
47. Hindson BJ, Ness KD, Masquelier DA et al (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem* 83(22):8604–8610. <https://doi.org/10.1021/ac202028g> [doi]
48. Bhat S, McLaughlin JL, Emslie KR (2011) Effect of sustained elevated temperature prior to amplification on template copy number estimation using digital polymerase chain reaction. *Analyst* 136(4):724–732. <https://doi.org/10.1039/c0an00484g> [doi]
49. Corbisier P, Pinheiro L, Mazoua S et al (2015) DNA copy number concentration measured by digital and droplet digital quantitative PCR using certified reference materials. *Anal Bioanal Chem* 407(7):1831–1840. <https://doi.org/10.1007/s00216-015-8458-z>
50. Jacobs BK, Goetghebeur E, Clement L (2014) Impact of variance components on reliability of absolute quantification using digital PCR. *BMC Bioinformatics* 15:283. <https://doi.org/10.1186/1471-2105-15-283>
51. Hindson CM, Chevillet JR, Briggs HA et al (2013) Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat Methods* 10(10):1003–1005. <https://doi.org/10.1038/nmeth.2633>
52. JCGM (2008) JCGM 200:2012 international vocabulary of metrology—basic and general concepts and associated terms (VIM), 3rd edn. BIPM, Sèvres Cedex, France
53. Griffiths KR, Burke DG, Emslie KR (2011) Quantitative polymerase chain reaction: a framework for improving the quality of results and estimating uncertainty of measurement. *Anal Methods* 3:2201–2211
54. Yukl SA, Kaiser P, Kim P et al (2014) Advantages of using the QIAshredder instead of restriction digestion to prepare DNA for droplet digital PCR. *BioTechniques* 56(4):194–196. <https://doi.org/10.2144/000114159>
55. Clinical and Laboratory Standards Institute (2004) Protocols for determination of limits of detection and limits of quantitation; approved guideline, vol vol 24. CLSI, Wayne, OA, USA. Contract No.: EP17
56. Milbury CA, Zhong Q, Lin J et al (2014) Determining lower limits of detection of digital PCR assays for cancer-related gene mutations. *Biomol Detect Quantif* 1:8–22



# Chapter 3

## Fundamentals of Counting Statistics in Digital PCR: I Just Measured Two Target Copies—What Does It Mean?

Svilen Tzonev

### Abstract

Current commercially available digital PCR (dPCR) systems and assays are capable of detecting individual target molecules with considerable reliability. As tests are developed and validated for use on clinical samples, the need to understand and develop robust statistical analysis routines increases. This chapter covers the fundamental processes and limitations of detecting and reporting on single molecule detection. We cover the basics of quantification of targets and sources of imprecision. We describe the basic test concepts: sensitivity, specificity, limit of blank, limit of detection, and limit of quantification in the context of dPCR. We provide basic guidelines how to determine those, how to choose and interpret the operating point, and what factors may influence overall test performance in practice.

**Key words** Statistics, Counting, Sensitivity, Specificity, Subsampling, Limit of detection, Limit of blank, Poisson distribution, False positive, False negative, Performance characteristics

---

### 1 Introduction

Digital PCR (dPCR) technology promises to change how genomic markers are detected and quantified [1, 2]. It has been demonstrated to effectively count amplifiable targets at single molecule resolution. This is possible because partitioning isolates the targets, while end-point amplification allows highly reliable detection of the reaction components in each partition.

This chapter will focus on the fundamental concepts of counting statistics as they apply to detection and quantification of targets in dPCR tests. We will cover basic detection and quantification, describe critical test performance parameters, and provide recommendations for choosing the test operating point.

We will first cover detection of single species and discuss detection and counting of generic target molecules.

## 2 Counting Molecules

In digital PCR the PCR reaction is first partitioned into independent, usually equal size, volumes (droplets or chambers on a device) [3]. During a standard thermocycling protocol, amplifiable genomic target molecules will be replicated and their number will increase exponentially. In a common format, hydrolysis probes matching specific amplicon counterparts are used. The probes have quencher and dye molecules attachments. When the polymerase enzyme cleaves the hybridized probes it releases unquenched dye molecules into the reaction. Thus, the concentration of free dye in the partition will increase exponentially with the PCR cycles. At end-point, any partition that contained at the beginning an amplifiable target will contain detectable high concentration of fluorescent dye and can be detected as “positive.” Any partition that did not contain a target will remain “negative.” In practice, we have to deal with damaged or otherwise poorly PCR-accessible target molecules that may not amplify from cycle 1, nonspecific probe hybridization that may lead to false positives, polymerase errors that may lead to false positive or false negatives, and other complications. We will cover such cases later.

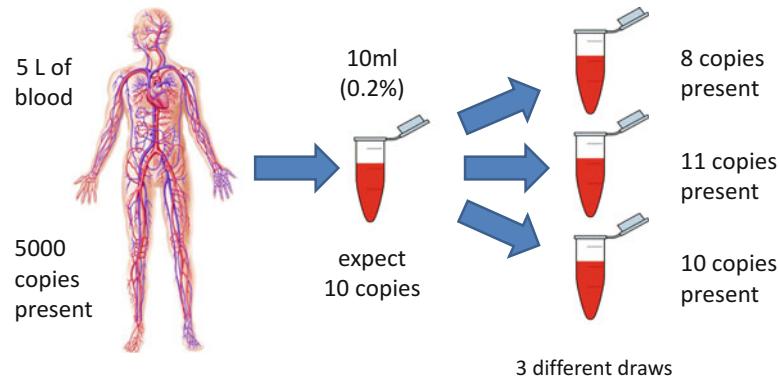
First we will focus on an “ideal dPCR counting machine” that can detect with 100% certainty and no confusion the presence or absence of a target molecule in a partition.

---

## 3 Subsampling

Let us first consider a test that detects a single species of target molecules. Even a perfect counting machine, which makes no mistakes of any kind, has to deal with the stochastic nature of many molecular phenomena. In a typical experiment or test, we want to measure the concentration of a particular biomolecule as it is *in vivo*. In order to do this, we have to use a subsample of the whole—a draw of blood or a tissue biopsy. We measure the subsample *in vitro* to estimate the state of the whole. In the following we will use the liquid biopsy paradigm to illustrate the basic concepts.

A hypothetical patient has 5 L of blood. If we draw 10 ml of blood at a time, we are subsampling 0.2% of the whole blood volume. Let us say that there are 5000 copies of the target molecules of interest in the 5 L of blood. This corresponds to a concentration of 1 target/1 ml of blood. In any 10 ml draw, we expect to collect on average 10 target molecules. Assuming perfect sample processing with no losses of any kind, we would expect to load on average 10 targets in our machine, which would detect all of them. However, any individual 10 ml draw may contain 8 or 11 or



**Fig. 1** Subsampling process. A patient has 5 L of blood in which 5000 molecules of interest are present. Any 10 ml draw of blood represents 0.2% of the total blood volume. We would expect to see on average 10 target molecules. However, in any actual 10 ml drawn, we may see different numbers of targets molecules due to the stochastic nature of molecule distribution in the 5 L of blood

10 **actual** target molecules. Our ability to make confident statements about the whole is still limited by the subsampling problem. See Fig. 1.

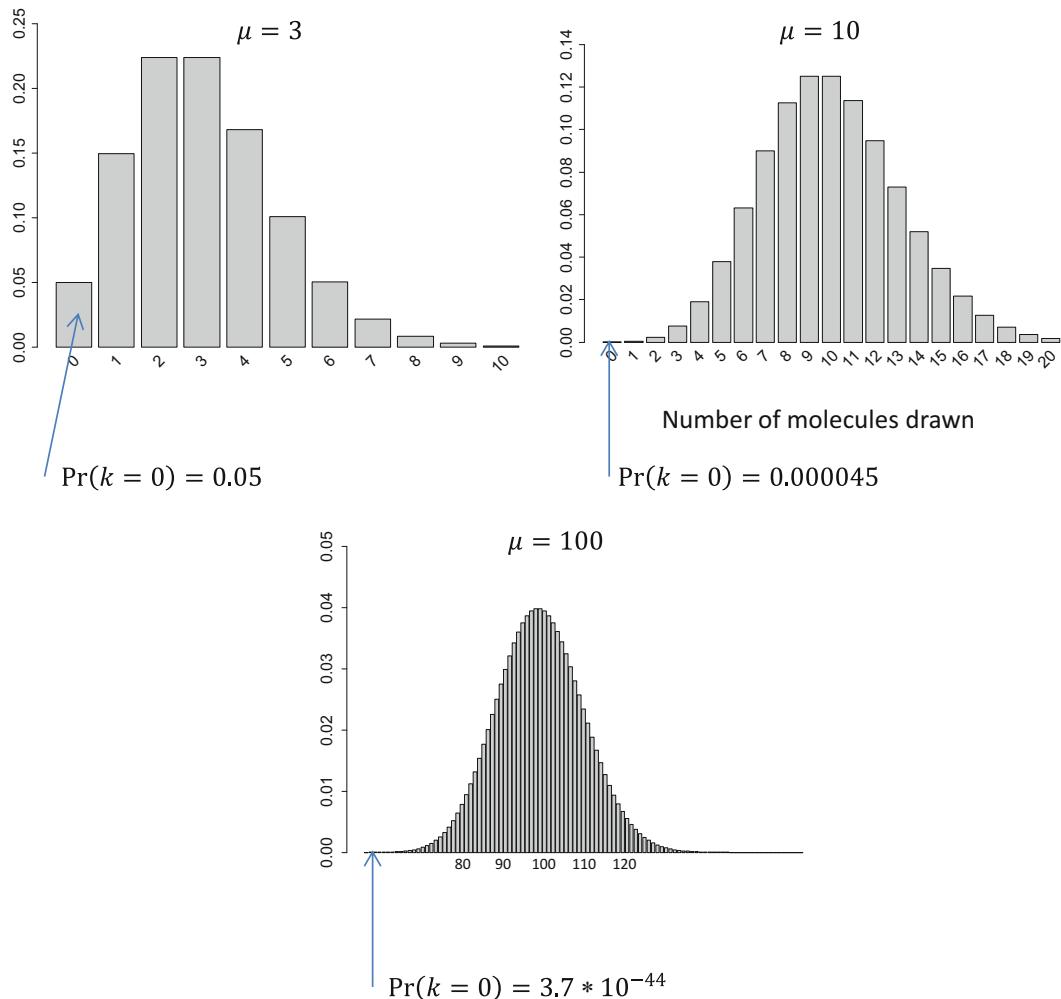
How variable is the number of actual targets per draw? When we are drawing a tiny fraction of the whole and, as long as we can assume that the targets are all independent and equally likely to be sampled, we can describe the probability to draw exactly  $k$  actual targets with the Poisson distribution:

Probability (to draw  $k$  molecules, given expected number molecules is  $\mu$ )

$$= \mu^k \frac{e^{-\mu}}{k!}$$

where  $\mu$  is the **expected** or average number drawn; in our case it is 10. Note that  $k$  must be an integer—we may draw 0, 1, 2, etc. molecules, while  $\mu$  can be a noninteger number—the average may be 1.5 molecules. Figure 2 illustrates the Poisson distribution for  $\mu = 3, 10$ , and  $100$ . As we can see, when  $\mu$  is small, there is significant variability of how many targets will be subsampled in a draw relative to the expected number. Note how in the case of three expected molecules, there is about 5% probability that the draw would contain no targets at all! This observation will form the basis of the fundamental bound on the limit of detection, this is discussed later. Note also that for large  $\mu$  the relative uncertainty shrinks as the distribution becomes more concentrated around the average. It also becomes more like the familiar normal or Gaussian distribution.

## Probability



**Fig. 2** Poisson distribution. Poisson distribution for three different values of the expected number of molecules drawn,  $\mu = 3, 10$  or  $100$ . Probability (to draw  $k$  molecules, given expected number molecules is  $\mu = \mu^k \frac{e^{-\mu}}{k!}$ ). The height of each bar for particular integer  $k$  represents the probability to draw  $k$  molecules. Note how for small  $\mu$ , the distribution shape is skewed and becomes increasingly more symmetrical and concentrated around the maximum for larger values of  $\mu$ . For each case, the probability to draw exactly zero molecules is also shown

---

## 4 Multiple Occupancy and Partitioning

The original concept of dPCR assumed a **limiting dilution** regime, i.e., each partition would not contain more than one target with any significant probability. Most of current commercial systems [4–7] allow robust measurement of target copies even when multiple molecules can occupy the same partition with a significant probability. Quantification is possible under the assumption that

the targets are allocated to a partition independently and with equal probability. (Such assumptions would be violated if the targets are in some way “entangled” or if the partitions are of different volumes.) It turns out that the appropriate statistical framework is also based on the Poisson distribution, since each partition is similar to a blood draw from the previous example—each partition is a small subsample of a larger whole. For specific details see [8, 9]; here we will only focus on how the target concentration is estimated.

A positive partition may have contained one or more targets. There is no way to be certain. A negative partition, on the other hand, must have contained exactly 0 targets, otherwise it would be positive. Once we have counted all partitions (positive and negative),  $N_{\text{tot}}$ , and all negative partitions,  $N_{\text{neg}}$ , we can use this formula to estimate the average target occupancy per partition,  $\lambda$ :

$$\lambda = \ln(N_{\text{tot}}) - \ln(N_{\text{neg}}) = -\ln\left(\frac{N_{\text{neg}}}{N_{\text{tot}}}\right)$$

The total number of targets,  $T$ , present at the beginning of the reaction is then simply

$$T = \lambda N_{\text{tot}}$$

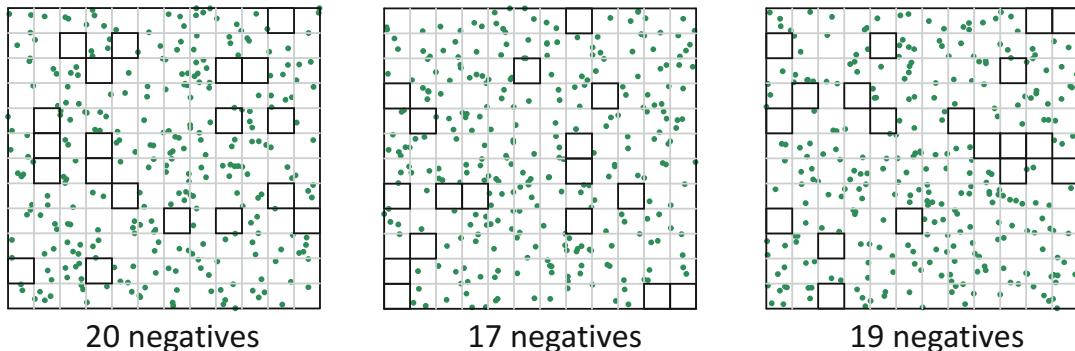
The value of  $\lambda$  calculated in this way is the “most likely value.” We cannot be sure of the exact number in reality since we do not know the precise distribution of targets into partitions.

To arrive at the target concentration, we divide  $\lambda$  by the partition volume:

$$[T] = \frac{\lambda}{V_{\text{partition}}}$$

Consider a given reaction mix with a fixed number of targets in it. If we perform multiple thought partitioning experiments we can understand the variability of the results around the most likely values. For each partitioning, depending on the exact (stochastic) pattern of target distribution into the partitions, we may see different number of negative partitions. This would lead us to estimate slightly different values of the target copies. We are again subject to the inevitable stochastic nature of how molecules “choose” their partitions.

Figure 3 illustrates the concept of partitioning uncertainty. We have partitioned the same reaction volume with the same number of molecules,  $T$ , three times into the same number of partitions of equal volume. Since the molecules are distributed stochastically into the partitions, the patterns in each realization will be different. The observed number of partitions with no molecules will also vary between realizations. In fact, the number of **actual** negative partitions is a variable that also follows a Poisson distribution, given the **expected** number of negative partitions. We can then use similar



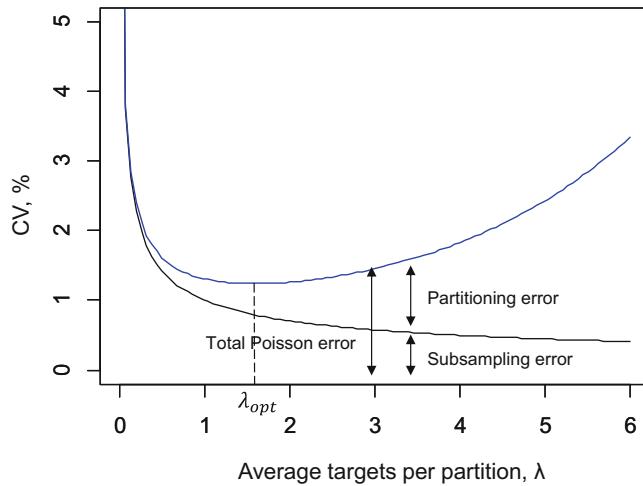
**Fig. 3** Partitioning uncertainty. A thought experiment where a reaction volume with fixed number of molecules is partitioned three times. Exact location of each molecule in the resulting partitions is stochastic. This leads to a different number of partitions with no molecules at all—negative partitions are highlighted. This will result in somewhat different “most likely estimate” of the number of molecules in the reaction. The number of negative partitions is itself a variable that follows a Poisson distribution

math as above to describe how uncertain the estimate of  $\lambda$  is going to be even when we know the exact number of molecules,  $T$ , in the total volume.

## 5 Total Poisson Uncertainty

So far, we saw how subsampling and partitioning variability limit our ability to estimate precisely the number of target molecules *in vivo* even with a perfect counting machine. When we combine the two sources of variability, we arrive at the frequently described dPCR uncertainty curve, shown also Chap. 2. When only a small number of targets is drawn (provided we have enough partitions), the total uncertainty is dominated by the subsampling processes. In other words, it does not matter if we put, five **actual** targets in 20,000 or 20,000,000 partitions, we will know with high confidence that there were five targets in the reaction we measured and **approximately** five targets per blood draw, if these were repeated multiple times. When the number of targets increases relative to the number of partitions (average occupancy  $\lambda$  is larger), at some point there will be too few negative partitions. Since the **actual** number of negative partitions is a stochastic variable, the uncertainty of  $\lambda$  increases and, thus, the uncertainty of  $T$ .

Figure 4 illustrates the contributions from subsampling and partitioning errors to the total Poisson statistical error. For low occupancy numbers,  $\lambda$ , (number of targets is much less than the number of partitions), total uncertainty is dominated by subsampling. While for large occupancy numbers (number of targets much larger than the number of partitions), total uncertainty is dominated by partitioning. The lowest overall uncertainty is reached at  $\lambda_{opt} \approx 1.6$ . For 10,000 partitions even when  $\lambda = 6$ , the coefficient of



**Fig. 4** Total Poisson Error = Subsampling + Partitioning Errors, expressed as % CV, as a function of average targets per partition,  $\lambda$ . The black curve shows the subsampling error. The blue curve shows the total Poisson error. The distance between the curves is due to the partitioning error. For small  $\lambda$ , the subsampling error dominates, while the partitioning error is small. For large  $\lambda$ , the subsampling error is small, while the partitioning error dominates. The total Poisson error is minimized at  $\lambda_{opt} \approx 1.6$ . Values for %CV are calculated for 10,000 partitions. Increasing the number of partitions brings all curves lower (smaller %CV) but does not change the general shapes. Optimum stays at 1.6 targets per partition

variation of  $\lambda$  and  $T$  is ~3%, i.e., sufficiently low to allow accurate and precise quantification in most cases.

A few words about the concept of confidence intervals. When we estimate experimentally some unknown variable we talk about a point estimate and a confidence interval around it. The point estimate is the most likely value of the variable; this is what is usually reported. A 95% confidence interval, around the point estimate, describes the range between a minimum and a maximum value of the variable, between which the true value of it will lie 95% of the time. In other words, if we knew the exact value of the variable and repeated the estimation experiment many times, 95% of the time it will indeed lie in the stated confidence interval. Most commercial dPCR systems report both point value estimates and 95% confidence intervals for target copies and concentrations.

## 6 False Positives and False Negatives

So far, we have discussed an idealized case where our perfect system makes no mistakes of any kind. In practice, we have to deal with false positive and false negative reporting. In the simplest case, we can define a false positive or negative at the level of the partitions—a

partition that should have been detected as a negative is reported as a positive and vice versa. The rates of such false reporting per partition define a false positive rate, FPR, and a false negative rate, FNR. Which of these is more important depends on the test and how much we care about each kind of error.

For example, when false (positive and negative) rates are small and there is a small number of targets, the FNR is usually not important in absolute terms, as we only have a few true positive partitions that may be converted into a false negative. In this case, the FPR is more significant as there is a large number of true negative partitions, each of which may turn into a false positive. In the opposite case—a large number of targets—it is the positive partitions that are numerous and thus FNR matters more, in absolute terms, than the FPR per partition.

In both cases, the prevalence of the state or condition measured plays a role in determining which false rate is more important. This is true at the patient or sample level as well, in the sense of the prevalence of the condition we are trying to measure with our test. There are cases where the danger of misdiagnosis at the clinical level influences greatly the relative importance of FPR and FNR. For example, in a prenatal screening test for Down syndrome, most patients are expected to have normal pregnancies. Due to the large number of individuals screened, even a small test FPR will produce a large number of false positive results that would require additional procedures. On the other hand, a false negative result is also very damaging as the overall cost to the patient is significant, since the fetus would be incorrectly classified as normal.

We can think of false positive and false negative rates at the level of a well, test (multiple wells) or sample (if multiple tests are performed). It is important to keep in mind which of these concepts is being used. The combination of the sample preparation protocol, detection assay and dPCR measurement system will be subject to the limitations imposed by both stochastic molecular processes and the FP/FN detection rates.

A few common examples of processes that can lead to false results include:

- Nonspecific primers and/or probe hybridization
- Polymerase errors
- Fluorescent dust particles
- Contamination and cross-contamination
- Inappropriate thresholding algorithms

In practice, these should be recognized and controlled at the assay design level, periodic equipment maintenance, robust laboratory practices, and frequent specific process monitoring.

---

## 7 Basic Test Concepts and Their Interpretation in dPCR

We distinguish two major types of tests—qualitative and quantitative. A qualitative test can only categorize an unknown sample as *positive* or *negative* (in some case there is also a third, *uncalled* category). A quantitative test will produce a *value* for the measurand of interest in an unknown sample when possible or perhaps produce an out-of-reportable-range result. For dPCR, quantification is most often as target counts or copies in the test or concentration of target copies in the test. There may be other derivative measures that are based on combinations of such measurements, for example a ratio of two concentrations or similar.

dPCR essentially always produces a direct estimate of the target copies in the reaction. Thus, it is by nature quantitative. To call a sample negative or positive, an analytical call cutoff value is applied—any sample with detected copies higher than or equal to the cutoff is declared as positive; any sample with fewer copies than the cutoff is declared as negative. In some cases, there may be a “grey zone” where no call is produced. We will focus on the simpler binary version of sample positive/negative determination.

Three fundamental parameters determine the performance of a given dPCR test: FPR, FNR, and the choice of analytical cutoff value (if needed for a qualitative test).

Next, we will define additional important test descriptors and discuss how they are influenced by these fundamental parameters in the context of counting assays in dPCR. Looking ahead at Fig. 5, discussed below, will help the reader visualize the concepts.

---

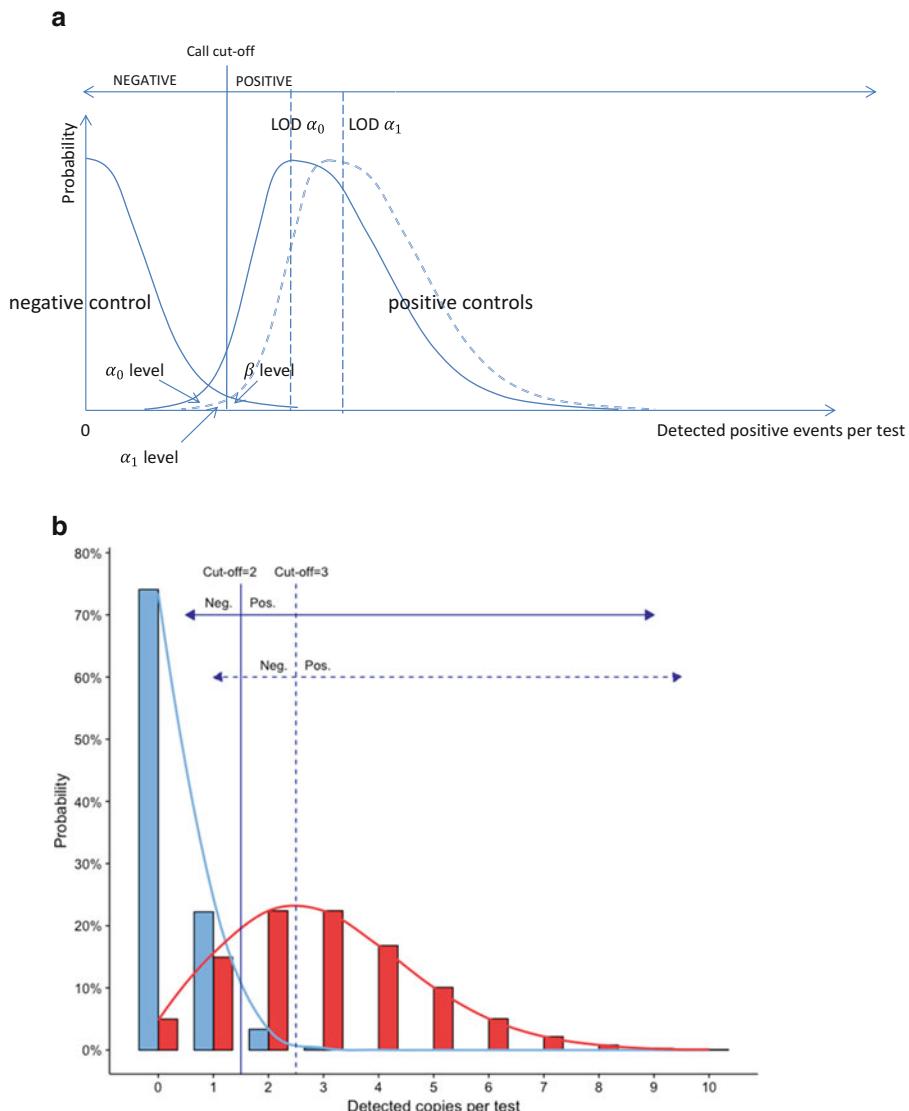
## 8 Sensitivity

For a qualitative test the *sensitivity* is defined as the probability that a truly positive (as determined by a prior test or clinical criterion) sample is measured as positive by the test. 100% sensitivity means that all truly positive samples do come out as positive by the test. Sensitivity may be less than 100% when the FNR per test is higher than zero or if the call cutoff value is set too high and truly positive samples are called as negatives. Intuitively, if we challenge the test with strongly positive or strongly negative samples, we can achieve better performance (See below on the concept of limit of detection).

---

## 9 Specificity

For a qualitative test, *specificity* is defined as the probability that a truly negative sample is measured as negative by the test. Again, 100% specificity means that the test is never wrong when negative



**Fig. 5** Relationship between sensitivity, specificity, and call cutoff (a). Classic description and definition of these test concepts. Horizontal axis—signal measured by the test (in dPCR this is detected targets/events). Vertical axis—probability of a particular level to be measured. Left curve—the test response to a negative/blank control, Right curves—the test response to two positive samples at different levels. Depending where we choose the value of the call cutoff, the areas under the curves to the right and to the left of the cutoff determine the values of specificity ( $\beta$  level) and sensitivity ( $\alpha$  level), respectively. Sensitivity =  $1 - \alpha$ . Specificity =  $1 - \beta$ . Moving the cutoff to the right will decrease  $\beta$  and increase  $\alpha$ , thus, increase specificity but decrease sensitivity. The opposite is true when moving the cut-off to the left. The value of the measurand for which  $\alpha_0 = 5\%$ , determines the 95% confidence level limit of detection of the test (indicated as the peak of the corresponding positive control curve, or most likely value). For increased Sensitivity,  $\alpha_1 < 5\%$ , the corresponding LOD will be greater than the LOD at 95%. When we use blank samples for the negative control, the cutoff value corresponds to the limit of blank at  $1 - \beta$  level. (b). Same concepts as they apply in dPCR. Continuous curves are replaced by histograms as dPCR reports integer values for the targets detected. Blue: histogram of possible copies for a negative sample, Red: histogram of possible copies for a positive sample. Moving the cutoff value is only possible in discrete ways, which results in jumps of possible values of sensitivity and specificity ( $\alpha$  and  $\beta$  levels not shown)

samples are measured. Any value higher than zero for the FPR per test will lead to specificity that is less than 100%. Alternatively, a cutoff that is set too low, causing truly negative samples to be called as positives, will also cause specificity of less than 100%.

Figure 5 illustrates the inherent trade-off between sensitivity and specificity. Note how by moving the cutoff value, one changes both the specificity and sensitivity of a given test, increasing one while decreasing the other. These change in opposite directions. For tests that produce integer values (molecules detected in test) cutoff values can only be discrete, which leads to a set of possible combinations of sensitivity/specificity values.

---

## 10 Limit of Blank (LOB)

The limit of blank, LOB, at a particular confidence level, usually 95%, is a critical concept. The formal definition, according to CLSI guidelines [10], is the maximum value of the analyte that may be reported 95% of the time when we measure a true blank sample. In dPCR we report counts of target molecules, thus the LOB must be an integer—0, 1, 2, etc. counts per test. When we say that the LOB for a test is 1 copy, we mean that 95% of the time when we measure a negative/blank sample we will report at most one copy. This is equivalent to stating that we will report 2, 3, or more copies at most 5% of the time (*See Fig. 5*).

The LOB is closely related to the *specificity* of the test and the underlying FPR per partition or per test. When we measure a blank sample, each partition should be reported as negative. However, when the FPR per partition is not 0, some truly negative partitions may be reported as false positives. In most tests we will still report 0 positive partitions and, thus, 0 target copies. But for some samples, we may report 1, 2 or more copies. These will all be false positive counts.

The robust way to experimentally measure LOB of a test is to run multiple blank samples, record the number of targets reported and rank order the results. When we draw a line at the 95% percentile, the value where this happens is the 95% LOB. It is good practice to verify the assumed Poisson distribution of the counts by fitting a model. Theory predicts that these counts will also follow a Poisson distribution, as long as each negative partition may be independently misreported as positive. The average number of false positive target counts per test will equal the FPR per partition times the number of partitions per test. If the counts per test do not appear to follow a Poisson distribution, we may still use the LOB number, but need to be aware there are other effects playing a role, other than unreliable detection of negative partitions. These may point to additional assay variability or inadequate lab practices.

## 11 Limit of Detection (LOD)

The limit of detection, LOD, at a particular confidence level, frequently 95%, is another critical concept. The formal definition, according to the CLSI guidelines, is the minimum level of the analyte in a sample that will be reported as detected with this same 95% probability. The LOD level depends on the analytical call cutoff threshold and FPR and FNR values. An increase of the FPR per partition will lead to an increase of the false positives and will cause the system to report on average higher counts than truth, pushing the apparent LOD lower if the cutoff is held constant (*see* below). An increase of the FNR has the opposite effect—it may cause positive partitions to appear as negative, effectively reporting a lower value for the analyte than truth. This will increase the LOD, if we hold the cutoff threshold constant.

The lowest possible value for the call cutoff is 1 copy per test, i.e., if we see a single copy of the target we will call the sample positive. Even if FPR and FPN are exactly zero, we already saw how subsampling puts a fundamental limit on what target levels we can detect confidently. Recall when we expect to draw on average three target copies, there is a 5% probability that any individual draw will contain no copies at all—there would be nothing to detect! This observation sets a critical level of three copies as the natural LOD of molecular counting assays like in dPCR when subsampling is a factor (*see* exceptions in “Finite Subsampling” section below). This limitation applies to any other counting based methods, like versions of next generation sequencing, that detect individual target molecules. Choosing a higher level for the call cutoff will produce a higher LOD value. Table 1 illustrates the LOD 95% values for different cutoff values.

A word of caution, the reader should not confuse the concepts of Sensitivity and Positive Predictive Value (PPV) of a test. The

**Table 1**  
**LOD 95% values for different test cutoff values (copies per test). FPR = 0, FNR = 0**

Call cutoff (positive sample) copies per test	LOD 95% copies per test
1	3.00
2	4.74
3	6.30
4	7.75
5	9.15

Sensitivity, again, relates to the probability that the test will call a positive sample as positive, while PPV inverts the logic—it defines the probability that a sample that is called positive by the test is indeed positive in reality.

## 12 Limit of Quantification (LOQ)

For quantitative tests the limit of quantification (LOQ) of a test is defined as the minimum level of an analyte that can be measured within a predefined level of uncertainty. The uncertainty is typically expressed by the standard deviation or the coefficient of variation based on multiple measurements of a given sample at the LOQ level. The preferred imprecision level depends on the situation, typical numbers include 20%, 35%, or higher and depend on the requirements of the application.

By definition, the values of the three parameters must be ordered

$$\text{LOB} < \text{LOD} \leq \text{LOQ}$$

For any test the LOD must be higher than the LOB, and the LOQ cannot be lower than the LOD. At the lowest LOD achievable, three copies, the minimal possible coefficient of variation is

$$CV_{\min} = \frac{1}{\sqrt{3}} \approx 58\%$$

In practice, the observed CV at this level will usually be higher as other sources of variability will inevitably add to this theoretical minimum. There are dPCR tests with remarkably low levels of LOB, LOD, and LOQ, which reach very close to the theoretical limits of single molecule detection levels.

## 13 Why and How Do These Performance Characteristics Matter

The concepts discussed above describe the overall performance of a test of interest. This is expressed in the context of “absolute truth” on the analytical side or a prior method of measurement of what matters on the clinical side. For example, if a patient has ten viral particles per 10 ml of blood (for a total of 5000 in the circulation), is this clinically significant? Should this be viewed as a positive or a negative patient? Typically, one chooses the clinical cutoff level on a population basis based on criteria relevant to patient health or prognosis.

The relationships between the test performance characteristics and the clinical cutoff level determine the ultimate utility of the test.

**Table 2**  
**Clinical and test concepts meaning and context**

	<b>Applies to</b>	<b>Determines</b>	<b>How known or chosen</b>	<b>When known relative to test</b>
Clinical cutoff	Patient population	Is a patient considered clinically positive or negative	Based on external clinical knowledge	Before the test is performed
Test/call cutoff	Test	When the test reports positive or negative result	Chosen to satisfy test requirements	Before the test is performed
LOB	Test	Maximum value (95% probability) that may be reported on a blank sample	Determined during test development and validation	Before the test is performed
LOD	Test	Level of analyte that can be reliably (95% probability) detected by the test	Determined during test development and validation	Before the test is performed
Test result	Sample + Test	What the test measured for a given sample	By performing a test on a sample	After the test is performed

The meaning of and the context of selected concepts are summarized in Table 2.

*In general, a test is suitable for determining patient status when its LOD matches the clinical cutoff and does not produce too many false positive results.* But there may be cases where higher specificity is preferred and the LOD differs from the clinical cutoff. Table 2 also illustrates why it may be acceptable to report below a stated LOD: the LOD is what we know about the test before we have measured any particular sample. The test result is what we know about the combination of test and sample—after we have performed a measurement (see “Reporting below LOD” below).

## 14 Effects of False Positives, False Negatives, and Call Cutoffs

When a test has a finite (nonzero) FPR or FNR values, the values of LOB, LOD, and LOQ will be different than when the FPR and FNR are zero. In general, higher rates of false positive partitions will lead to higher values of LOB for a fixed specificity. Since the cutoff values must be integers, in practice there is a quantization effect in this relationship.

As an illustration, let us consider first a test with zero FPR and FNR. In this case the LOB will be zero copies per test at 100% specificity level, the LOD will be three copies per test at the 95% sensitivity level for a call cutoff of one copy per test. (The LOB will of course also be zero at 95% specificity level.) The LOQ could be 11 copies per test at 30% CV, or perhaps, five copies per test at 45% CV. These are the best possible performance characteristics of any

test based on counting single molecules, *no matter what technology is used.*

Now let us increase the FPR to 0.01 copies per test. The LOB is zero copies but at 99% specificity level, the LOD will be 2.99 copies at 95% sensitivity. As we increase further the FPR per test, we can keep LOB at zero but will lose specificity; at the same time LOD will keep going lower for a fixed sensitivity level. We are still keeping the cutoff value at one copy per test. Refer to Fig. 5b. At some value of the FPR per test, the specificity may become unacceptably low and we may be forced to choose a different cutoff—to the next integer up. When we do this, we regain higher specificity, but have to jump to a different zone for the LOD value—as a result, there will be a significant change in the LOD value. This happens because the specificity, sensitivity, and cutoff values are interrelated, and in digital we must use an integer value for the cutoff. This “integerness” requirement leads to specific zones of operation for any digital test.

Table 3 illustrates the critical values of various parameters for selected specificity and sensitivity levels and the zones defined by the choice of cutoff level. Here is how to read it: The first column contains the value of the call cutoff. Column 2 contains the maximum value of the FP counts per test allowed so we can operate at 95% specificity. Column 3 contains the corresponding value for the LOD at 95% sensitivity (given the cutoff value and the max FPR). All units are copies per test. The last column repeats the numbers from Table 1 and illustrates the LOD level if we were operating at exactly zero FPR.

Let us look at the row for cutoff of 1. If we need to have specificity of at least 95% and use this cutoff, our test cannot have FPR of higher than 0.05 copies per test. At this FPR level, the LOD at 95% sensitivity will be 2.94 copies per test.

**Table 3**  
**Critical levels of cutoff value, FPR, and LOD for selected specificity and sensitivity values**

<b>Call cutoff (positive sample)</b>	<b>Spec 95%, Sens 95%</b>		
	<b>max FPR copies per test</b>	<b>LOD 95%</b>	<b>LOD 95% at 0 FPR</b>
1	0.05	2.94	3.00
2	0.36	4.39	4.74
3	0.82	5.48	6.30
4	1.37	6.38	7.75
5	1.97	7.19	9.15

For a test with FPR of 0.1 copies per test, in order to keep specificity above 95%, we must choose a cutoff of 2. For this cutoff, maximum FPR is 0.36 copies per test and the LOD at this FPR will be 4.39 at 95% Sensitivity (row 2). Actual specificity for a test with 0.1 FPR and cutoff of 2 copies will be higher than 95% and we can estimate the LOD 95% by subtracting 0.1 from 4.74 to get 4.64. The careful reader will notice that the values of columns 2 and 3 add up to the value of column 5 for any given cutoff value (subject to rounding effects). Intuitively, this roughly translates to the statement that “all positive calls are either a true or a false positive for any given cutoff level.”

Our last example will be a test with a FPR of one copy per test and also negligible FNR. Since we cannot choose a cutoff between three and four, we have to go to four. A sample is positive when we see at least four copies. In this case the LOB is three copies (at a level higher than 95% but we have no choice) per test. The 95% LOD is approximately 6.75 copies per test (7.75–1.00). In this framework, we call a sample positive when we measure four copies and we will correctly call a sample as positive when the expected number of copies in our test is 6.75 as drawn from the larger whole.

We need to emphasize once more that these relationships between the critical values are governed by fundamental counting statistics considerations. As long as we are trying to ascertain something about the sample *in vivo* based on small representative subsample drawn from it, they will always apply. Subsampling at the molecular level is governed by Poisson statistics and no technology or approach can do better. dPCR technology and its commercial implementation do allow us to reach very close to these physical limitations, indeed.

## 15 Additional Considerations

### 15.1 Selecting the Operating Point

We saw that the value of call cutoff, sensitivity, and specificity are interrelated. How one chooses where to operate depends on the requirement of the test: Is it more acceptable to call false positives or false negatives? A commonly used scheme is to first select the specificity level required. Together with the measured FPR (and related LOB), this will determine the choice for the analytical cutoff value. Then we evaluate the sensitivity required, which will finally determine the LOD level for the test.

Referring to Fig. 5a, we first measure the distribution of results on blank samples (left curve). We choose desired specificity,  $(1 - \beta)$ , which sets the value of the call cutoff (area under the curve to the right of the cutoff). We next select the desired Sensitivity,  $(1 - \alpha)$ . To do so, measure multiple replicates of samples at different levels of positivity to generate a family of response curves (see Fig. 5a). The curve (i.e., level of positivity) for which  $\alpha\%$  of the curve area is

to the left of the cutoff value determines the LOD at  $\alpha\%$ . Again, for quantized measurements, we may be forced to settle on particular values for specificity and sensitivity, since the cutoff must be an integer.

## **15.2 Finite Subsampling**

There are circumstances where subsampling is minimal and we therefore do not subsample a very small fraction of the whole. For example, if experiments are performed on a single cell level and most of the biological material is actually assayed. In such cases we will not suffer the subsampling penalties to the same level. We need to deal with hypergeometric distributions instead of Poisson, but such details go beyond the scope of this chapter. In the extreme case where we do measure the whole, “what we see is what was there” principle applies and the only uncertainty would come from partitioning variability and any dead volume in the system.

## **15.3 Nonstandard Variability**

We described above the cases where pure molecular stochasticity is involved—the factors we know will be always present and that we cannot control. In practice, there may be other sources of variability that could extend uncertainty. Issues like inadequate lab practices, contamination, reagent quality, and intermittent system performance may invalidate the rules described here. Regular monitoring of predicted versus measured behavior is important. This is accomplished by running controls with any unknown samples and periodic evaluation of performance on such controls. Monitoring the LOB on negative controls and verification of a Poisson distribution of the false positives is helpful as well.

## **15.4 Multitarget and Related-Target Considerations**

A very common type of diagnostic tests is where a reference target is measured together with the biomarker of interest. For example, the number of wild-type genome copies detected is used as normalization for the number of mutant target copies. In this case a natural quantity of interest is the Minor Allele Frequency (MAF), the ratio of (mutant)/(mutant + wild type) copies. Clearly, samples with a higher number of wild-type copies may show a higher number of false mutant copies. The FPR per wild-type copy, then, becomes the natural metric to watch. Careful estimation and monitoring of this rate is important to develop and apply appropriate cutoff criteria. The concepts of LOB, LOD, and LOQ, as described above, are applicable for the MAF value. In many cases, the LOD when expressed in MAF space, will still be determined by the availability of actual mutant copies. As we saw, the critical value is three mutant copies for 95% confidence, which, when divided by the wild-type copies present, will establish the critical level for the MAF at 95% confidence.

When a large number of wild-type copies are present, the cluster of partitions with both wild-type and mutant copies, i.e., double-positives, becomes more diffuse and it may be harder to

select appropriate amplitude threshold values. One technique that can minimize the false calls is to use only pure mutant partitions as a reliable source of mutant copies and compare against the double-negative partitions' count, although this will somewhat diminish the sensitivity of the test.

### **15.5 Reporting Below LOD**

Should we report a result which falls below the LOD? The nature of dPCR allows us to do this with confidence, provided that certain criteria are met.

Let us first consider what the LOD means. This is what we know about the detection capabilities of our test *before* we have performed an experiment on an unknown sample. At this stage we usually know nothing about the sample—this is why it is called unknown. *After* we have performed the measurement, we know more about the sample. The LOD applies to the test, while actual reported level applies to the sample.

Because dPCR essentially counts individual molecules, when we know that the FPR is sufficiently low, we may reliably call sample levels below formal LOD. Consider a test with FPR less than 0.05 per test and, thus, with 95% LOB of zero target copies. If we detect one or two targets, we are 95% confident that these are real and can be reported with such confidence, even though this is below the 95% LOD of three copies per test. Operating at very low FPR is always desirable from a performance guarantee perspective and allows one to report confidently below formal test LOD levels.

### **15.6 Increasing the Measured Volume**

Many of the limitations described above arise due to low levels of target molecule counts. If we subsample a small volume of the whole, we will always be limited by the molecular stochasticity processes. One way to counteract that is to increase the subsampling volume so we do not operate so close to the fundamental limit of three copies. In other words, *increasing the expected number of molecules to be detected will always help to improve test performance*. However, if preamplification is used, one has to watch for potential biases and nonuniformity and accept the likely loss of absolute quantification—a major benefit for dPCR.

---

## **16 Conclusions**

Digital PCR technology and available commercial systems and assays have reached the point where they are becoming mainstream and preferred approaches when most precise and sensitive detection of genomic targets is required. With increased availability of such solutions, the field needs to understand the issues that are specific to counting statistics so better tests can be developed, validated, and put into practice for the benefit of science and patients alike.

While we cannot beat the fundamental limits due to molecular stochasticity, dPCR technology allows us to operate at the best possible mode.

## Acknowledgments

Many thanks to my Bio-Rad colleagues George Karlin-Neumann, Dianna Maar, Xitong Li, Lucas Frenz, and Francisco Bizouarn for multiple suggestions on content and clarity of this chapter.

## References

1. Sykes PJ, Neoh SH, Brisco MJ et al (1992) Quantitation of targets for PCR by use of limiting dilution. *BioTechniques* 13(3):444–449
2. Vogelstein B, Kinzler KW (1999) Digital PCR. *Proc Natl Acad Sci U S A* 96(16):9236–9241
3. Baker M (2012) Digital PCR hits its stride. *Nat Meth* 9:541–544
4. Bhat S, Herrmann J, Armishaw P et al (2009) Single molecule detection in nanofluidic digital array enables accurate measurement of DNA copy number. *Anal Bioanal Chem* 394(2):457–467
5. Hindson BJ, Ness KD, Masquelier DA et al (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem* 83(22):8604–8610. <https://doi.org/10.1021/ac202028g>
6. Taly V, Pekin D, Benhaim L, Kotsopoulos SK et al (2013) Multiplex picodroplet digital PCR to detect KRAS mutations in circulating DNA from the plasma of colorectal cancer patients. *Clin Chem* 2013(59):1722–1731
7. Madic J, Zocevic A, Senlis V et al (2016) Three-color crystal digital PCR. Biomolecular Detection and Quantification 2016 (10):34–46. <https://doi.org/10.1016/j.bdq.2016.10.002>
8. Dube S, Qin J, Ramakrishnan R (2008) Mathematical analysis of copy number variation in a DNA sample using digital PCR on a nanofluidic device. *PLoS One* 3(8):e2876
9. Whale AS, Huggett JF, Tzonev S (2016) Fundamentals of multiplexing with digital PCR. Biomolecular Detection and Quantification. <https://doi.org/10.1016/j.bdq.2016.05.002>
10. CLSI (2012) Evaluation of detection capability for clinical laboratory measurement procedures; approved guideline - second edition. CLSI document EP17-A2. Clinical and Laboratory Standards Institute, Wayne, PA



# Chapter 4

## Control Materials and Digital PCR Methods for Evaluation of Circulating Cell-Free DNA Extractions from Plasma

Alexandra S. Whale, Ana Fernandez-Gonzalez, Alice Gutteridge, and Alison S. Devonshire

### Abstract

Cell-free DNA is an accessible source of genetic material found naturally in plasma that could be used in many diagnostic applications. Translation of cfDNA analysis methods from research laboratories into the clinic would benefit from controls for monitoring the efficiency of patient sample purification and for quality control of the whole workflow from extraction through to analysis. Here we describe two types of control materials that can be “spiked” into plasma samples to monitor and evaluate different aspects of the workflow. The first control material is an internal control that enables evaluation of extraction efficiency, fragment size bias, and sample inhibition. The second control material serves as a parallel quality control material for measurement of specific genetic targets such as tumor mutations.

**Key words** Cell-free DNA, cfDNA, Calibration, Control, Digital PCR, Droplet digital PCR, dPCR, ddPCR, Efficiency, Extraction, Plasma, Standardization

---

### 1 Introduction

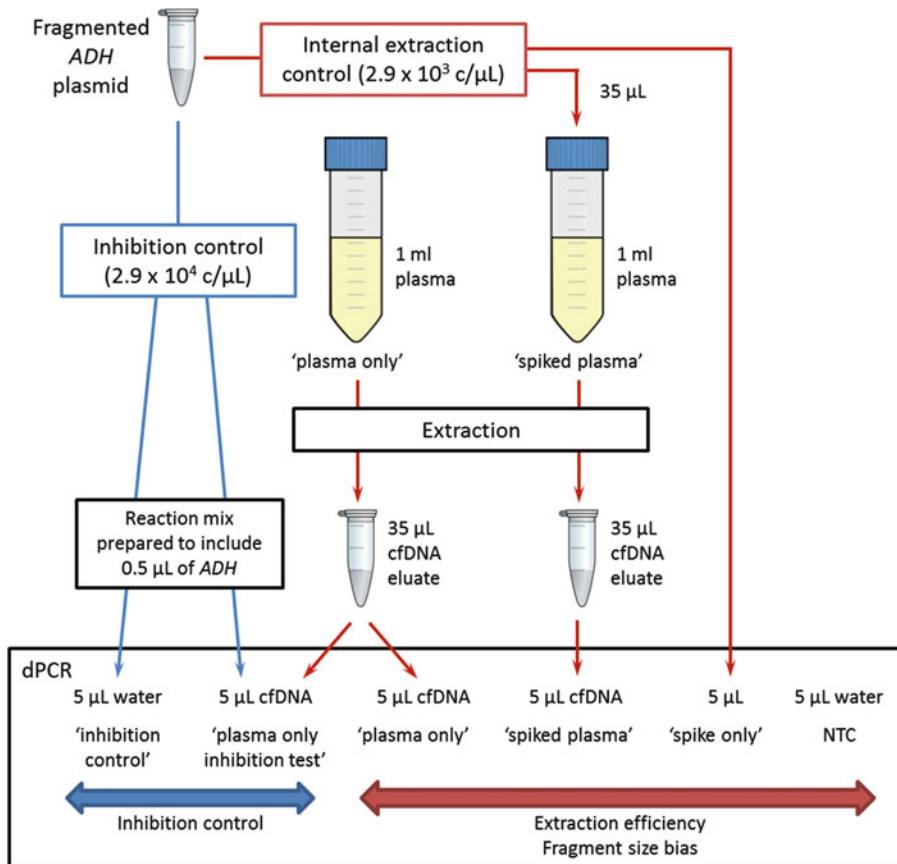
Cell-free DNA (cfDNA) is non-cellular fragmented DNA that is found naturally in body fluids such as the plasma. The vast majority of cfDNA originates from apoptotic whole blood cells and other normal living cells and is less than 200 bp in size [1–4]. However, the shedding of cfDNA can also derive from other tissue components such as fetus [5], tumor [6, 7], or graft organs [8, 9]. This presents a tremendous opportunity for human health research and diagnostic applications as it provides an accessible source of genetic material that in the past may have been difficult to obtain. Genetic alterations that can be detected include point mutations [10], single nucleotide polymorphisms [11, 12] and copy number variations [13]. Furthermore, the methylation status of cfDNA can also be measured and this has considerable potential for guiding treatment decisions [14, 15].

Although cfDNA was first isolated from plasma 60 years ago [16], several reports have highlighted a number of factors which may be hindering the translation of cfDNA into a routine analyte for clinical diagnostics [17–21]. These include a lack of appropriate controls for pre-analytical stages and for standardization of the whole process from extraction through to analysis, thereby making it difficult to compare results from different studies [17]. Such differences can be caused by choice of extraction and quantification method, the reporting metric used or the storage condition of the sample [18, 19]. Additional considerations include the assessment of PCR inhibition and fragment size bias [20] – important considerations for analysis of biological and clinical matrices given the short fragment size of cfDNA. Furthermore, cfDNA has low abundance with concentrations in the region of 1.8 to 44 ng/mL plasma reported [18]. This coupled with the fact that only a proportion of the cfDNA originates from the fetus, tumor, or graft organ being evaluated compounds the difficulty in the analytical measurement and the inherent variation associated with low level detection [22].

In this chapter we describe two different types of control material that are “spiked” into plasma samples before cfDNA extraction to enable the investigator to monitor different aspects of the analytical workflow:

1. **Internal extraction control:** DNA fragments of 115 bp, 461 bp, and 1448 bp generated by restriction digestion of the “*ADH*” plasmid, which lacks homology to human and mammalian sequences [20, 22, 23]. This control is suitable for monitoring sample-specific extraction efficiency, and identifying fragment size bias effects and the presence of inhibitor carryover from the extraction process (Fig. 1).
2. **Whole process quality control material:** fragmented human genomic DNA (gDNA), to mimic cfDNA that is found *in vivo*, with a specified abundance of the target genetic alteration to enable whole process monitoring from extraction efficiency through to analytical performance of the assay (Fig. 2).

To generate the data for this chapter, we used the QIAamp Circulating Nucleic Acid kit (Qiagen P/N 55114) and have included details specific to this kit. However, the above controls enable evaluation of the performance of the extraction process and so are not specific for a particular extraction method. A list of the currently available commercial kits developed for, or commonly applied to cfDNA extraction from plasma, can be found in Table 1. Specific details on how to perform a dPCR experiment have been omitted as these are covered in other chapters in this book (Chaps. 2 and 3) and elsewhere [20, 24].

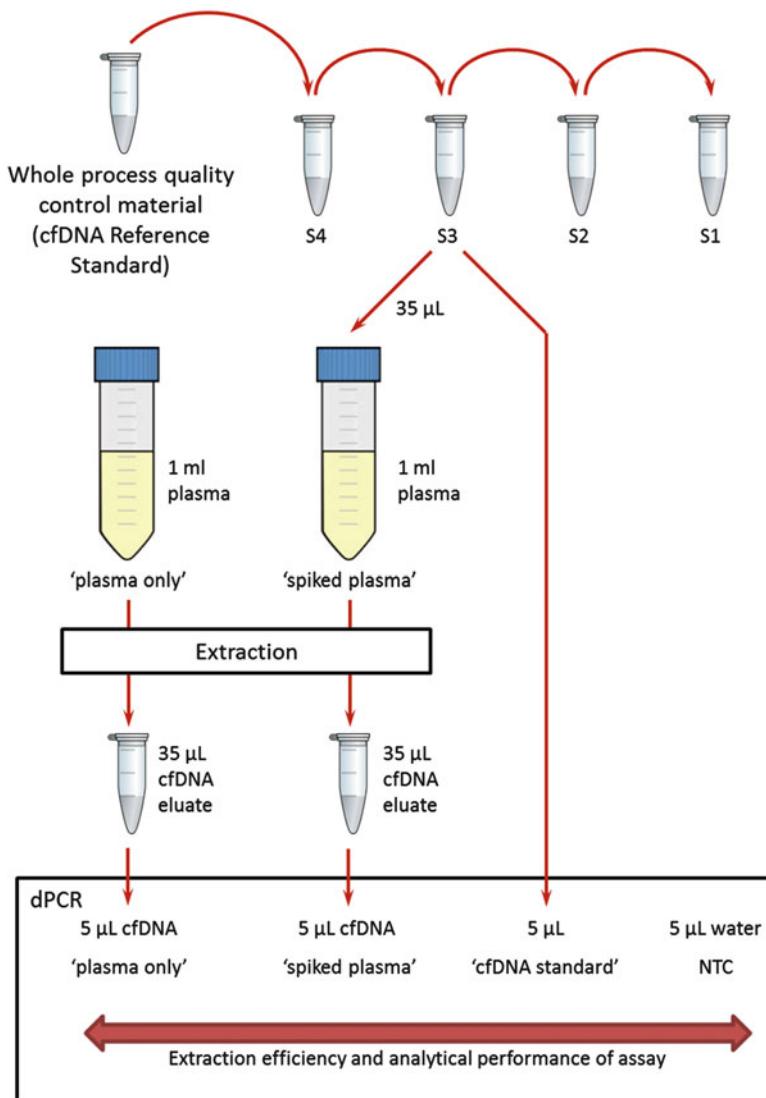


**Fig. 1** Workflow schematic for the internal extraction control. The fragmented *ADH* plasmid can be used to assess two different aspects of the cfDNA extraction: an internal extraction control that is “spiked” into the plasma sample to evaluate the extraction efficiency and fragment size bias (red workflow), and as an inhibition control that is spiked into the reaction mix (blue workflow). The different reactions are shown in the dPCR box with template volumes given as referenced in the main protocol in a total reaction volume of 20 μL

## 2 Materials

### 2.1 Internal Extraction Control

1. The restriction digested pSP64 poly(A) *ADH* plasmid (*see Note 1*) containing various fragment lengths including 115 bp, 461 bp, and 1448 bp (Fig. 3a) (*see Note 2*) is available upon request at two dilutions: approximately 2.9 × 10<sup>3</sup> *ADH* copies/μL and 2.9 × 10<sup>4</sup> *ADH* copies/μL (*see Note 3*) in non-human carrier (*see Note 4*). Store in 50 μL aliquots at –20 °C to reduce freeze–thaw effects and vortex upon thawing for 10 s to mix.
2. Oligonucleotides for *ADH*-115, *ADH*-461, and *ADH*-1448 assays (Table 2).



**Fig. 2** Workflow schematic for the whole process quality control material. The whole process quality control material can be diluted to different concentrations and then “spiked” into plasma; one dilution (S3) is shown in the workflow as used to generate Fig. 5. The different reactions are shown in the dPCR box with template volumes given as referenced in the main protocol in a total reaction volume of 20 µL

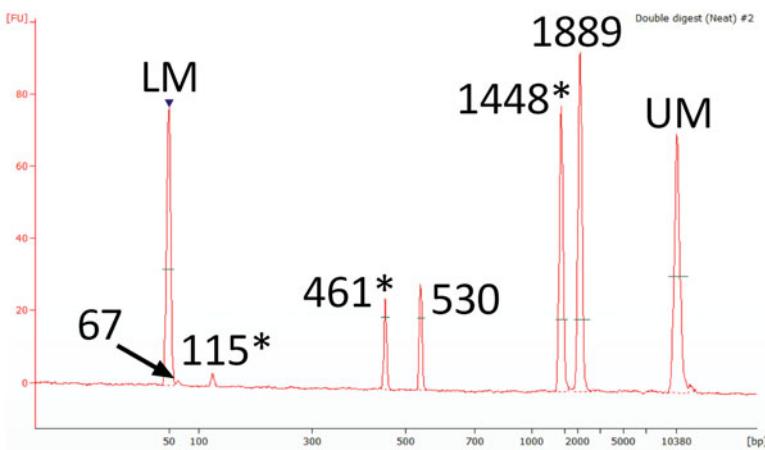
## 2.2 Whole Process Quality Control Material

- Human cfDNA Reference Standards at approximately 50 ng/µL (Horizon Diagnostics) (see Note 5). The example described in this chapter is gDNA heterozygous for *KRAS* G12C (allelic frequency of 50% verified by dPCR) prepared by sonication to a mean fragment size of ~160 bp (Fig. 3b). (Horizon Diagnostics cfDNA Reference Standards currently available: Multiplex I (8 targets) (P/N HD780) and BRAF V600E (P/N HD781)).

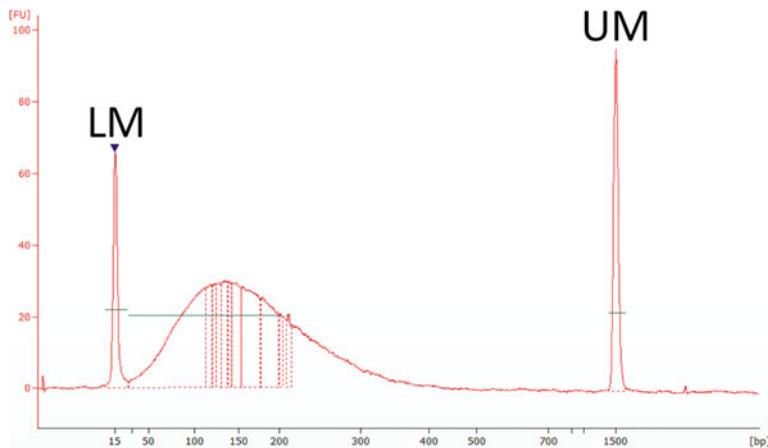
**Table 1**  
**Commercially available kits for cfDNA extraction from plasma**

Kit name	Manufacturer and part number	Purification method	Plasma volume (mL)	Elution volume (μL)	Other matrices	Citation
PME free-circulating DNA Extraction Kit	Analytik Jena 845-IR-0003010	Polymer beads	≤ 5	50	Serum, urine	
NextPrep-Mag™ cfDNA Isolation Kit	Bioo Scientific	Magnetic beads	<1 mL–3 mL	12–36	Serum	
FitAmp™ Plasma/Serum DNA Isolation Kit	Epigentek P-1004	Column	0.5	8–20	Serum and body fluids	[20, 25–28]
EpiQuik Circulating Cell-Free DNA Isolation Kit	Epigentek P-1064	Magnetic beads	0.5	20	Serum	
Nucleospin XS kit	Macherey Nagel 7409000	Column	0.24–0.72	5–30	Serum, bronchial lavage	[19, 20, 25, 29, 30]
Plasma/Serum Cell-Free Circulating DNA Purification	Norgen Bioteck Corp. 55600	Column	1–4	25	Serum	[30]
Plasma/Serum Circulating DNA Isolation	Norgen Bioteck Corp. 51200	Slurry	0.4–2	≤100	Serum	[30–32]
Chemagic Circulating NA Kit special	PerkinElmer CMG-1096	Magnetic beads	1–4	60–100		[30, 33]
Maxwell® RSC ctDNA Plasma Kit	Promega AS1480	Magnetic beads	0.2–1	50		
QIAamp Circulating Nucleic Acid kit	QIAGEN 55114	Column	1–5	20–150	Serum and urine	[20, 25, 30, 34–36]
QIAamp DNA Blood Mini kit	QIAGEN 51104	Column	0.2 (single loading)	50–200	Whole blood, body fluids	[25, 36–39]
MagMAX™ Cell-Free DNA Isolation Kit	Thermo Fisher Scientific	Magnetic beads	0.5–10	15–50	Liquid samples (serum, plasma)	
Quick-ctDNA™ Serum & Plasma Kit	Zymo Research D4076	Column	≤10	35–50	Serum, amniotic fluid, and cerebrospinal fluid	

As of March 2016, these were the commercially available kits; however, this list is unlikely to be exhaustive. At the time of compilation, new kits were being brought to market on a regular basis and so may not be included in this table. For the same reason, at the time of this publication to our knowledge, not all of the kits had been cited in peer-review publications

(A) Fragmented *ADH* plasmid

## (B) Human cfDNA Reference Standards



**Fig. 3** Example of electropherograms showing the fragment size of the extraction controls. Control samples were analyzed using the 2100 Bioanalyzer with the DNA 1000 series II kit (Agilent). The lower and upper markers are indicated as LM and UM, respectively. (a) Electropherogram of the internal extraction control showing the peaks corresponding to the six fragments generated from the restriction digest of the linearized *ADH* plasmid (67, 115, 461, 530, 1448, and 1889 bp). The corresponding *ADH* assays that target the different fragment sizes are indicated with an asterisk (\*). (b) Electropherogram of the human cfDNA Reference Standards fragmented to a mean size of 160 bp

2. Non-human carrier RNA. For the example given in this study, yeast tRNA (Sigma-Aldrich P/N R3508 at 10 ng/mL) was diluted in 1 × TE to a final concentration of 5 ng/μL (see Note 4).
3. 1× Tris-EDTA (TE), buffer pH 8.0, (molecular biology grade), for example Fluka P/N 93283.

**Table 2**  
***ADH* assay oligonucleotide sequences**

Assay name	ADH fragment size	Oligonucleotide sequence (5'-3')	[20× stock] (μM)	[Final reaction] (μM)
<i>ADH</i> -115	115 bp	F: GGGCCGAGCGCAGAA R: ACTCTAGCTTCCC GGCAACA P: HEX-TGGTCCTGCAACTTATCCG CCTCC-BHQ1	18 18 5	0.9 0.9 0.25
<i>ADH</i> -461 ( <i>Adh</i> β)	461 bp	F: TTGAGAGTGTGGAGAAGGGAGTGA R: CGGTAAAGATCGGCAACACA P: FAM-TCTTCAGGCCAGGAGATC-MGB	18 18 4	0.9 0.9 0.2
<i>ADH</i> -1448 ( <i>Adh</i> δ)	1448 bp	F: TGAACCCGAAAGACCATGACA R: CCCACCATCCGT CATCTCA P: FAM-CCAATTCAACAGGTGATC-MGB	18 18 4	0.9 0.9 0.2

It is recommended that all primers and probes are HPLC-purified. For the example given in this study, the fluorophore is given in the table; however, these probes have been published conjugated with a number of different fluorophore dyes depending on the experimental set up [20, 22, 23]. Key: *bp* base pairs, *F* forward primer, *R* reverse primer, *P* probe, *HEX* Hexachlorofluorescein, *FAM* 6-carboxyfluorescein, *BHQ1* black hole quencher 1, *MGB* minor groove binder (Life Technologies)

4. Specific assay to detect the genetic alteration present in whole process quality control material. In this chapter, the PrimePCR™ ddPCR™ Mutation Assays (Bio-Rad): *KRAS* WT for p.G12C, Human (P/N 10031249, dHsaCP2000008) and *KRAS* p.G12C, Human (P/N 10031246, dHsaCP2000007) were used in a duplex reaction to quantify the *KRAS* G12C mutation.

### 2.3 cfDNA Extraction from Plasma

1. Blood samples collected in EDTA K2 vacutainers followed by plasma separation using standard methods (see Note 6). Prepare 1.1 mL plasma aliquots and freeze at –80 °C (see Note 7).
2. 0.5 mL and 1.5 mL microfuge tubes (see Note 8).
3. Vortex mixer.
4. Centrifuge and microfuge.
5. Extraction reagents or kit (Table 1). For the experimental workflow illustrated in this book chapter, the QIAamp® Circulating Nucleic Acid Kit (Qiagen P/N 55114) was used.

### 2.4 Digital PCR Analysis

1. Digital PCR system: the worked examples in this chapter were performed using the QX200™ Droplet Digital™ PCR System with the QuantaSoft™ analysis program (Bio-Rad). This protocol can be adapted for other digital PCR platforms.

2. Compatible mastermix for the selected digital PCR system: ddPCR™ Supermix for Probes (Bio-Rad P/N 186-3010) and ddPCR™ Supermix for Probes (No dUTP) (Bio-Rad P/N 186-3024) were used to generate the results shown in this chapter for the internal extraction control and whole process quality control, respectively.
3. Nuclease-free water (molecular biology grade), for example, Ambion™ (P/N AM9937).

---

### 3 Methods

#### 3.1 Internal Extraction Control: Fragmented ADH Plasmid

##### 3.1.1 Spiking and Extraction of Internal Extraction Control from Plasma

This protocol is designed for the assessment of extraction efficiency, presence of downstream inhibitors and fragment size bias. This protocol is described for extraction of cfDNA from 1 mL plasma, however, the principles hold for smaller or larger volumes of plasma. A schematic illustrating the work-flow and the different uses for the internal extraction control is provided (Fig. 1). Spike-in volumes can be scaled up or down according to the volume of plasma processed.

1. Thaw the required 1.1 mL plasma aliquots at room temperature and maintain on ice thereafter (*see Note 7*).
2. Perform an additional centrifugation step at  $1000 \times g$  for 10 min at 4 °C (*see Note 9*).
3. Remove 1 mL of plasma supernatant into a fresh tube for the first stage of the extraction protocol (a 50 mL tube in the case of the QIAamp® Circulating Nucleic Acid Kit).
4. Add 35 µL of the diluted fragmented *ADH* plasmid ( $2.9 \times 10^3$  copies/µL stock) per 1 mL plasma to generate the “spiked plasma” sample (*see Note 10*). Vortex to mix. Ensure that remaining plasmid is kept for dPCR analysis as the “spike only” control.
5. Prepare a “plasma only” control for each plasma sample containing 1 mL of plasma but no fragmented *ADH* plasmid (*see Note 11*).
6. Follow preferred extraction method (examples given in Table 1) according to the manufacturer’s specifications (*see Note 12*).
7. Elute cfDNA in 35 µL elution buffer (*see Note 10*).
8. Store eluted cfDNA at 4 °C for up to 1 month and –20 °C for longer term.

**Table 3**  
Preparation of reactions for analysis with the *ADH* assays

Component	Unit	[stock]	<i>ADH-461 and ADH-115 duplex assay</i>		<i>ADH-1448 assay</i>	
			1× [work]	Volume (μL)	1× [work]	Volume (μL)
Nuclease-free water				3.00	4.00	
ddPCR™ Supermix for probes	×	2	1	10.00	1	10.00
<i>ADH</i> -461 assay (FAM)	×	20	1	1.00	—	—
<i>ADH</i> -1448 assay (FAM)	×	20	—	—	1	1.00
<i>ADH</i> -115 assay (HEX)	×	20	1	1.00	—	—
cfDNA extract				5.00	5.00	
Total				20.00	20.00	

For the setup, each *ADH* assay was prepared as a 20× stock and stored in aliquots at –20 °C to reduce freeze–thaw effects. *ADH* assays were conjugated to FAM or HEX fluorophore dyes as indicated.

**Table 4**  
Preparation of reactions for inhibition analysis with the *ADH-461* assay

Component	Unit	[stock]	<i>ADH-461 assay</i>	
			1× [work]	Volume (μL)
Nuclease-free water				3.50
ddPCR™ Supermix for probes	×	2	1	10.00
<i>ADH</i> -461 assay (FAM)	×	20	1	1.00
Fragmented <i>ADH</i> plasmid	c/μL	$2.9 \times 10^4$		0.50
cfDNA extract (plasmid only)				5.00
Total				20.00

*3.1.2 Digital PCR Analysis of Internal Extraction Control*

1. Prepare each reaction for dPCR as per manufacturer’s instructions using 5 μL of template per 20 μL reaction for each *ADH* assay as shown in Table 3 (see Note 13).
2. Prepare triplicate reactions (see Note 14) for each sample extract from “spiked plasma” and corresponding “plasma only” control (if applicable).
3. Prepare additional triplicate reactions (see Note 16) for the “spike only” control (no extraction performed) and no-template controls (NTCs) using water in place of template.
4. To test for inhibitors carried over from the extraction process, prepare additional reactions with the fragmented *ADH* plasmid ( $2.9 \times 10^4$  *ADH* copies/μL stock) added directly to the reaction mix as shown in Table 4. For these reactions, only the

“plasma only” samples can be analyzed. Prepare additional reactions for the “inhibition control” using water in place of cfDNA template (will only contain fragmented *ADH* plasmid from the reaction mix).

5. Partition the reaction as per manufacturer’s instructions.
6. Perform PCR using the following cycling conditions: Enzyme activation for 10 min at 95 °C, 40 cycles of denaturation for 30 s at 94 °C and annealing/extension for 1 min at 60 °C, followed by signal stabilization for 10 min at 98 °C and then hold at 4 °C. Follow manufacturer’s instructions for the cycling, for example, using the QX200™ Droplet Digital™ PCR System requires all temperature ramping to be set at 2 °C/s.
7. Analyze the cycled partitions and ensure the thresholds are set correctly to separate the positive and negative partitions (Figs. 4a–c) (*see Note 15*).
8. Check that the “plasma only” controls (if applicable) and NTCs have no positive partitions (*see Note 16*).
9. Calculate the concentration (copies/μL) for each “spiked plasma” and “spike only” samples for the three *ADH* assays using the digital platform software or previously published equations for digital PCR analysis [23, 40].
10. Calculate the extraction efficiencies for the different *ADH* fragment sizes (*see Notes 17 and 18*) using the following equation:  
Calculation of extraction efficiency of the fragmented *ADH* plasmid

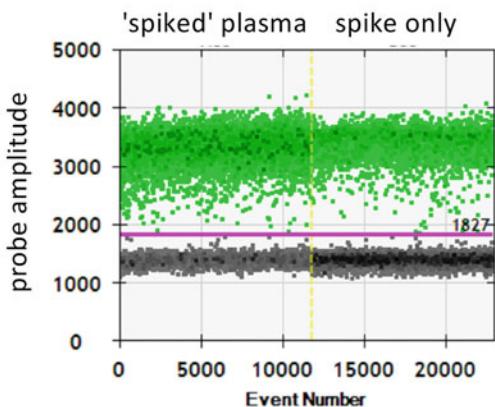
$$\text{Extraction efficiency (\%)} = \frac{[\text{ADH}_{\text{spiked plasma}}]}{[\text{ADH}_{\text{spike only}}]} \quad (1)$$

11. Plot the extraction efficiencies of the three *ADH* fragments and perform ANOVA statistical analysis to identify any fragment size bias (Fig. 4d).
12. For the inhibition analysis, calculate the difference in concentration of the “plasma only” samples using the reactions containing the fragmented *ADH* plasmid, “plasma only inhibition test” (Table 4) with the “inhibition control” using the following equation:

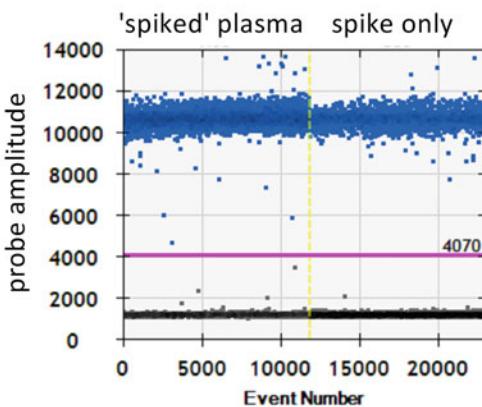
Calculation of inhibition

$$\text{Relative inhibition} = \frac{[\text{ADH-461}_{\text{plasma only inhibition test}}]}{[\text{ADH-461}_{\text{inhibition control}}]} \quad (2)$$

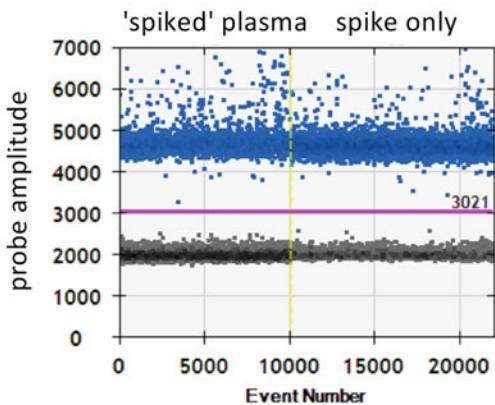
(A) ADH-115 assay



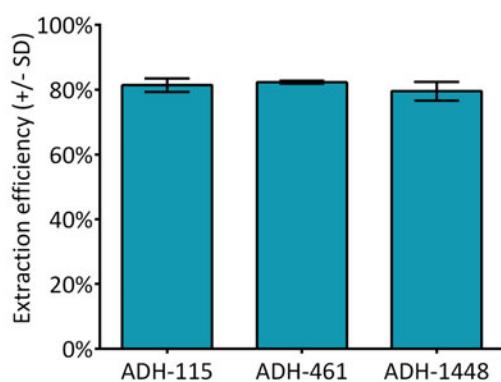
(B) ADH-461 assay



(C) ADH-1448 assay



(D) Extraction efficiency



**Fig. 4** Example of results obtained using the internal extraction control. Plasma was extracted using the QIAamp® Circulating Nucleic Acid Kit. The QX200™ Droplet Digital™ PCR System with the QuantaSoft™ analysis program (Bio-Rad) was used to generate the data. (a–c) Examples of the QuantaSoft™ generated one-dimensional scatter plots of the *ADH* assays for the “spiked” plasma and spike only samples. The x-axis shows the partition count (event number) with the y-axis showing the fluorescent amplitude of the probe. The horizontal pink line represents the threshold that defines the positive (above) and negative (below) partitions. Positive partitions are either blue (FAM probe) or green (HEX probe) while negative partitions are grey. (d) Bar graph showing the extraction efficiency of each assay as calculated with Eq. 1. The error bars represent the standard deviation between triplicate dPCR measurements

### 3.2 Whole Process

#### Quality Control

#### Material

This protocol illustrates how the whole process from extraction through to dPCR analysis can be tested across a range of concentrations of the analyte (in the example, the analyte is the *KRAS* G12C sequence found in plasma). A schematic illustrating the work-flow for the whole process quality control material is provided (Fig. 2). As with the protocol described in Sect. 3.1, this protocol is written for extraction of cfDNA from 1 mL plasma, however, the

**Table 5**  
**Preparation of a dilution series of cfDNA Reference Standards from stock solution**

Dilution ID	[cfDNA Standard] (ng/ $\mu$ L)	[cfDNA Standard] (KRAS G12C copies/ $\mu$ L) <sup>a</sup>	[cfDNA Standard] (KRAS G12C copies/mL plasma) <sup>b</sup>	Dilution factor	Sample volume ( $\mu$ L)	Diluent volume ( $\mu$ L)	Final volume ( $\mu$ L)
Stock	50	~4250					
S4	2.3	195	6840	21.3	25.0	507	432.3
S3	0.47	37	1306	5.0	100	400	400
S2	0.094	9	312	5.0	100	400	400
S1	0.019	2	69	5.0	100	400	500

Avoid using small volumes (<10  $\mu$ L) in dilution series as precision of pipetting is lower for small volumes. Mix each dilution by vortexing and brief centrifugation to pool contents before subsequent dilution. Sufficient diluted cfDNA Reference Standard should be prepared for spiking into all plasma extractions and for the “spike only” dPCR analysis

<sup>a</sup>As measured by dPCR. <sup>b</sup>Based on addition of 35  $\mu$ L cfDNA Reference Standard/mL plasma

principles hold for smaller or larger volumes of plasma and spike-in volumes can be scaled up or down according to the volume of plasma processed.

### 3.2.1 Spiking and Extraction of cfDNA Reference Standards from Plasma

1. Prepare a dilution series of cfDNA Reference Standards in 5 ng/ $\mu$ L non-human carrier RNA. An example of a dilution series is given in Table 5 that is suitable for testing the linearity of extraction efficiency and analytical sensitivity of cfDNA extractions. Store the dilutions in 50  $\mu$ L aliquots at 4 °C for short term (<1 month) and –20 °C for longer term.
2. Thaw a 1.1 mL plasma aliquot at room temperature and maintain on ice thereafter (see Note 7).
3. Perform additional centrifugation step at 1000  $\times g$  for 4 min at 4 °C (see Note 8).
4. Remove 1 mL of plasma supernatant into a fresh tube for the first stage of the extraction protocol (a 50 mL tube in the case of the QIAamp® Circulating Nucleic Acid Kit).
5. Add 35  $\mu$ L of the diluted cfDNA Reference Standard (Table 5) to 1 mL plasma in a fresh 1.5 mL tube (see Note 10). Vortex to mix and pulse down. Ensure that remaining cfDNA Reference Standard dilutions are kept for analysis as the “cfDNA standard” control.
6. Prepare a “plasma only” control for each plasma sample containing 1 mL of plasma but no cfDNA Reference Standard (see Note 19).

**Table 6****Preparation of reactions for analysis with the PrimePCR™ ddPCR™ Mutation Assays (Bio-Rad)**

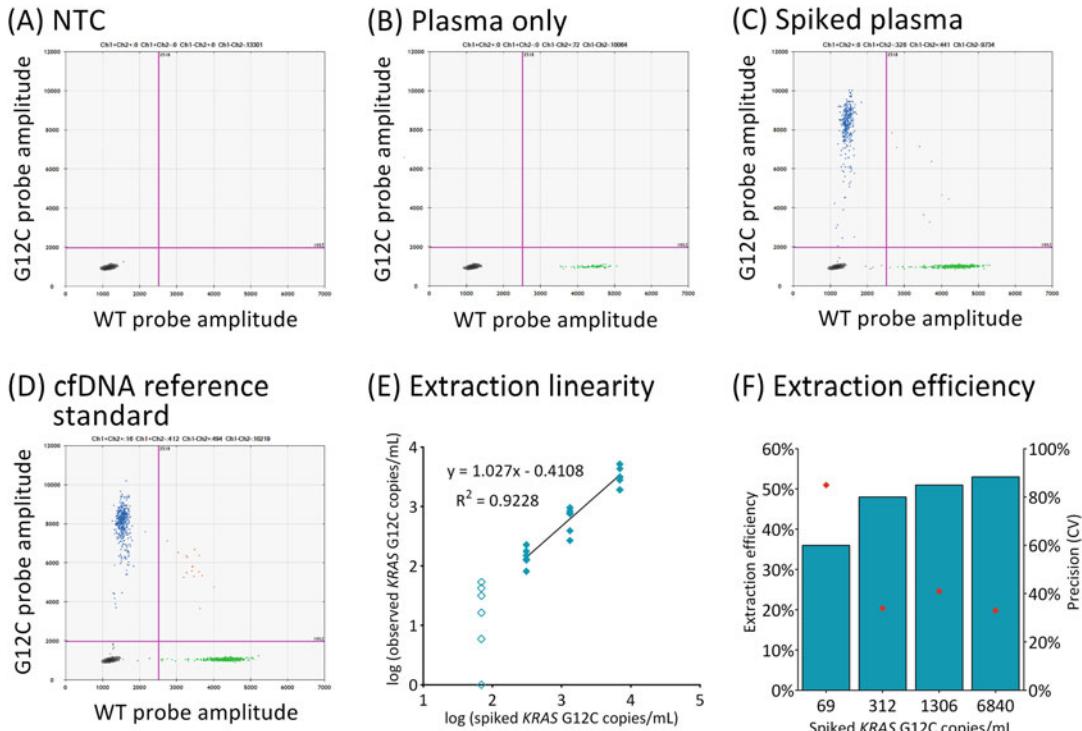
<b>Component</b>	<b>Unit</b>	<b>[stock]</b>	<b>1 × [working]</b>	<b>Volume (µL)</b>
Nuclease-free water				3.00
ddPCR™ Supermix for probes (no dUTP)	×	2	1	10.00
PrimePCR KRAS p.G12C (FAM)	×	20	1	1.00
PrimePCR KRAS WT for p.G12C (HEX)	×	20	1	1.00
cfDNA extract				5.00
Total				20.00

“The PrimePCR assays were conjugated to FAM (6-carboxyfluorescein) or HEX (Hexachlorofluorescein)” fluorophore dyes as indicated

7. Follow preferred extraction method (examples given in Table 1) according to the manufacturer’s specifications (*see Note 12*).
8. Elute cfDNA in 35 µL elution buffer (*see Note 10*).
9. Store eluted cfDNA at 4 °C for up to 1 month and –20 °C for longer term.

### 3.2.2 Digital PCR Analysis of cfDNA Reference Standards

1. Prepare each reaction for dPCR as per manufacturer’s instructions using 5 µL of template per 20 µL reaction using the assay to detect the genetic alteration present in human cfDNA Reference Standards. In the example, a duplex reaction containing the assays for detecting the KRAS WT and KRAS G12C sequence were used (Table 6).
2. Prepare triplicate reactions for each sample extract from “spiked plasma” and the corresponding “plasma only” control (*see Note 14*).
3. Prepare additional triplicate reactions to analyze the cfDNA Reference Standard (“cfDNA standard”) dilutions (no extraction performed) and no template controls (NTCs) using water in place of template.
4. Partition the reaction as per manufacturer’s instructions.
5. Perform PCR using the cycling conditions recommended by the manufacturer for the PrimePCR assay: Enzyme activation for 10 min at 95 °C, 40 cycles of denaturation for 30 s at 94 °C and annealing/extension for 1 min at 55 °C, followed by signal stabilization for 10 min at 98 °C and then hold at 4 °C. For all stages the ramp rate is set at 2 °C/s.



**Fig. 5** Example of results obtained using the Human KRAS G12C cfDNA Reference Standard as a whole process quality control material. Plasma was extracted using the QIAamp® Circulating Nucleic Acid Kit. The QX200™ Droplet Digital™ PCR System with the QuantaSoft™ analysis program (Bio-Rad) was used to generate the data. (a–d) Examples of the QuantaSoft™ generated two-dimensional scatter plots of the (a) NTC, (b) “plasma only,” (c) “spiked” plasma (dilution S3 as detailed in Table 5) and (d) cfDNA Reference Standard for the duplex KRAS G12C assay (also dilution S3). The fluorescent amplitude for each probe is shown on the x-axis (WT) and y-axis (G12C). The vertical and horizontal pink lines represent the two thresholds that define the positive and negative partitions. Positive partitions are either blue (G12C; FAM probe) or green (WT; HEX probe) while negative partitions are grey. Double positive partitions, containing both G12C and WT sequences are orange. (e–f) Analysis of the extraction of the four dilutions of the KRAS G12C cfDNA Reference Standard (as given in Table 5). Six replicate extractions were performed for each dilution with a single dPCR measurement. (e) Correlation analysis demonstrated good linearity down to log 2.5 (312 copies/mL plasma; equivalent to 45 copies per dPCR). The limit of detection (LOD) for the whole process was approximately log 1.8 (69 copies/mL plasma; equivalent to nine copies per dPCR) with KRAS G12C undetectable in 1 of 6 extracts analyzed. (f) Graph showing the extraction efficiency (bars) plotted on the left y-axis as calculated with Eq. 2 with the associated precision (red diamonds) plotted on the right y-axis

6. Analyze the cycled partitions and ensure the thresholds are set correctly to separate the positive and negative partitions (Figs. 5a–d) (*see Note 15*).
7. Check the NTCs have no positive partitions (Fig. 5a).
8. Check the “plasma only” controls for positive partitions that target the analyte of interest. In the example, the “plasma only” control is from a normal healthy blood donor and so should have no positive partitions for the KRAS G12C point

mutation. The *KRAS* WT sequence will be present in the endogenous cfDNA (see Notes 16 and 20) (Fig. 5b).

9. Calculate the concentration (copies/ $\mu$ L) for each “plasma only,” “spiked plasma” (Fig. 5c), and “spike only” samples (Fig. 5d) for the *KRAS* G12C and *KRAS* WT using the digital platform software or previously published equations for digital PCR analysis [23, 40]. The observed concentration of the spiked *KRAS* G12C DNA can be expressed as copies per mL plasma by multiplying the concentration (copies/ $\mu$ L) by the elution volume (35  $\mu$ L) (Fig. 5e).
10. Calculate the extraction efficiency using the following equations (see Note 20):

Calculation of extraction efficiency of mutant *KRAS* G12C sequences

$$\text{Extraction efficiency (\%)} = \frac{[\text{KRAS G12C}_{\text{spiked plasma}}]}{[\text{KRAS G12C}_{\text{cfDNA standard}}]} \quad (3)$$

Calculation of extraction efficiency of *KRAS* WT sequences

$$\text{Extraction efficiency (\%)} = \frac{[\text{KRAS WT}_{\text{spiked plasma}}] - [\text{KRAS WT}_{\text{plasma only}}]}{[\text{KRAS WT}_{\text{cfDNA standard}}]} \quad (4)$$

11. Plot the extraction efficiencies of the plasma samples (Fig. 5f).

#### 4 Notes

1. The pSP64 poly(A) *ADH* plasmid is a 4510 bp plasmid that contains a portion of the *Arabidopsis thaliana* alcohol dehydrogenase (*ADH*) gene cloned into the pSP64 poly(A) plasmid [22]. The cloned *ADH* sequence is nucleotides 774 to 2274 of the GenBank sequence M12196 and encodes the “landsberg” allele.
2. Fragmentation was achieved as described in Devonshire et al., 2014 [20]. Briefly, 1.5  $\mu$ g of pSP64 poly(A) *ADH* was linearized with 10 units of *Bgl*II in NEBuffer 2 in a 50  $\mu$ L reaction at 37 °C for 2 h. Fragmentation of 0.75  $\mu$ g (25  $\mu$ L) of the linearization reaction was performed with 10 units *Alw*NI, 2 units *Bsr*DI and 5  $\mu$ g BSA in NEBuffer 2 in a 50  $\mu$ L reaction at 37 °C for 1 h followed by 65 °C for 1 h and enzyme inactivation at 80 °C for 20 min. All restriction endonucleases and buffers were purchased from NEB.

3. The concentration of the fragmented *ADH* plasmid is approximately 15 ng/ $\mu$ L as determined by fluorimetry (Qubit® 2.0 fluorometer with the dsDNA HS Assay Kit (Invitrogen™)). 1 ng of the fragmented *ADH* plasmid is approximately  $2.015 \times 10^8$  copies.
4. Alternative non-human carrier molecules can be used, however, the investigator will need to verify that no cross-reactivity exists between the analytical assay and carrier molecules. The available fragmented *ADH* plasmid is diluted in sonicated salmon sperm DNA (ssDNA) (Agilent P/N 201190 at 10  $\mu$ g/ $\mu$ L) to a final concentration of 50 ng/ $\mu$ L.
5. A list of the currently available cfDNA Reference Standards can be found at <https://www.horizondiscovery.com/reference-standards/our-formats/cfdna>.
6. Plasma from normal healthy individuals in the described protocol was purchased from SeraLab Ltd. (UK), which was prepared using standard procedures as follows. Whole blood from normal healthy individuals was drawn into Vacutainer tubes containing anticoagulants (EDTA) and mixed pursuant to BD specifications and allowed to sit at room temperature for 30 min. Plasma was separated by refrigerated centrifugation at  $1303 \times g$  for 20 min and then removed by aspiration, ensuring no Red Blood Cell contamination, into a fresh tube prior to shipment. Shipment should be on dry ice and plasma stored at  $-80$  °C upon arrival. When thawed, plasma should be a clear liquid. A cloudy appearance indicates protein or cell lysis contamination; such samples should be discarded.
7. Plasma is a biohazard as it may contain infectious agents. It is advisable to handle all plasma samples in a Class II Biological Safety Cabinet until lysis or proteinase treatment in buffers containing chaotropic agents is complete. The risk of spillages can be reduced whilst mixing by sealing tubes with Parafilm or in double bags. Plasma should be homogenized by mixing on a rotator at 4 °C for 30 min prior to aliquoting.
8. Wherever possible it is recommended to use tubes suitable for low concentration nucleic acid samples, for example Lo-bind® tubes (Eppendorf).
9. A further centrifuge step at  $1000 \times g$  for 10 min at 4 °C to remove any residual precipitated material or cellular debris can also be performed [25].
10. For the experiment set up described in this chapter, 35  $\mu$ L of the diluted fragmented *ADH* plasmid was spiked into each plasma sample extraction to give a total of  $1 \times 10^5$  copies per extraction ( $2.9 \times 10^3$  copies  $\times$  35  $\mu$ L =  $1 \times 10^5$  copies). This protocol is developed so that the volume spiked into the plasma is the same as the elution volume. This enables a more

straightforward calculation of the extraction efficiency and recovery yield. Different spike and elution volumes can be used, but these need to be accounted for in the calculations. It is advisable to avoid using small volumes (<10 µL) as the precision of pipetting is lower for small volumes.

11. “Plasma only” extractions are included as a control for cross-contamination at the extraction stage. “Plasma only” extracts can also be used for inhibition testing (Table 4).
12. Several of the example methods in Table 1 are readily scalable and can extract cfDNA from 0.2 mL up to 10 mL plasma. This protocol is developed using 1 mL plasma with 35 µL of internal control material extracted using the QIAamp® Circulating Nucleic Acid Kit and eluted 35 µL; the volume and concentration of the internal control material and elution volumes should be scaled for larger/smaller volumes accordingly.
13. There are three *ADH* assays that target the different fragment sizes of the fragmented *ADH* plasmid: *ADH*-115 (115 bp fragment), *ADH*-461 (461 bp fragment), *ADH*-1448 (1448 bp fragment). To evaluate the fragment size bias of the extraction method all three assays can be used in parallel. The example set up described here shows a duplex reaction with the *ADH*-115 and *ADH*-461 assays and a separate reaction with the *ADH*-1448 assay. These assays can be used in duplex or uniplex reactions with changes in the conjugated fluorophore, however, investigators will need to validate any changes made to the published assays.
14. It is preferable to prepare triplicate reactions as technical replicates (set up three separate reactions with mastermix and template) instead of preparing a triple volume reaction and aliquoting into three separate reaction wells as this will capture the technical error of the experiment set up in addition to the instrument error.
15. It is important to set the thresholds that separate the positive and negative partitions globally across all reactions for an assay. Where possible, it is advisable to classify each droplet into a category; in the given example, this was achieved by selecting all wells and using the “cross-hair” tool in the QuantaSoft™ analysis program for the QX200™ Droplet Digital PCR System.
16. The identification of positive partitions in the NTC or “plasma only” control for an assay that targets a sequence unique to the control material indicates either (1) nonspecificity of the assay or (2) a background source of contamination. Both scenarios should be investigated by comparison with NTC and “carrier only” reactions to identify non-specificity or sources of contamination. To determine if the non-specific amplification

originates from the endogenous cfDNA extracted from the plasma samples, control reactions can be performed using genomic DNA as a template; presence of positive partitions indicates non-specific amplification of the *ADH* assays or assay to detect the genetic alteration present in the human cfDNA Reference Standards.

17. As the *ADH* assays target a sequence present in the fragmented *ADH* plasmid, the recovery of the fragmented *ADH* plasmid can be calculated directly without subtraction of the “plasma only” control (*see Note 15* to confirm specificity).
18. If the elution volume does not equal the spiked volume then a modified version of Eq. 1 can be used:

$$\text{Extraction efficiency (\%)} = \frac{[\text{Adh}_{\text{spiked plasma}}] \times V_e}{[\text{Adh}_{\text{spike only}}] \times V_s}$$

where  $V_e$  and  $V_s$  are the elution and spiked volumes, respectively.

19. The presence of positive partitions for the *KRAS* G12C mutation in the “plasma only” sample may indicate the presence of the mutation in the endogenous cfDNA. This should be verified with further analysis of separate plasma extractions. The false positive rate of the *KRAS* G12C assay should also be characterized using gDNA which is 100% wild-type for the mutation of interest [41], for the example *KRAS* Wild Type Reference Standard (Horizon Discovery P/N HD710).
20. Selection of a target (*KRAS* G12C) that is only present in the cfDNA Reference Standard allows calculation of the recovery yield of the spiked DNA without subtraction of the “plasma only” control. Selection of a target that is present in both the plasma and cfDNA Reference Standard needs the “plasma only” control concentration to be subtracted.

## Acknowledgment

The work described in this chapter was in part funded by the UK National Measurement System (NMS) and Innovate UK under the project “Enabling stratified medicine through cell and tissue reference standards for minimally invasive cancer testing” (Project Number: 101862) in partnership with Horizon Discovery. The remaining funding was from the European Metrology Research Programme (EMRP) joint research project [SIB54] “Bio-SITrace” (<http://biositrace.lgcgroup.com>) which is jointly funded by the EMRP participating countries within EURAMET and the European Union. The authors would like to acknowledge Dr. Karin Schmitt and Dr. Hadas Raveh-Amit at Horizon

Diagnostics for providing the Human cfDNA Reference Standard material, Dr. Tim Forszew at UCL for support with dPCR testing of the *ADH* internal extraction control, and Dr. George Karlin-Neumann and Dr. Svilen Tzonev at Bio-Rad for their assistance in the development of ddPCR experiments.

## References

1. Stroun M, Lyautey J, Lederrey C, Olson-Sand A, Anker P (2001) About the possible origin and mechanism of circulating DNA: apoptosis and active DNA release. *Clinica Chimica Acta* 313(1-2):139–142. [https://doi.org/10.1016/S0009-8981\(01\)00665-9](https://doi.org/10.1016/S0009-8981(01)00665-9)
2. Lui YY, Chik KW, Chiu RW, Ho CY, Lam CW, Lo YM (2002) Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. *Clin Chem* 48(3):421–427
3. Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, Knippers R (2001) DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res* 61(4):1659–1665
4. van der Vaart M, Pretorius PJ (2007) The origin of circulating free DNA. *Clin Chem* 53(12):2215. <https://doi.org/10.1373/clinchem.2007.092734>
5. Lo YM, Corbett N, Chamberlain PF, Rai V, Sargent IL, Redman CW, Wainscoat JS (1997) Presence of fetal DNA in maternal plasma and serum. *Lancet* (London, England) 350(9076):485–487. [https://doi.org/10.1016/s0140-6736\(97\)02174-0](https://doi.org/10.1016/s0140-6736(97)02174-0)
6. Chen XQ, Stroun M, Magnenat JL, Nicod LP, Kurt AM, Lyautey J, Lederrey C, Anker P (1996) Microsatellite alterations in plasma DNA of small cell lung cancer patients. *Nat Med* 2(9):1033–1035
7. Nawroz H, Koch W, Anker P, Stroun M, Sidransky D (1996) Microsatellite alterations in serum DNA of head and neck cancer patients. *Nat Med* 2(9):1035–1037
8. Lo YM, Tein MS, Pang CC, Yeung CK, Tong KL, Hjelm NM (1998) Presence of donor-specific DNA in plasma of kidney and liver-transplant recipients. *Lancet* (London, England) 351(9112):1329–1330
9. Lui YY, Woo KS, Wang AY, Yeung CK, Li PK, Chau E, Ruygrok P, Lo YM (2003) Origin of plasma cell-free DNA after solid organ transplantation. *Clin Chem* 49(3):495–496
10. Taly V, Pekin D, Benhaim L, Kotsopoulos SK, Le Corre D, Li X, Atochin I, Link DR, Griffiths AD, Pallier K, Blons H, Bouche O, Landi B, Hutchison JB, Laurent-Puig P (2013) Multiplex Picodroplet digital PCR to detect KRAS mutations in circulating DNA from the plasma of colorectal cancer patients. *Clin Chem* 59(12):1722–1731. <https://doi.org/10.1373/clinchem.2013.206359>
11. Beck J, Bierau S, Balzer S, Andag R, Kanzow P, Schmitz J, Gaedcke J, Moerer O, Slotta JE, Walson P, Kollmar O, Oellerich M, Schutz E (2013) Digital droplet PCR for rapid quantification of donor DNA in the circulation of transplant recipients as a potential universal biomarker of graft injury. *Clin Chem* 59(12):1732–1741. <https://doi.org/10.1373/clinchem.2013.210328>
12. Snyder TM, Khush KK, Valentine HA, Quake SR (2011) Universal noninvasive detection of solid organ transplant rejection. *Proc Natl Acad Sci U S A* 108(15):6229–6234. <https://doi.org/10.1073/pnas.1013924108>
13. Shaw JA, Page K, Blighe K, Hava N, Guttery D, Ward B, Brown J, Ruangpratheep C, Stebbing J, Payne R, Palmieri C, Cleator S, Walker RA, Coombes RC (2012) Genomic analysis of circulating cell-free DNA infers breast cancer dormancy. *Genome Res* 22(2):220–231. <https://doi.org/10.1101/gr.123497.111>
14. Potter NT, Hurban P, White MN, Whitlock KD, Lofton-Day CE, Tetzner R, Koenig T, Quigley NB, Weiss G (2014) Validation of a real-time PCR-based qualitative assay for the detection of methylated SEPT9 DNA in human plasma. *Clin Chem* 60(9):1183–1191. <https://doi.org/10.1373/clinchem.2013.221044>
15. Barault L, Amatu A, Bleeker FE, Moutinho C, Falcomatà C, Fiano V, Cassingena A, Siravegna G, Milione M, Cassoni P, De Braud F, Rudà R, Soffietti R, Venesio T, Bardelli A, Wesseling P, de Witt Hamer P, Pietrantonio F, Siena S, Esteller M, Sartore-Bianchi A, Di Nicolantonio F (2015) Digital PCR quantification of MGMT methylation refines prediction of clinical benefit from alkylating agents in glioblastoma and metastatic colorectal cancer. *Ann Oncol*. <https://doi.org/10.1093/annonc/mdv272>

16. Mandel P, Metais P (1948) [Not Available] Comptes rendus des seances de la Societe de biologie et de ses filiales 142 (3-4):241–243
17. Ilie M, Hofman V, Long E, Bordone O, Selva E, Washetine K, Marquette CH, Hofman P (2014) Current challenges for detection of circulating tumor cells and cell-free circulating nucleic acids, and their characterization in non-small cell lung carcinoma patients. What is the best blood substrate for personalized medicine? Annals of Translational Medicine 2 (11):107
18. Fleischhacker M, Schmidt B (2007) Circulating nucleic acids (CNAs) and cancer--a survey. Biochim Biophys Acta 1775(1):181–232. <https://doi.org/10.1016/j.bbcan.2006.10.001>
19. Fleischhacker M, Schmidt B, Weickmann S, Fersching DM, Leszinski GS, Siegela B, Stotzer OJ, Nagel D, Holdenrieder S (2011) Methods for isolation of cell-free plasma DNA strongly affect DNA yield. Clinica chimica acta; international journal of clinical chemistry 412(23–24):2085–2088. <https://doi.org/10.1016/j.cca.2011.07.011>
20. Devonshire AS, Whale AS, Gutteridge A, Jones G, Cowen S, Foy CA, Huggett JF (2014) Towards standardisation of cell-free DNA measurement in plasma: controls for extraction efficiency, fragment size bias and quantification. Anal Bioanal Chem 406 (26):6499–6512. <https://doi.org/10.1007/s00216-014-7835-3>
21. Oxnard GR, Paweletz CP, Kuang Y, Mach SL, O'Connell A, Messineo MM, Luke JJ, Butaney M, Kirschmeier P, Jackman DM, Janne PA (2014) Noninvasive detection of response and resistance in EGFR-mutant lung cancer using quantitative next-generation genotyping of cell-free plasma DNA. Clinical cancer research : an official journal of the American Association for Cancer Research 20 (6):1698–1705. <https://doi.org/10.1158/1078-0432.ccr-13-2482>
22. Sanders R, Huggett JF, Bushell CA, Cowen S, Scott DJ, Foy CA (2011) Evaluation of digital PCR for absolute DNA quantification. Anal Chem 83(17):6474–6484. <https://doi.org/10.1021/ac103230c>
23. Whale AS, Cowen S, Foy CA, Huggett JF (2013) Methods for applying accurate digital PCR analysis on low copy DNA samples. PLoS One 8(3):e58177. <https://doi.org/10.1371/journal.pone.0058177>
24. Huggett JF, Foy CA, Benes V, Emslie K, Garrison JA, Haynes R, Hellemans J, Kubista M, Mueller RD, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT, Bustin SA (2013) The digital MIQE guidelines: minimum information for publication of quantitative digital PCR experiments. Clin Chem 59 (6):892–902. <https://doi.org/10.1373/clinchem.2013.206375>
25. Page K, Guttery DS, Zahra N, Primrose L, Elshaw SR, Pringle JH, Blighe K, Marchese SD, Hills A, Woodley L, Stebbing J, Coombes RC, Shaw JA (2013) Influence of plasma processing on recovery and analysis of circulating nucleic acids. PLoS One 8(10):e77963. <https://doi.org/10.1371/journal.pone.0077963>
26. Iba T, Miki T, Hashiguchi N, Tabe Y, Nagaoka I (2014) Combination of antithrombin and recombinant thrombomodulin modulates neutrophil cell-death and decreases circulating DAMPs levels in endotoxemic rats. Thromb Res 134(1):169–173. <https://doi.org/10.1016/j.thromres.2014.04.015>
27. Garcia-Gimenez JL, Sanchis-Gomar F, Lippi G, Mena S, Ivars D, Gomez-Cabrera MC, Vina J, Pallardo FV (2012) Epigenetic biomarkers: a new perspective in laboratory diagnostics. Clinica chimica acta; international journal of clinical chemistry 413 (19–20):1576–1582. <https://doi.org/10.1016/j.cca.2012.05.021>
28. Egert S, Boesch-Saadatmandi C, Wolffram S, Rimbach G, Muller MJ (2010) Serum lipid and blood pressure responses to quercetin vary in overweight patients by apolipoprotein E genotype. J Nutr 140(2):278–284. <https://doi.org/10.3945/jn.109.117655>
29. Kirsch C, Weickmann S, Schmidt B, Fleischhacker M (2008) An improved method for the isolation of free-circulating plasma DNA and cell-free DNA from other body fluids. Ann N Y Acad Sci 1137:135–139. <https://doi.org/10.1196/annals.1448.035>
30. Mauger F, Dulary C, Daviaud C, Deleuze JF, Tost J (2015) Comprehensive evaluation of methods to isolate, quantify, and characterize circulating cell-free DNA from small volumes of plasma. Anal Bioanal Chem 407 (22):6873–6878. <https://doi.org/10.1007/s00216-015-8846-4>
31. Keim SA, Kulkarni MM, McNamara K, Geraghty SR, Billock RM, Ronau R, Hogan JS, Kwiek JJ (2015) Cow's milk contamination of human milk purchased via the internet. Pediatrics. <https://doi.org/10.1542/peds.2014-3554>
32. Tritten L, O'Neill M, Nutting C, Wanji S, Njouendou A, Fombad F, Kengne-Ouaffo J, Mackenzie C, Geary T (2014) Loa Loa and

- Onchocerca ochengi miRNAs detected in host circulation. Mol Biochem Parasitol 198 (1):14–17. <https://doi.org/10.1016/j.molbiopara.2014.11.001>
33. How-Kit A, Tost J (2015) Pyrosequencing ® -based identification of low-frequency mutations enriched through enhanced- ice -COLD-PCR. In: Lehmann U, Tost J (eds) Pyrosequencing: Methods and Protocols. Springer, Berlin, Germany, pp 83–101
34. Dawson S-J, Tsui DWY, Murtaza M, Biggs H, Rueda OM, Chin S-F, Dunning MJ, Gale D, Forshaw T, Mahler-Araujo B, Rajan S, Humphray S, Becq J, Halsall D, Wallis M, Bentley D, Caldas C, Rosenfeld N (2013) Analysis of circulating tumor DNA to monitor metastatic breast cancer. New England Journal of Medicine 368(13):1199–1209. <https://doi.org/10.1056/NEJMoa1213261>
35. Holmberg RC, Gindlesperger A, Stokes T, Lopez D, Hyman L, Freed M, Belgrader P, Harvey J, Li Z (2013) Akonni TruTip((R)) and Qiagen((R)) methods for extraction of fetal circulating DNA-evaluation by real-time and digital PCR. PLoS One 8(8):e73068. <https://doi.org/10.1371/journal.pone.0073068>
36. Breitbach S, Tug S, Helmig S, Zahn D, Kubiak T, Michal M, Gori T, Ehlert T, Beiter T, Simon P (2014) Direct quantification of cell-free, circulating DNA from unpurified plasma. PLoS One 9(3):e87838. <https://doi.org/10.1371/journal.pone.0087838>
37. Fong SL, Zhang JT, Lim CK, Eu KW, Liu Y (2009) Comparison of 7 methods for extracting cell-free DNA from serum samples of colorectal cancer patients. Clin Chem 55 (3):587–589. <https://doi.org/10.1373/clinchem.2008.110122>
38. Keshavarz Z, Moezzi L, Ranjbaran R, Aboualizadeh F, Behzad-Behbahani A, Abdullahi M, Sharifzadeh S (2015) Evaluation of a modified DNA extraction method for isolation of cell-free fetal DNA from maternal serum. Avicenna journal of medical biotechnology 7(2):85–88
39. Jorgez CJ, Dang DD, Simpson JL, Lewis DE, Bischoff FZ (2006) Quantity versus quality: optimal methods for cell-free DNA isolation from plasma of pregnant women. Genetics in Medicine: Official Journal of the American College of Medical Genetics 8 (10):615–619. doi: 10.1097/gim.0000241904.32039.6f
40. Dube S, Qin J, Ramakrishnan R (2008) Mathematical analysis of copy number variation in a DNA sample using digital PCR on a nanofluidic device. PLoS One 3(8):e2876. <https://doi.org/10.1371/journal.pone.0002876>
41. Milbury CA, Zhong Q, Lin J, Williams M, Olson J, Link DR, Hutchison B (2014) Determining lower limits of detection of digital PCR assays for cancer-related gene mutations. Biomolecular Detection and Quantification 1 (1):8–22. <https://doi.org/10.1016/j.bdq.2014.08.001>

## **Part II**

### **Absolute Quantification**



# Chapter 5

## Multiplex Droplet Digital PCR Protocols for Quantification of GM Maize Events

**David Dobnik, Bjørn Spilsberg, Alexandra Bogožalec Košir, Dejan Štebih, Dany Morisset, Arne Holst-Jensen, and Jana Žel**

### Abstract

The standard-curve based simplex quantitative polymerase chain reaction (qPCR) has been the gold standard for DNA target quantification for more than a decade. The large and growing number of individual analyses needed to test for genetically modified organisms (GMOs) is reducing the cost-effectiveness of qPCR. Droplet digital PCR (ddPCR) enables absolute quantification without standard curves, avoids the amplification efficiency bias observed with qPCR, allows more accurate estimations at low target copy numbers and, in combination with multiplexing, significantly improves cost efficiency. Here we describe two protocols for multiplex quantification of GM maize events: (1) nondiscriminating, with multiplex quantification of targets as a group (12 GM maize lines) and (2) discriminating, with multiplex quantification of individual targets (events). The first enables the quantification of twelve European Union authorized GM maize events as a group with only two assays, but does not permit determination of the individual events present. The second protocol enables the quantification of four individual targets (three GM events and one endogene) in a single reaction. Both protocols can be modified for quantification of any other DNA target.

**Key words** Droplet digital PCR, ddPCR, GM maize, GMO, Multiplex, Quantification

---

### 1 Introduction

Standard-curve based quantitative PCR (qPCR) is regarded as the gold standard technology for analysis of genetically modified organisms (GMOs) due to its sensitivity and robustness. In the last decade the number of GMOs found on a given market, and for example authorized in the EU, has steadily increased. Consequently, testing for GMOs must target a large number of sequence motifs to ensure detection and quantification of all GMOs in consideration can be detected and quantified. Analysis employing separate event-specific qPCRs is therefore no longer cost efficient. The use of multiplex PCR can mitigate the problem of cost effectiveness. However, a multiplex PCR is a complex experimental system

where potential interference between oligonucleotides and amplification products can produce false positive/negative results. Several multiplex qPCRs for quantification of GMOs have been developed and reported, but only two (duplexes) are validated in an interlaboratory Scheme [1, 2].

The concept of digital PCR (dPCR) was presented already in 1992 [3]. In dPCR the reaction mixture is distributed into a large number of partitions and at end point of the reaction these partitions are qualitatively scored as positive or negative, based on the fluorescence signal. Initial DNA concentration is then calculated using the Poisson distribution [4]. In contrast to qPCR, partial inhibition of the PCR does not affect dPCR-based quantitation, as it is only necessary to distinguish partitions positive for the target of interest from those that are negative in order to calculate the sample concentration of the target. Two available dPCR technologies, microfluidic/chip based dPCR and emulsion (droplet) based dPCR [5], have been tested for applicability for GMO detection [6–10].

Droplet digital PCR (ddPCR) does not need standard curves for absolute quantification, thus avoiding the amplification efficiency bias observed with qPCR [7, 8]. Targets even at low copy numbers can be accurately determined [11] and cost-efficiency can be significantly improved in combination with multiplexing. The protocols described below are focused on the ddPCR implementation, as this technology has already been shown to be suitable for routine GMO diagnostics using duplex reactions [9] and uses high number of partitions per well, thus enabling a wide linear range of the response.

Two different protocols for multiplex quantification of GM maize events are described below: (1) nondiscriminative multiplex quantification of GM maize targets as a group, i.e., per ingredient or plant species (multiplex quantification per ingredient, MQI) and (2) discriminative multiplex quantification of individual GM events (multiplex event quantification, MEQ).

In the MQI protocol, all 12 GM maize events authorized in the EU by April 1, 2015 are quantified simultaneously as a group using two assays per sample, without possibility of distinguishing individual targets present in the sample. This protocol is in line with the EU labeling regulation [12], which states that labeling of a product is not required, when the combined concentration of all EU authorized GM events per individual plant species is below 0.9% and the presence is unintended and technically unavoidable. The two assays, 4- and 10-plex, also include, besides the 3 and 9 GM events (see Tables 1 and 2), respectively, a maize endogene (*hmgA*) as a target. The inclusion of *hmgA* is necessary for calculation of GM maize content relative to all maize content in the sample. The full performance of these multiplex assays has

**Table 1**  
**Primers and probes used in the MQI 4-plex assay**

Target and method reference	Name	DNA sequence of the oligonucleotide (5'-sequence-3')	Final concentration in the PCR [nmol/l]	Volume of 40 µM solution needed for 72 µl of PP mix [µl]
<i>hmgA</i> , QT-TAX-ZM-002	Fw-hmgA	TGGACTAGAAATCTCGTGCTGA	900	5.40
	R-hmgA	GCTACATAGGGAGCCTTGTCT	900	5.40
	P-hmgA	HEX-CAATCCACACAAACGCGGTAA-BHQ-1	180	1.08
Bt11 <sup>a</sup> , Brodmann et al. [13]	Fw-Bt11	GCGGCTTATCTGTCTCAGGG	600	3.60
	R-Bt11	CAACTGGTCTCTCTCCGGGA	600	3.60
	P-Bt11	6-FAM-CGTGTTCCCTCGGATCTCGACATGT-BHQ-1	180	1.08
NK603, QT-EVE-ZM-008	Fw-NK603	ATGAATGACCTCGAGTAAGCTTGTAA	900	5.40
	R-NK603	AAGAGATAACAGGATCCACTCAAACACT	900	5.40
	P-NK603	6-FAM-TGGTACACACGCGACACACTCCACTC-BHQ-1	180	1.08
DAS59122, QT-EVE-ZM-012	Fw-DAS59122	GGGATAAGCAAGTAAAAGGGCTC	600	3.60
	R-DAS59122	CCTTAATTCTCCGCTCATGATCAG	600	3.60
	P-DAS59122	6-FAM-TTTAAAACCTGAAGGGAAACGACAA-BHQ-1	200	1.20
Nuclease-free water				31.56

Method reference: reference for the method for quantification of event/gene. Methods starting with QT were validated by the EURL GMFF (<http://gmo-crl.jrc.ec.europa.eu/gmomethtods/>)

*PP mix*: a mix of forward primers, reverse primers, and probes, 6-FAM 6-carboxyfluorescein, BHQ1 black hole quencher 1, HEX hexachloro-6-carboxyfluorescein

<sup>a</sup>Construct specific method

**Table 2**  
**Primers and probes used in the MQI 10-plex assay**

<b>Target and method reference</b>	<b>Name</b>	<b>DNA sequence of the oligonucleotide (5'-sequence-3')</b>	<b>Volume of 40 µM solution needed for 120 µl of PP mix [µl]</b>		
			<b>Final concentration in the PCR [nmol/l]</b>	<b>Final concentration in the sequencing reaction [nmol/l]</b>	<b>Volume of 40 µM solution needed for 120 µl of PP mix [µl]</b>
hmGA, QT-TAX-ZM-002	Fw-hmGA R-hmGA P-hmGA	TGGGACTAGAAATCTCGTGCAGA GCTACATAGGGAGCCCTTGCTCT HEX-CAATCCACACAAACGCACGCCGTABHQ-1	900 900 180	9.0 9.0 1.8	9.0 9.0 1.8
DAS1507, QT-EVE-ZM-010	Fw-DAS1507 R-DAS1507 P-DAS1507	TAGTCTTCGGGCCAGAACATGG CTTTGCCAAGATCAAGCG 6-FAM-TAACTCAAGGCCCTCACTCCG-BHQ-1	450 450 90	4.5 4.5 0.9	4.5 4.5 0.9
GA21, QT-EVE-ZM-014	Fw-GA21 R-GA21 P-GA21	C GTTATGCTTATTGCAACCTTAGAACAA GGGATCCTCTCGCGTT 6-FAM-TTTCTCAACAGCAGGGTGGT CCGGGT-BHQ-1	450 450 90	4.5 4.5 0.9	4.5 4.5 0.9
MIR604, QT-EVE-ZM-013	Fw-MIR604 R-MIR604 P-MIR604	GCGCACGCCAATTCAACAG GGTCATAACGTGACTCCCTTAATTCT 6-FAM-AGGGGGAAACGACAATCTGAT CATG-BHQ-1	150 150 90	1.5 1.5 0.9	1.5 1.5 0.9
MIR162, QT-EVE-ZM-022	Fw-MIR162 R-MIR162 P-MIR162	GCGGGGTGTCATCTATGTTACTAG TGCCTTATCTGTGTCCTCAGA 6-FAM-TCTAGACAAATTCACTACATAA AAACGTCGCCA-BHQ-1	150 150 90	1.5 1.5 0.9	1.5 1.5 0.9
MON810, QT-EVE-ZM-020	Fw-MON810 R-MON810 P-MON810	TCGAAGGAGCGAACGACTCTAACGT GCCACCTTCCCTTCCACTATCTT 6-FAM-AAACATCCCTTGCCATTGCCACGC-BHQ-1	150 150 90	1.5 1.5 0.9	1.5 1.5 0.9
MON863, QT-EVE-ZM-009	Fw-MON863 R-MON863 P-MON863	TGTACGGCTAAATGCTGAACT GTAGGATCGGAAAGCTGGTAC 6-FAM-TGAACACCACCGAACAAAG TAGGGTCA-BHQ-1	450 450 90	4.5 4.5 0.9	4.5 4.5 0.9

MON88017, QT-EVE-ZM-016	Fw-MON88017 R-MON88017 P-MON88017	GAGCAGGACCTGAGAAGCT GCCGGAGTTGACCATCCA 6-FAM-TCCCGCCTTCAGTTAACAGAGTC GGGT-BHQ-1	150 150 90	1.5 1.5 0.9
MON89034, QT-EVE-ZM-018	Fw-MON89034 R-MON89034 P-MON89034	TTCTCCATAATTGACCATCATACTCATT CGGTATCTATAATAACCGTGGTTTAA 6-FAM-ATCCCCGGAAATTATGTT-MGBNFQ	450 450 180	4.5 4.5 1.8
T25, QT-EVE-ZM-011	Fw-T25 R-T25 P-T25	ACAAGCGTGTCTGCTGCTCCAC GACATGATACTCCCTCCACCG 6-FAM-TCAATTGAGTCGGTCCGCCATTGTCG-BHQ-1	450 450 90	4.5 4.5 0.9
		Nuclease-free water	34.2	

Methods were validated by the EURL-GMFF (<http://gmo-ctrl.jrc.ec.europa.eu/gmometodshome/>). Method reference: reference for the method for quantification of event/gene *PP mix* a mix of forward primers, reverse primers, and probes, 6-FAM 6-carboxyfluorescein, *BHQ1* black hole quencher 1, *HEX* hexachloro-6-carboxyfluorescein, *MGB-NFQ* minor groove binding non-fluorescent quencher

been reported in Dobnik et al. (2015) [14]. Additionally a MQI system for quantification of soybean GM lines, has been assessed by [15].

In the MEQ protocol, in contrast, three GM maize events (*see* Table 3) and a maize endogene (*hmga*) are quantified in one single reaction as distinguishable DNA targets (appearing in distinguishable droplet clusters). The performance of this multiplex was recently reported [16]. Note that this multiplex assay targets different GMO events than the MQI 4-plex.

A flowchart of experimental steps for both protocols is shown in Fig. 1. Both of these protocols can be modified for quantification of any other DNA target by designing new assays and after adjustments of primer and probe concentrations as described in Subheading 3.

For the purpose of these protocols DNA isolated from different kinds of samples (food, feed, seed, and plants) can be used. The extraction technique and sample properties have a crucial influence on the results of GMO detection and quantification [17]. Efficient and reliable extraction of nucleic acids is a prerequisite for obtaining accurate results from molecular analyses such as amplification of specific DNA or RNA sequences by polymerase chain reaction (PCR). The extraction method should yield sufficient DNA of adequate structural integrity and purity, independently of the matrix to which it is applied [18]. Acceptance criteria applicable to DNA extraction methods were recently defined in the new European Network of GMO Laboratories (ENGL) document on Definition of minimum performance requirements for analytical methods of GMO testing [19]. Briefly, concentration, yield, structural integrity, and purity are to be considered. Although these requirements were prepared for real-time PCR, they can also all be considered as minimum requirements for ddPCR. Regardless of the method used for DNA extraction, it is essential to include quality controls to exclude the possibility of sample contamination during extraction. For this purpose the use of extraction blank control and environment control is necessary. These controls are later on used in PCR, together with positive and no template control.

---

## 2 Materials

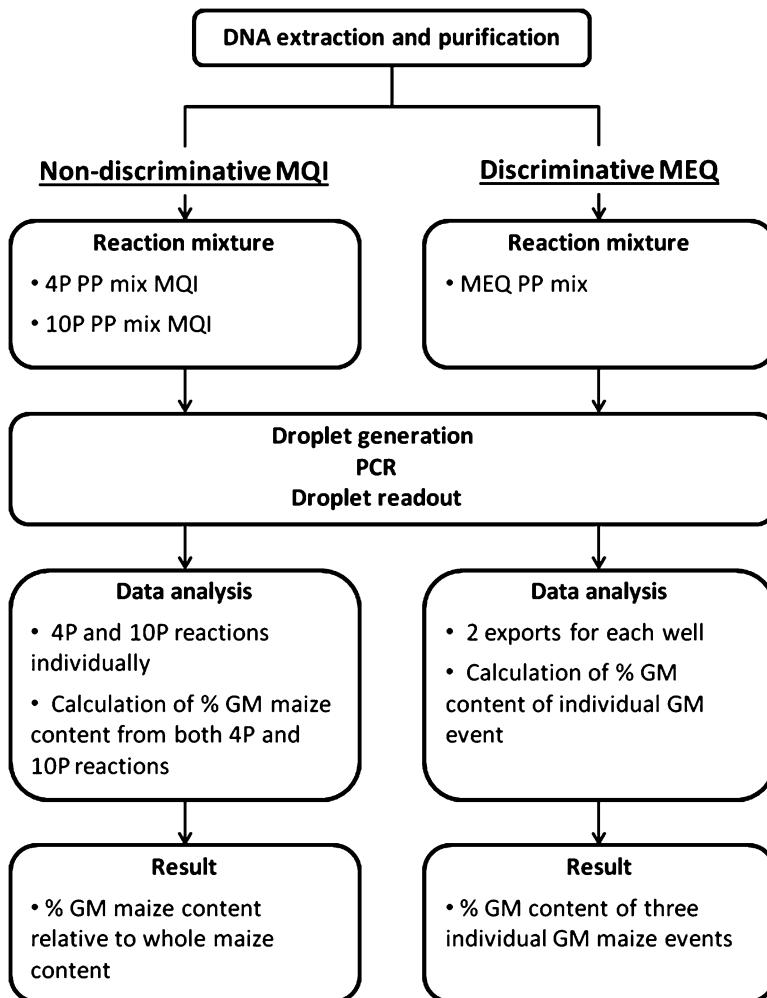
### 2.1 Droplet Generation

1. Droplet Generation Oil for Probes (Bio-Rad, Pleasanton, CA).
2. Droplet Generator DG8™ Cartridges and Gaskets (Bio-Rad, Pleasanton, CA).
3. 2× ddPCR™ Supermix for Probes (No dUTP) (Bio-Rad, Pleasanton, CA).
4. Nuclease-free water (Sigma-Aldrich®, MO) or equivalent.

**Table 3**  
Primers and probes used in MEQ assay

Target and method reference	Name	DNA sequence of the oligonucleotide (5'-sequence-3')	Final concentration in the PCR [nmol/l]	Volume of 40 µM solution needed for 120 µl of PP mix [µl]
<i>bmgA</i> , QT-TAX-ZM-002	Fw-hmgA R-hmgA P-hmgA	TTCGACTAGAAATCTCGTGTGA GCTACATAGGGAGCCCTGTCTT 6-FAM-CAATCCACACAAACGCCACGCGTA-BHQ-1	300 300 100	3 3 1
MON810, QT-EVE-ZM-020	Fw-MON810 R-MON810 P-MON810	TCGAAAGGACGAAAGGACTCTAACGT GCCACCTTCCTTCCATTCACTATCTT HEX-AACATCCCTTGCCTTGCCAGC-BHQ-1	900 900 300	9 9 3
MON863, QT-EVE-ZM-009	Fw-MON863 R-MON863 P-MON863	TCTTACGGCTAAATGCTGAACCT GTAGGATCAGGAAAGCTTGGTAC 6-FAM-TGAACACCCATCCGAACAAGTAGGGTCA-BHQ-1	900 900 300	9 9 3
DP98140, QT-EVE-ZM-021	Fw-DP98140 R-DP98140 P-DP98140	GTGTGTATGTCTCTTGTCTGGTCTT GAITGTCTGTTCCCCGCTTC HEX-CCTCATCGATCCCCCTCTTGTATAAATCTT-BHQ-1	300 300 200	3 3 2
		Nuclease-free water		63

Method reference: reference for the method for quantification of event/gene. Methods were validated by the EUR-L-GMFF (<http://gmo-crjrc.europa.eu/gmometodo/>)  
PP mix a mix of forward primers, reverse primers, and probes, 6-FAM 6-carboxyfluorescein, BHQ1 black hole quencher 1, HEX hexachloro-6-carboxyfluorescein



**Fig. 1** Flowchart of experimental steps for both protocols. For both protocols the steps of droplet generation, PCR, and droplet readout are the same, but the reaction mixture and analysis steps with final reporting of results are different

5. Primers and probes for each of the two reactions of MQI protocol are described in Table 1 (4-plex (4P) reaction) and Table 2 (10-plex (10P) reaction).
6. Primers and probes for MEQ 4-plex protocol are described in Table 3.
7. DG8™ Cartridge Holder (Bio-Rad, Pleasanton, CA).
8. QX100™/200™ Droplet Generator (Bio-Rad, Pleasanton, CA).

## 2.2 PCR

1. Rainin presterilized BioClean LTS tips with filter (*see Note 1*).
2. Eppendorf™ 96-Well twin.tec™ PCR Plates, Semi-skirted (Eppendorf, Germany) (*see Note 2*).

3. Pierceable Foil Heat Seal (Bio-Rad, Pleasanton, CA) or equivalent.
4. Heat Sealer (Eppendorf, Germany) or equivalent.
5. T100™ Thermal Cycler (Bio-Rad, Pleasanton, CA) or equivalent (*see Note 3*).

### **2.3 Droplet Readout**

1. Droplet Reader Oil (Bio-Rad, Pleasanton, CA).
2. QX100™/200™ Droplet Reader (Bio-Rad, Pleasanton, CA).

## **3 Methods**

### **3.1 Sample Acquisition**

Sampling methodology is not described in this protocol, as it is usually not performed by the laboratory personnel. However, to get an insight into this process and the procedure of subsampling, we propose reading of the following document: Guidelines for sample preparation procedures in GMO analysis [20].

### **3.2 Sample Preparation**

Most of the samples, including seeds and feed, should be ground with a mill enabling good particle size reduction (e.g., Retsch ZM200 rotor mill). Samples with high fat content are to be cooled with liquid nitrogen before grinding with a Retsch GM200 knife mill. Soft samples, such as sausages or tofu, are to be homogenized with a Bioreba HOMEX 6 homogenizer. For leaf samples, first a small piece is cut from each individual leaf and all the pieces are then combined and homogenized as one sample with a FASTprep instrument (MP Biomedicals) using 15 ml tubes, a ceramic ball, and quartz sand.

Different procedures can be applied for DNA extraction from the homogenized sample, but for the purpose of this protocol the following extraction methods can be used, based on the sample type [21]:

- NucleoSpin® Food (Macherey-Nagel) kit, which should be used as described by the manufacturer.
- The cetyltrimethylammonium bromide (CTAB) method with RNase-A and proteinase-K solutions for removing RNA and proteins from the sample should be performed as described in ISO 21571, Annex A.3—Preparation of PCR-quality DNA using the CTAB-based DNA extraction methods [22]. Because the CTAB method uses chloroform, the relevant steps should be performed in a fume hood.
- DNeasy Plant Mini Kit (Qiagen) should be used as described by the manufacturer to extract DNA from plant leaves.

DNA prepared by following one of these protocols is usually of such quality that it is suitable for further analysis with ddPCR.

Regardless of the protocol used for DNA extraction, two controls must be used during the process: (1) the extraction blank control, where water is used instead of a sample and (2) environment control, where a tube with a volume of water equal to elution volume of samples is left opened during the DNA extraction [23].

### **3.3 Optimization of Multiplex Assays**

The primers and probes combined into the described multiplex assays were originally designed for real-time PCR with one primer pair and probe per event (event specific primer–probe combination, ESPPC), but the combinations were optimized to achieve reliable detection and quantification of the targets also in ddPCR.

The first step when combining different ESPPCs into multiplex is the *in silico* analysis of interactions between the primers and probes. For the purpose of the described protocols, this was performed with the software Autodimer [24]. All primers and probes were aligned with a local alignment algorithm with a score of +1 for match and a penalty of -1 for mismatch. The interactions with a score  $\geq 9$  were considered as the ones potentially causing significant interference. Thus the ESPPCs with such scores were separated. Nevertheless, as not all interactions can be predicted by *in silico* analysis, the performance of multiplex assays was first experimentally evaluated on DNA samples containing single targets (e.g., in MQI two of the assays were removed from one group, based on the experimental results).

As the described multiplex assays were developed for GMO quantification, their performance (sensitivity, specificity, accuracy, trueness, ...) have to be compliant with Minimum Performance Requirements for Analytical Methods of GMO Testing [19]. For that reason a thorough in-house validation of the new assay should be performed (following reported guidelines [19]), including testing the precision/repeatability on serial dilutions to determine limits of quantification and detection [19].

Additionally, PCR cycling protocol can be optimized (e.g., by using gradient PCR) to get the best possible separation of droplet clusters and minimize the “rain” of droplets between clusters. The term “rain” refers to the droplets with fluorescence amplitude ranging from explicitly negative and explicitly positive. The definite origin of rain is not clear, but it was speculated that it occurs due to delayed onset of PCR, partial inhibition of individual droplets or damaged positive or negative droplets, what affects the fluorescence signal [25, 26]. However, for the described protocols the cycling parameters were not optimized and were identical to those from the simplex real-time PCR methods. Specific considerations for developments of new multiplexes falling in line with the two different protocols described in this chapter are given below.

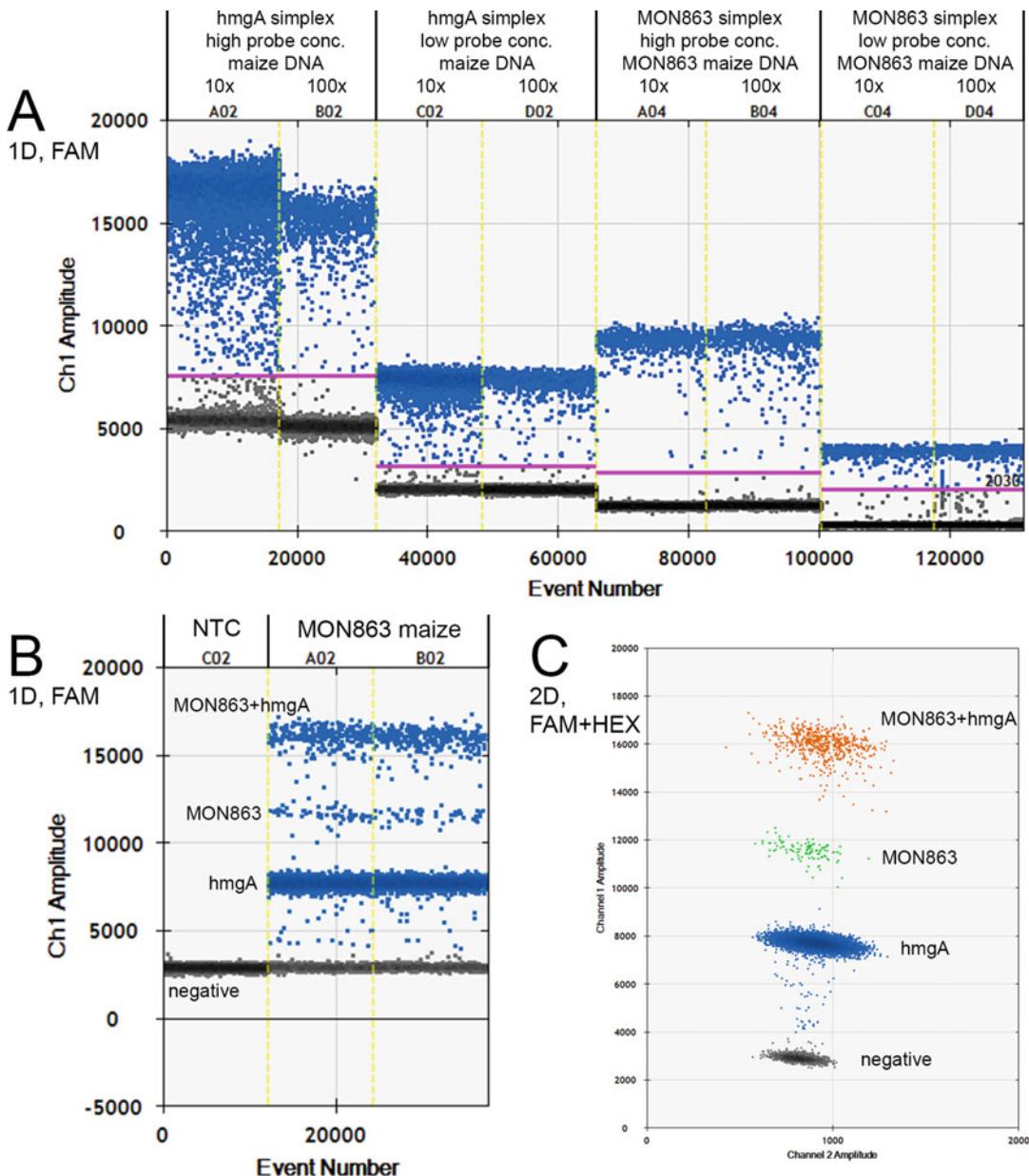
### 3.3.1 Nondiscriminative Multiplex

This kind of multiplexing which quantifies a group of targets without the possibility to discriminately identify them requires that the assay is tested for interactions between primers and probes and their potential influence on the performance (especially false negative or false positive signals). Additionally, the concentrations of primers and probes must be adjusted for some individual targets. The performance of multiplex assay must be evaluated against simplex reaction on the same DNA sample containing only one target. Whenever the absolute difference between the results is  $\geq 25\%$ , the concentrations of primers and probes must be adjusted, so that results would fit in this limit.

Increasing the number of targets in the multiplex within one fluorescent reporter is a major limitation. With the increased number of targets, the concentration of probes is also increasing, leading to the increase of the negative background fluorescence. When the fluorescence amplitude of negative droplets is too high it reaches the amplitude of positive droplets and the clusters can no longer be separated. Therefore, to allow appropriate discrimination between negative and positive droplets the number of individual assays that can be combined into multiplex is limited. The described process of optimization can be done for both fluorescence channels and finally only one group of targets per channel can be quantified simultaneously.

### 3.3.2 Discriminative Multiplex

Bio-Rad's ddPCR devices (QX100 or QX200) enable detection in two fluorescence channels, what makes them ready for at least duplex quantification (one target in each channel). Because the fluorescence amplitude measured in negative and positive droplets after amplification depends on the probe concentration, the multiplexing level of ddPCR can be higher. For the MEQ protocol described below, we have exploited this characteristic and prepared a quadruplex reaction, where two targets are quantified within one fluorescence channel. It is at most important, in this kind of multiplexing, to optimize primer and probe concentrations to achieve a clear separation of droplet clusters. With that in mind, assay developers should test different concentration combinations (e.g., low probe concentration for one target and high probe concentration for second target in FAM channel). In the case of low concentrations, the cluster of positive droplets should still be clearly above the cluster of negative droplets. Examples of different fluorescence amplitudes when using high or low probe concentrations are presented in Fig. 2. Note that when optimizing, low concentration conditions might give amplitude of positive droplets at the same level as is the cluster of negative droplets in high concentration conditions (Fig. 2). However, when combining these two together, the negative and two positive clusters would still be separated, as the fluorescence amplitude will be summed up.



**Fig. 2** Example of the effect of different probe concentrations on the fluorescence amplitude when developing the MEQ assays. Simplex reactions for *hmgA* and MON863 (both in FAM channel) were performed on two dilutions of DNA using high and low probe concentrations and readout is visualized in 1D (**a**). Both simplex assays were combined into duplex (both in FAM channel) where *hmgA* had low probe concentration and MON863 had high probe concentration (**b**). Clear separation of negative and different positive clusters can be observed in 1D view. The same cluster separation can be observed also in 2D view (**c**)

### **3.4 Preparation of Reaction Mixtures**

Primer and probe stock solutions should be kept at a final concentration of 100 µM, stored at –20 °C. To reduce freezing–thawing cycles, smaller volumes (for one time use) of 40 µM aliquots of working solution should be prepared (diluted in nuclease-free water). Such aliquots can be stored for 3 months at –20 °C.

1. To prepare a primers–probe (PP) mix for MQI, for each of the assays (4P and 10P) mix the specified volumes of 40 µM working solutions of individual primers and probes according to the Tables 1 and 2. For MEQ, prepare a PP mix by mixing the specified volumes of 40 µM working solutions of individual primers and probes according to the Table 3 (*see Note 4*).
2. To prepare the reaction mixture for MQI for each of the assays (4P and 10P), mix ddPCR™ Supermix for Probes and 4P PP mix or 10P PP mix, respectively. For MEQ protocol mix ddPCR™ Supermix for Probes with MEQ PP mix. For one reaction prepare 17.6 µl of Mastermix composed of 11 µl ddPCR™ Supermix for Probes and 6.6 µl PP mix (*see Note 5*).
3. Distribute 17.6 µl of Mastermix into the wells of nuclease-free 8-well strips or nuclease-free 96-well plates (*see Note 6*).
4. To get final reaction mixture, add 4.4 µl of sample DNA to each well with Mastermix (*see Note 7*). It is essential to include controls in the experiment, and therefore, instead of sample DNA, a prepared mixture of DNA containing all the targets (as positive control), nuclease-free water (as no template control, NTC), a sample of extraction blank control, and environment control are added to individual wells with Mastermix (*see Note 8*).
5. Briefly mix by pipetting up and down (at least five times) or vortexing and centrifuge to collect liquid at the bottom of wells/tubes prior to loading the cartridge (*see Note 9*).

### **3.5 Droplet Generation**

1. Insert 8-well cartridge for droplet generation into the cartridge holder.
2. Pipette 20 µl of reaction mixture in each of the middle wells (*see Note 10*).
3. Pipette 70 µl of Droplet Generation Oil in each of the lower wells.
4. Place the gasket over the cartridge holder.
5. Insert the cartridge holder with filled cartridge into the QX100/200™ droplet generator.
6. After the droplets are generated, carefully transfer 40 µl of droplet suspension (water-in-oil emulsion) from a cartridge to a 96-well plate (*see Note 11*).
7. Cover the plate with Pierceable Heat Seal Foil, seal the plate by using Heat Sealer and transfer it to a thermal cycler (*see Note 12*).

**3.6 PCR Reaction**

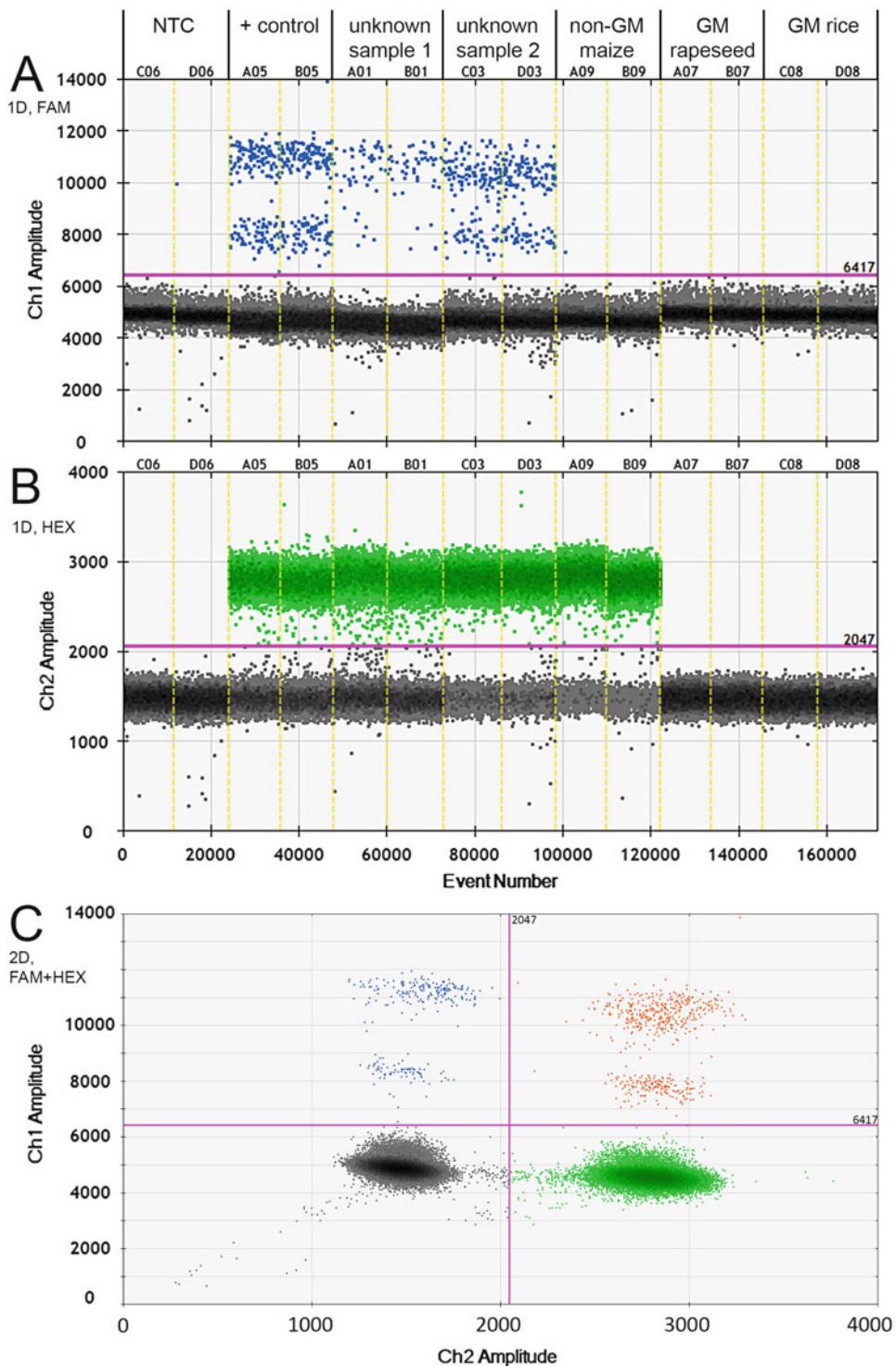
1. Insert the sealed 96-well plate into a conventional thermal cycler (*see Note 3*).
2. Set the cycling program to:
  - 2 min at 50 °C.
  - 10 min at 95 °C.
  - 40 cycles of a two-step thermal profile comprising of 15 s at 95 °C and 60 s at 60 °C (at ramp rate 2.5 °C/s).
  - 10 min at 98 °C (*see Note 13*).
  - Infinite hold at 4 °C.

**3.7 Droplet Reading**

1. When thermal cycling is completed transfer the sealed 96-well plate into the QX100/200™ Droplet Reader.
2. Start the QuantaSoft software (Bio-Rad).
3. First flush the system by performing the following operation: in Quantasoft software under “Setup” and “Instrument Routines” select “Flush System”.
4. Set up new experiment by using “Setup,” “Template,” and “New,” and then use well-editor to complete the required information for each well (*see Note 14*).
5. Save the created experiment.
6. Select “Run” to start the analysis.
7. When prompted, choose appropriate Color compensation option (FAM/HEX).
8. After the analysis has finished save experiment data and remove the plate from the machine.

**3.8 Data Analysis for Nondiscriminating MQI Testing**

1. To access the droplets analysis results in QuantaSoft, use “Setup,” “Plate Load,” select and open the saved experiment and select “Analyze.”
2. To distinguish between positive and negative droplets, set the threshold for each well, for each detector separately. Use NTC and positive control results as guidance for threshold setting (*see example in Figs. 3 and 4*) (*see Note 15*). An example of droplet readout on a dilution series of DNA mix is presented in Fig. 5.
3. Export the results as .csv file, which can be later on opened and further analyzed in Excel (Microsoft Office) or equivalent software.
4. Wells with accepted number of droplets <8000 should not be considered for further analysis (*see Note 16*)
5. The reaction in each well is considered as positive, if three or more droplets show the fluorescent signal above the threshold value (*see Note 17*).



**Fig. 3** Example of a threshold setting for 4P MQI with NTC, positive control, and different samples. Ch1 (**a**) represents the fluorescence signal originating from FAM, and Ch2 (**b**) represents the fluorescence signal originating from HEX. **a** and **b** represent the 1D view of analyzed droplets for transgene and *hmgA* targets, respectively. Two replicates for each of the following samples are shown: NTC (C06-D06), positive (+) control

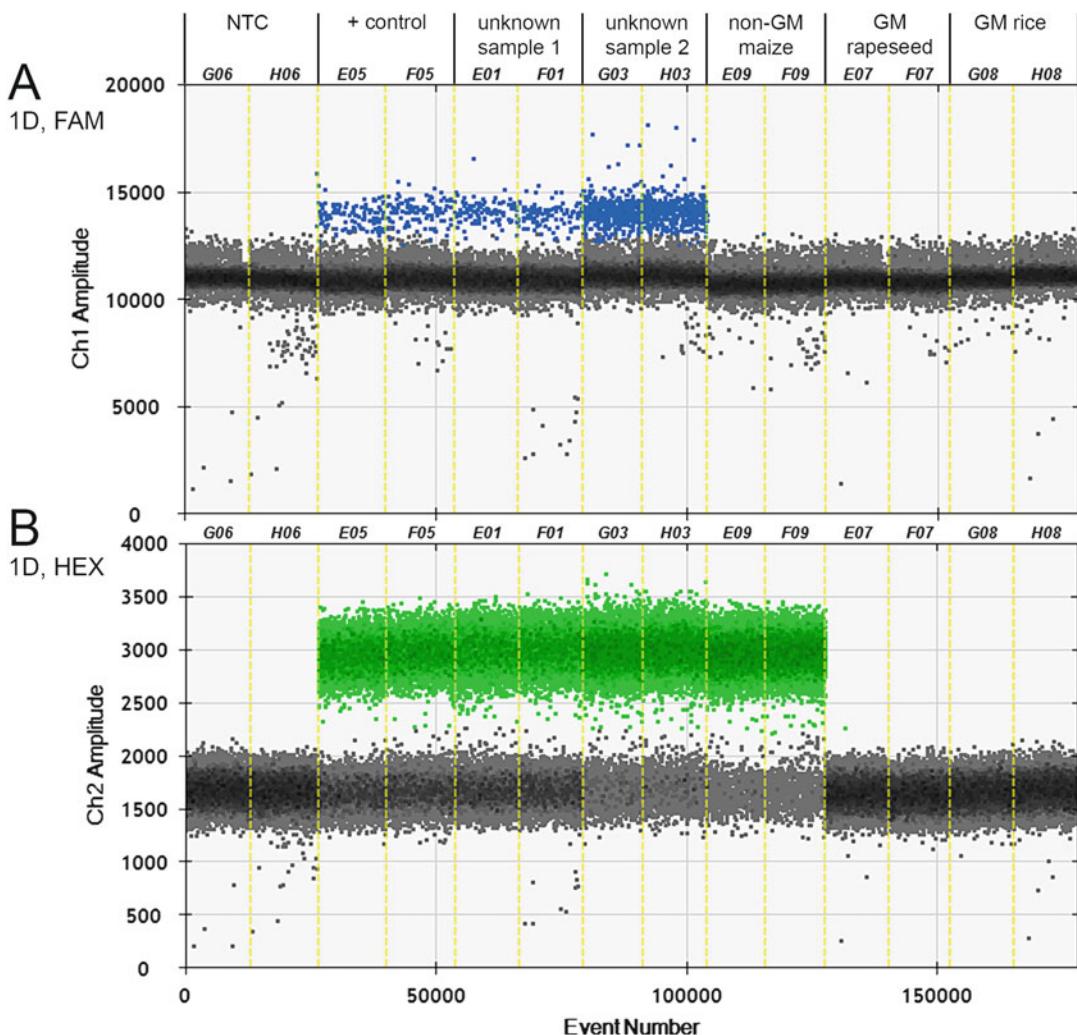
6. The reaction in each well is considered as negative, if <3 droplets show the fluorescent signal above the threshold value.
7. The target sequence is considered as detected, if for the 4P or 10P assay the reaction is positive, the NTC is negative, and the positive amplification control (positive DNA target control) is positive.
8. The target sequence is considered as not detected, if for the 4P and 10P assay the reaction is negative, the NTC is negative, and the positive amplification control (positive DNA target control) is positive.
9. The analysis must be repeated, if the positive amplification control reaction is negative or the negative controls are positive.
10. For calculation of GM maize content use the concentration values (copies/ $\mu$ l of reaction) converted to a concentration of undiluted DNA (the dilution of the sample is used for this calculation).
11. Percentage of GM maize in each sample (relative to all maize content) is calculated using the following equation, where average is calculated from concentration values of undiluted DNA (see Table 4 for example of calculation):

$$\text{GM maize content}[\%] = \frac{\text{average transgene concentration of 4P replicates} + \text{average transgene concentration of 10P replicates}}{\text{average concentration of } hmgA \text{ for replicates of both assays (4P and 10P)}} \times 100$$

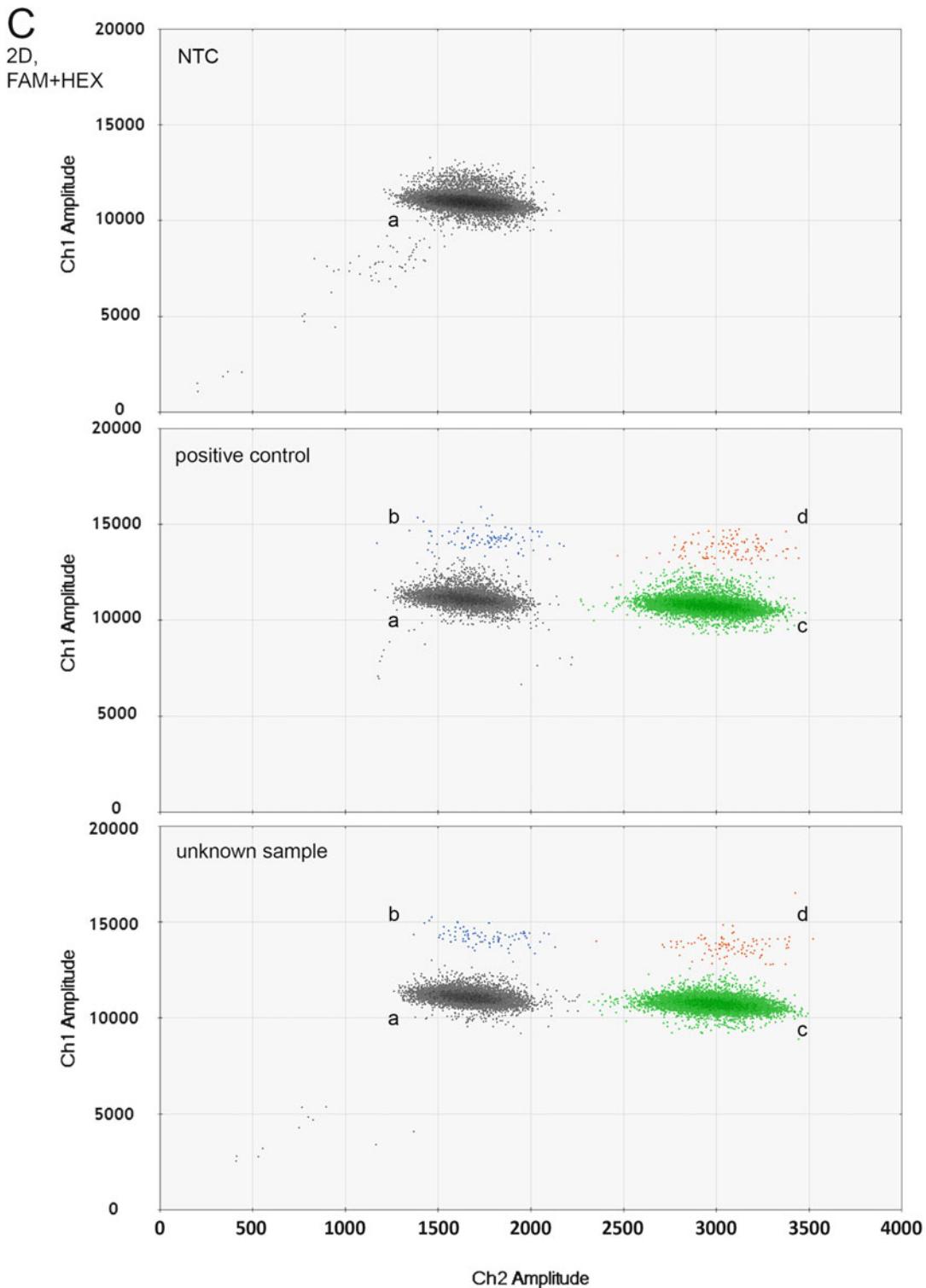
### **3.9 Data Analysis for Discriminating MEQ Testing**

1. To access the droplets analysis results in QuantaSoft, use “Setup,” “Plate Load,” select and open the saved experiment and select “Analyze.”
2. QuantaSoft software does not enable setting of more than one threshold for a given fluorophore, but there are several clusters that need to be separated (see Fig. 6). Thus two separate exports must be performed for each reaction of MEQ protocol (see Note 18 and Fig. 7). Use positive control results as guidance for setting the thresholds. The appropriate clusters for

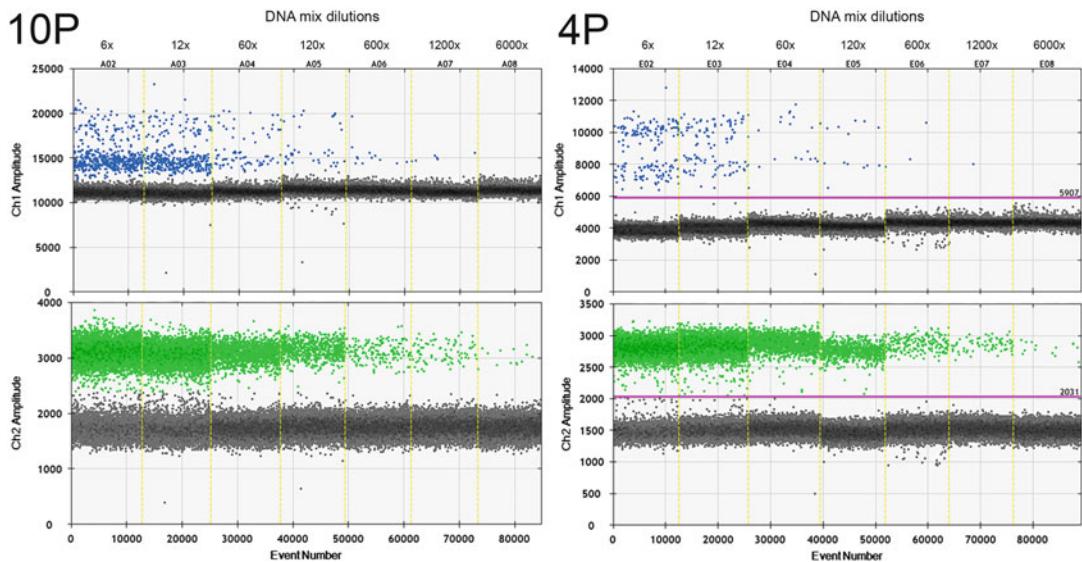
**Fig. 3** (continued) for transgenes (A05-B05), two unknown samples containing GM maize at different concentrations (A01-B01 and C03-D03), non-GM maize (A09-B09), GM rapeseed (A07-B07), and GM rice (C08-D08). The two clusters that appear in the FAM channel originate from different targets, as the primer and probe concentration of some is higher than of the other (but as there are three targets altogether and only two clusters, they cannot be discriminated). The positive droplets in NTC and non-GM maize sample (in FAM channel) do not make a reaction positive, because the number of such droplets per sample is less than three, and thus reactions are considered negative. The threshold should be set just above the cluster of negative droplets at fluorescence amplitude around 6400 and 2050 in **a** and **b**, respectively (note that these values may vary between the experiments and individual wells). **c** represents the 2D overlay view of analyzed droplets of all samples shown in **a** and **b**. Both thresholds in **c** are dividing the clusters in four groups: negative droplets (bottom left), droplets positive for transgenes (top left), droplets positive for *hmgA* (bottom right), and droplets positive for both the transgenes and *hmgA* (top right)



**Fig. 4** Example of a threshold setting for 10P MQI with NTC, positive control, and different samples. Ch1 (**a**) represents the fluorescence signal originating from FAM, and Ch2 (**b**) represents the fluorescence signal originating from HEX. **a** and **b** represent the 1D view of analyzed droplets for transgene and *hmgA* targets, respectively. Two replicates for each of the following samples are shown: NTC (G06-H06), positive (+) control for transgenes (E05-F05), two unknown samples containing GM maize at different concentrations (E01-F01 and G03-H03), non-GM maize (E09-F09), GM rapeseed (E07-F07), and GM rice (G08-H08). The positive droplet in GM rapeseed sample (in HEX channel) does not make a reaction positive, because the number of such droplets per sample is less than three, and thus reactions are considered negative. Due to the close position of positive and negative clusters with some rain between, the threshold for these reactions was set with lasso tool in 2D view, always comparing NTC, positive control and sample of interest (**c**). In **c** cluster of negative droplets is marked with “a,” cluster of droplets positive for transgenes with “b,” cluster of droplets positive for *hmgA* with “c” and cluster of droplets positive for transgenes and *hmgA* with “d”



**Fig. 4** (continued)



**Fig. 5** Example of MQI results on serial dilution of DNA sample positive for 12 GM maize lines. The droplet readouts in 1D view for 10P and 4P assay are shown on left and right panel, respectively. Dilution factor for DNA sample used in individual reactions is marked above each lane. Ch1 amplitude represents the response in FAM for the group of transgene targets and Ch2 amplitude represents the response in HEX for the *hmgA* target. For 10P analysis the negative and positive droplet clusters were separated using the lasso tool, and therefore no threshold is visible on the readout (lower cluster of droplets represents the negative droplets). For 4P analysis a single threshold was set in each of the fluorescence channels to separate the positive and negative droplet clusters (positive are above the line)

each of the target must be selected with lasso tool to get groups as presented in Fig. 7.

3. After selecting targets in one channel, the export of the results in .csv file must be made. Same process is then repeated for the other two targets in the second channel (*see Note 19*).
4. From each export we get the result (concentrations in target copies per  $\mu\text{l}$  of reaction) for two of the targets:
  - *hmgA* concentration = concentration from Ch1 export 1.
  - MON863 concentration = concentration from Ch2 export 1.
  - DP98140 concentration = concentration from Ch1 export 2.
  - MON810 concentration = concentration from Ch1 export 2.
5. For final percentage content of each of the three GM event targets (listed in Table 3) in each sample the following equation is used, where average is calculated from concentration values of all replicates per undiluted DNA (*see Table 4* for example of calculation):

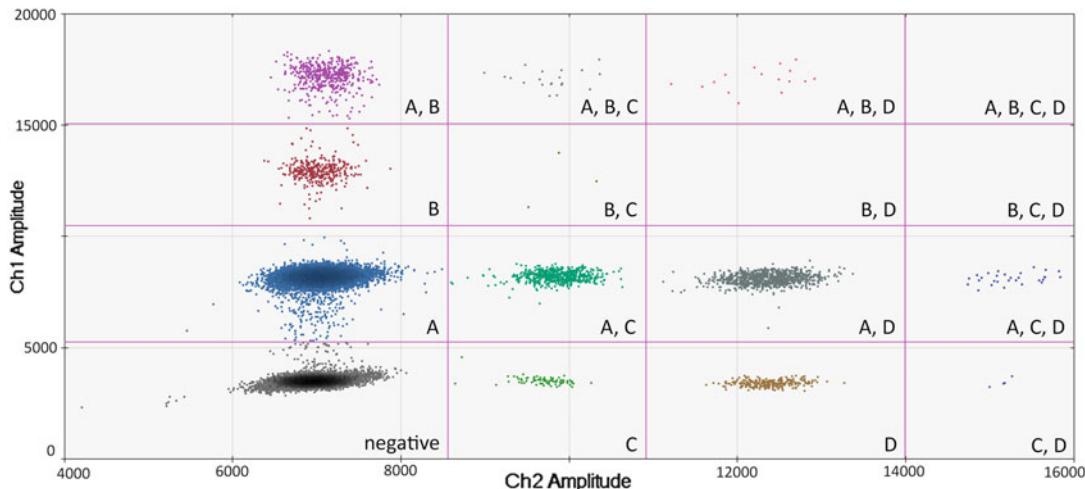
$$\text{GMcontent}[\%] = \frac{\text{average concentration of individual event}}{\text{average concentration of } hmgA} \times 100$$

**Table 4**  
**Examples of calculation of GM content using MQI and MEQ protocols using the formulas provided in the protocol**

Well	Sample	Assay	Target	Concentration (cp/ $\mu$ l) <sup>a</sup>	Concentration of undiluted sample (cp/ $\mu$ l)	Average concentration	GM content (%)
B01	DNA sample 1 ×	MEQ	hm <sub>G</sub> A	1013	1013		
B02	DNA sample 1 ×	MEQ	hm <sub>G</sub> A	1027	1027		
B03	DNA sample 5 ×	MEQ	hm <sub>G</sub> A	200	1000		
B04	DNA sample 5 ×	MEQ	hm <sub>G</sub> A	196	980	1005	
B01	DNA sample 1 ×	MEQ	MON810	89	89		
B02	DNA sample 1 ×	MEQ	MON810	96	96		
B03	DNA sample 5 ×	MEQ	MON810	16	80		= 90/1005 × 100
B04	DNA sample 5 ×	MEQ	MON810	19	95	90	= 8.96
D01	DNA sample 1 ×	4P MQI	hm <sub>G</sub> A	971	971		
D02	DNA sample 1 ×	4P MQI	hm <sub>G</sub> A	1071	1071		
D03	DNA sample 5 ×	4P MQI	hm <sub>G</sub> A	175	875		
D04	DNA sample 5 ×	4P MQI	hm <sub>G</sub> A	178	890		
D05	DNA sample 1 ×	10P MQI	hm <sub>G</sub> A	1084	1084		
D06	DNA sample 1 ×	10P MQI	hm <sub>G</sub> A	869	869		
D07	DNA sample 5 ×	10P MQI	hm <sub>G</sub> A	174	870		

D08	DNA sample 5×	10P MQI	hmgA	171	855	936
D01	DNA sample 1×	4P MQI	3 GM maize lines	50	50	
D02	DNA sample 1×	4P MQI	3 GM maize lines	43	43	
D03	DNA sample 5×	4P MQI	3 GM maize lines	9.6	48	
D04	DNA sample 5×	4P MQI	3 GM maize lines	8	40	45
D05	DNA sample 1×	10P MQI	9 GM maize lines	21	21	
D06	DNA sample 1×	10P MQI	9 GM maize lines	26	26	
D07	DNA sample 5×	10P MQI	9 GM maize lines	7	35	$= \frac{(45 + 28)}{936 \times 100}$
D08	DNA sample 5×	10P MQI	9 GM maize lines	6	30	$= 7.80$
				28		

<sup>a</sup>Output from Quantasoft software

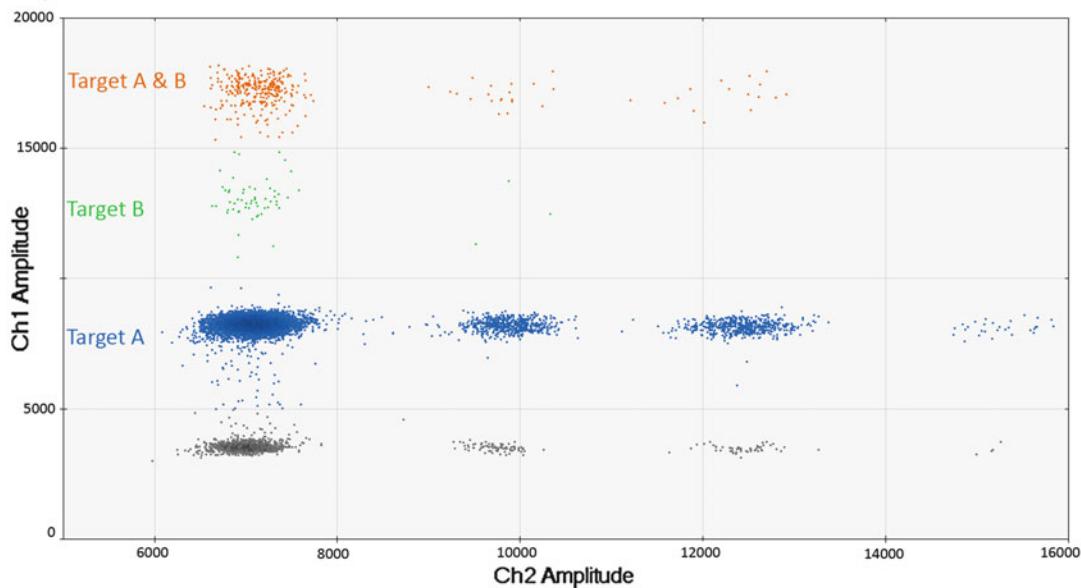


**Fig. 6** Example of possible droplet clusters in MEQ in 2D view of analyzed droplets. The cluster in the left bottom corner contains the fully negative droplets. All other clusters contain the droplets positive for one or more targets. The combination of positive targets in each cluster is marked with letters A (*hmgA*), B (MON863), C (DP98140), and D (MON810). Note that in some clusters there are no positive droplets, although all four targets were present in the sample. To always have an idea of where the clusters should be positioned, a positive control containing high concentration of all the targets must be used in each experiment

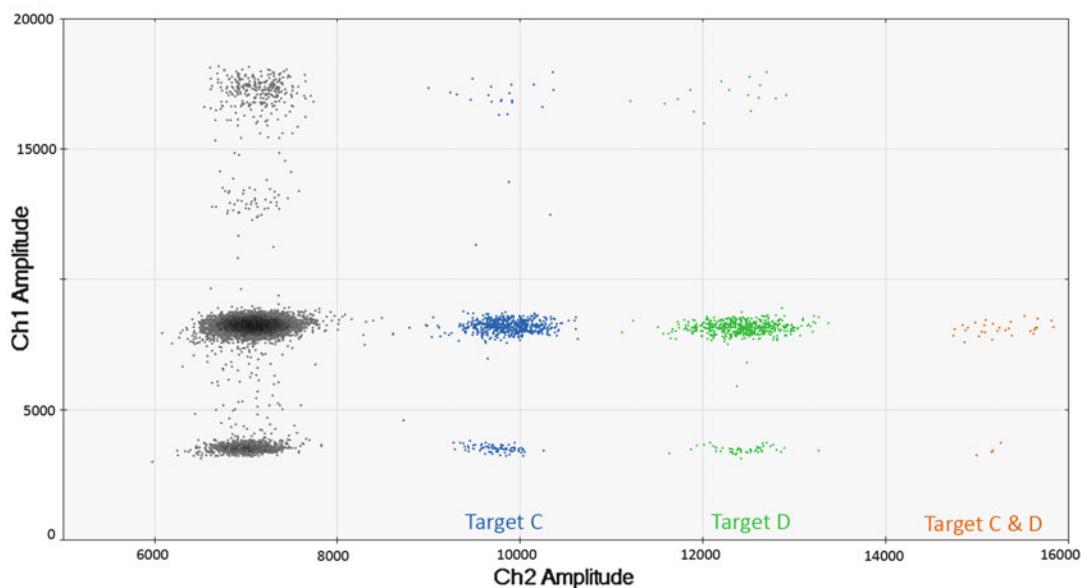
#### 4 Notes

1. Use of the proposed tips is strongly recommended as others can have rough tip openings and can lead to shedding of microscopic plastic debris in the sample reaction, which can cause shearing of the droplets. Thus, the use of Rainin tips is also recommended for upstream sample preparation process and ddPCR reaction setup.
2. The proposed Eppendorf™ 96-Well twin.tec™ PCR Plates must be used, as the use of other plates can have significant deleterious interactions with the droplets.
3. The cycler must enable setting of the ramp rate, which should be set to 2.5 °C/s, to ensure the correct temperature of each droplet during the cycling.
4. The tables give instruction on minimum volume of PP mix that needs to be prepared to avoid pipetting of small volumes. Bigger volume of PP mix can be prepared, aliquoted in smaller volumes (for one time use to avoid several freezing–thawing cycles) and stored at –20 °C. Aliquots can be used in subsequent experiments for at least 3 months. Concentrations of primers and probes in the mix were optimized in a way that all targets are successfully amplified and that the cluster of positive droplets is clearly separated from the cluster of negative droplets (based on the fluorescence).

## Export 1



## Export 2



**Fig. 7** Selecting the cluster groups in MEQ. After selection of cluster groups with lasso tool for the targets A and B the results are exported (Export 1). For the second export the cluster groups of targets C and D are selected (Export 2). See **Note 19** for further explanations

5. To ensure proper loading of 20 µl of reaction mixture into the cartridge without bubbles (which can cause subsequent shearing of droplets), the proposed volume of final reaction mixture is slightly larger (22 µl) than the volume loaded into the cartridge. Additionally, due to small pipetting losses, we also

recommend to prepare a 10% excess volume of Mastermix prior to distributing to individual sample preparation wells.

6. The amount of Mastermix needed for the whole experiment depends on the number of samples to be tested. For each experiment we recommend that individual sample is tested in two different dilutions (DNA concentrations of 50 and 10 ng/ $\mu$ l; 5 $\times$  dilution between samples is recommended), each in duplicate for each individual assay. For example, for MQI that would mean one full cartridge (8 wells) for one sample as two assays (4P and 10P must be run in parallel) and for MEQ two samples can be tested per cartridge. For controls (*see Note 8*) we recommend running them in duplicates per plate (altogether 8 wells).
7. It is important to evaluate the concentration, yield, structural integrity, purity and amplifiability of the extracted DNA to be used as a sample. This could be performed by spectrophotometry, fluorimetry, capillary/gel-electrophoresis and/or by the means of assessment of multiplication efficiency of a taxon-specific target (with real-time PCR). DNA concentration should be high enough to fulfill the requirements of the protocol (at least 50 ng/ $\mu$ l) with yield sufficient to perform all the tests. Lower concentrations might still be used, but then there is the possibility of false negative results. The minimum size of the majority of DNA fragments should be larger than the size of the amplicon produced by the PCR assay. However, less fragmented DNA (large DNA fragments) is recommended for these protocols to reduce the possibility of having breaks within the amplicon. The presence of inhibitors, when performing the DNA extraction and purification according to the instructions, is not problematic for these protocols. When checking the amplifiability of DNA with real-time PCR, the slope of the serial dilution curve should be between -3.1 and -3.6. As stated in **Note 6**, for further ddPCR reactions we recommend use of two different dilutions of the DNA sample.
8. Inclusion of appropriate controls is necessary for reliable results. NTC, positive control, environmental control, and extraction blank control need to be included in duplicates, filling one whole cartridge. For MQI each DNA sample must be tested with 4P and 10P assay to enable quantification of all GM maize events currently authorized in the EU. When a full 96-well plate is to be prepared, for ease of use, we suggest a scheme of pipetting following the columns as one cartridge contains eight wells corresponding to the number of wells per 96-well plate column.
9. Ensure that the sample DNA and the Mastermix were mixed adequately, as the Supermixes are fairly viscous and inadequate mixing will lead to less reproducible results. When using vortexing, use lower speed as vigorous shaking might cause bubble

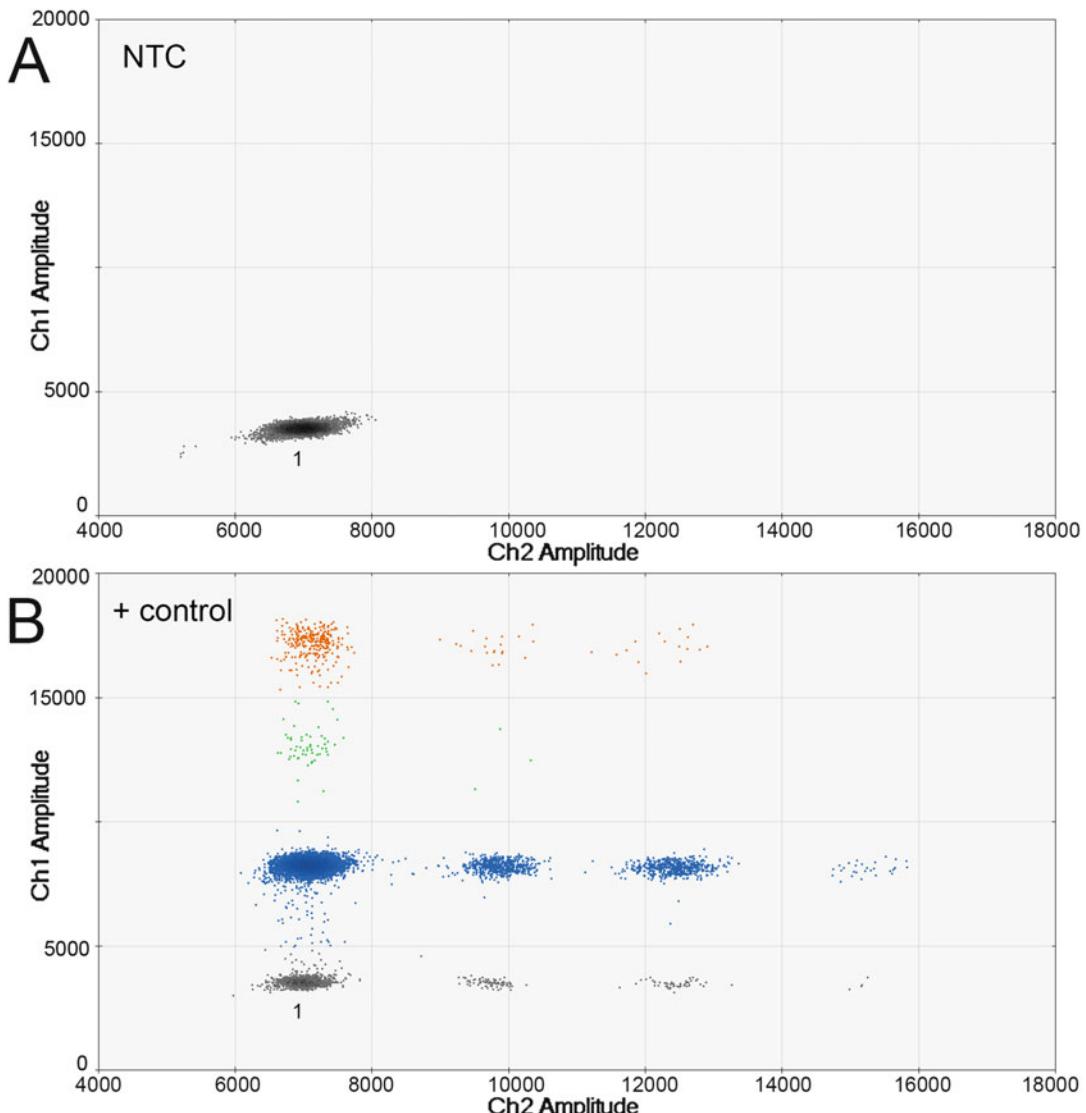
formation, which can reduce the available volume of the reaction mixture that can be reliably pipetted in the following step.

10. It is essential to avoid bubble formation at the bottom of the cartridge wells, as they will disturb the droplet formation. It is also critical to load sample wells first and only then continue with loading of Droplet Generation Oil for Probes. We recommend the use of special 20 µl aerosol-barrier (filtered) Rainin tips and pipetting at around 15° angle to the well wall. In case all wells of the cartridge are not filled with samples, special buffer must be used (ddPCR™ Buffer Control Kit [Bio-Rad, Pleasanton, CA]) in the otherwise empty cartridge wells (or none of your sample wells will form droplets).
11. We suggest using the cooling block as the plate holder, to keep the reaction mixture cold, if several cartridges are planned for the experiment. Keeping the plate with droplets cold also helps to prevent evaporation. Alternatively, this could be done by covering the wells with caps (which need to be removed prior to the sealing of the plate). When transferring the droplets from cartridge to the PCR plate it is essential to use slow pipetting speed to avoid shearing or coalescing of droplets. The use of 8-channel pipette is recommended. We suggest 5 s aspiration by holding the pipette in an angle between 30° and 45° and pressing the tip near the lower edge of the well, but not pressing against the bottom in such a way as to reduce the size of the tip inlet or you will shred your droplets. Dispensing of droplets should also take around five seconds by holding the pipette vertically, touching the lower part of the well wall with the tip (again avoiding pressing the pipette tip against a surface such that the opening is reduced in size).
12. It is of great importance not to centrifuge the plate containing the droplets. If using manual Heat sealer, be careful not to press the sealer to the plate for more than 5 s, as the applied heat could cause deformation of the plastic and evaporation could occur.
13. Final incubation at 98 °C is essential for polymerase deactivation. If used, the plate does not need to be put into droplet reader right after the cycling is completed, but can be stored up to three days at 4 °C.
14. The information important for the analysis is: Sample name, Target name for Target 1 and 2 and Type for Target 1 and 2.
15. There is an option of 1D (individual channel) or 2D (both channels simultaneously) threshold settings. QuantaSoft software also enables automatic threshold detection. For the purpose of this protocol we advise to use 2D chart, as sometimes the clusters might be shifted and manual threshold setting with lasso tool must be used. The threshold should be set manually at the lowest amplitude that captures true negative cluster

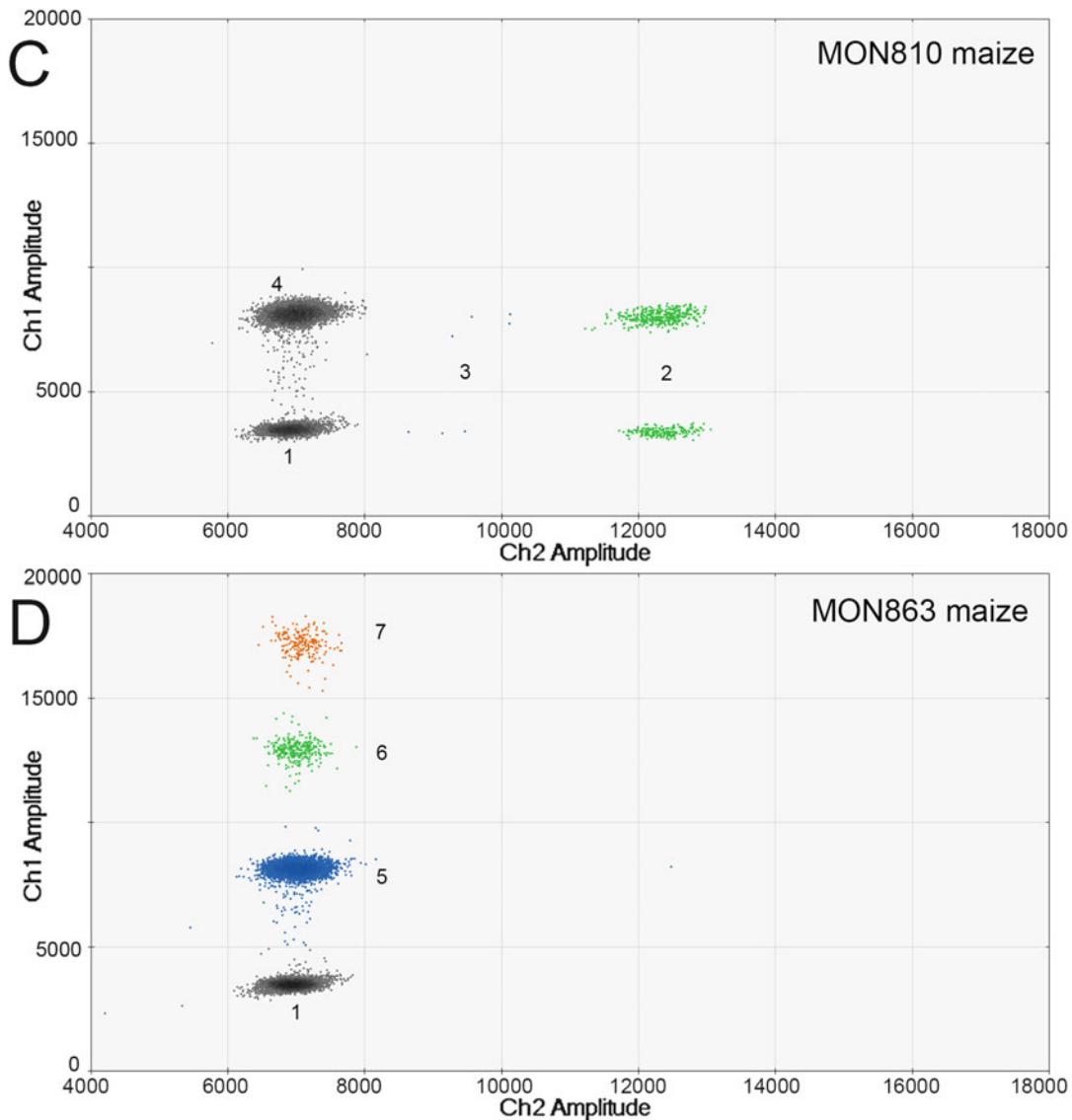
droplets of highest fluorescence amplitude, as visualized using both the fluorescence amplitude vs. event number and the histogram of events vs. amplitude data streams. At the same time, by referring to the positive control cluster amplitude and spread, the threshold should not be drawn so high as to lose true positive droplets (losing sensitivity), nor so low as to produce false positive droplets from the negative droplet cluster (losing specificity). Normally the threshold can be set for a group of wells (with same assay) at the same time, but sometimes slight shifts of fluorescence amplitude can be observed, and thus we recommend some caution when performing the threshold setting for higher number of wells to avoid misidentification of positive or negative droplets.

16. The machine is specified to produce up to 20,000 droplets. However, the actual number recovered and read may vary mostly dependent on droplet pipetting. We believe the readouts below 8000 copies might not be trusted.
17. Through several experiments, we have noticed random occurrence of positive droplets in no template controls (*see* Fig. 3a, NTC lane and non-GM maize lane). However, the number of positive droplets was never higher than two per well. Thus, we use three positive droplets as a cutoff for scoring a well as positive. At the opposite side, a reaction containing high target concentration must produce at least four negative droplets, to enable the calculations with Poisson statistics.
18. The total number of clusters of droplets with this protocol can be up to 16, where clusters have different combinations of targets (Fig. 6). Therefore, setting of one threshold per channel cannot give the result for concentration of individual targets. It is essential to use a positive control containing all four targets to be able to determine the cluster groups for the samples, which may not contain all of the targets. The value of each threshold separating the cluster groups may vary between the experiments, and therefore no fixed value can be offered. The most optimal way to select the individual cluster groups is by using lasso tool.
19. First export (Fig. 7, upper panel) covers all positive droplets in FAM fluorescence channel (droplets positive for target A (*bmgA*) and droplets positive for target B (MON863) and droplets positive for both targets). Second export (Fig. 7, lower panel) covers all positive droplets in HEX/VIC fluorescence channel (droplets positive for target C (DP98140) and droplets positive for target D (MON810) and droplets positive for both targets. Note that the result for targets A and C in the .csv file are marked as Ch1 and targets B and D are marked as Ch2. This does not reflect the actual fluorescent labeling, but is

merely a cause of the multiplex analysis procedure that is different and cannot be performed otherwise. Examples of droplet readout of two different samples, negative and positive control are given in Fig. 8.



**Fig. 8** Example of MEQ results with negative and positive control and two GM maize samples in 2D view. Negative cluster that can be distinguished using negative control (**a**) is marked with (1) in all panels. Positive control (**b**) helps determining the correct position of clusters in the unknown samples. In panel **c** the MON810 DNA was used and the positive clusters for MON810 target (in HEX channel) are marked with (2). Few droplets in the clusters marked with (3) are false positive signal for DP98140, which is actually the rain effect between negative and MON810 clusters. Cluster marked with (4) is considered negative for HEX channel, but represents positive droplets for *hmgA* in FAM channel. In panel D the MON863 DNA was used and the positive clusters are marked with (5), (6), and (7) for *hmgA*, MON863, and both together, respectively



**Fig. 8** (continued)

---

### Acknowledgment

This work was supported by the European Union under grant agreement no. 613908, (project DECATHLON). Complementary financial support was given by the Norwegian Research Council and the Slovenian Research Agency (contract numbers P4-0165 and 1000-15-0105). Consumables for the research resulting in the MEQ protocol were kindly provided by Bio-Rad.

## References

1. Takabatake R, Koiwa T, Kasahara M, Takashima K, Futo S, Minegishi Y, Akiyama H, Teshima R, Oguchi T, Mano J et al (2011) Interlaboratory validation of quantitative duplex real-time PCR method for screening analysis of genetically modified maize. *Food Hyg Saf Sci* 52:265–269
2. Waiblinger H-U, Ernst B, Anderson A, Pietsch K (2007) Validation and collaborative study of a P35S and T-nos duplex real-time PCR screening method to detect genetically modified organisms in food products. *Eur Food Res Technol* 226:1221–1228
3. Sykes PJ, Neoh SH, Brisco MJ, Hughes E, Condon J, Morley AA (1992) Quantitation of targets for PCR by use of limiting dilution. *BioTechniques* 13:444–449
4. Pinheiro LB, Coleman VA, Hindson CM, Herrmann J, Hindson BJ, Bhat S, Emslie KR (2012) Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Anal Chem* 84:1003–1011
5. Baker M (2012) Digital PCR hits its stride. *Nat Methods* 9:541–544
6. Demeke T, Gräfenhan T, Holigroski M, Fernando U, Bamforth J, Lee S-J (2014) Assessment of droplet digital PCR for absolute quantification of genetically engineered OXY235 canola and DP305423 soybean samples. *Food Control* 46:470–474
7. Bhat S, Herrmann J, Armishaw P, Corbisier P, Emslie KR (2009) Single molecule detection in nanofluidic digital array enables accurate measurement of DNA copy number. *Anal Bioanal Chem* 394:457–467
8. Corbisier P, Bhat S, Partis L, Xie VRD, Emslie KR (2010) Absolute quantification of genetically modified MON810 maize (*Zea mays* L.) by digital polymerase chain reaction. *Anal Bioanal Chem* 396:2143–2150
9. Morisset D, Štebih D, Milavec M, Gruden K, Žel J (2013) Quantitative analysis of food and feed samples with droplet digital PCR. *PLoS One* 8:e62583
10. Burns MJ, Burrell a M, Foy C a (2010) The applicability of digital PCR for the assessment of detection limits in GMO analysis. *Eur Food Res Technol* 231:353–362
11. Whale AS, Cowen S, Foy C a, Huggett JF (2013) Methods for applying accurate digital PCR analysis on low copy DNA samples. *PLoS One* 8:e58177
12. European Commission (2003) Regulation (EC) no 1829/2003 of the European Parliament and of the Council of 22 September 2003 on genetically modified food and feed. *Off J Eur Union L* 268:0001–0023
13. Brodmann PD, Ilg EC, Berthoud H, Herrmann A (2002) Real-time quantitative polymerase chain reaction methods for four genetically modified maize varieties and maize DNA content in food. *J AOAC Int* 85:646–653
14. Dobnik D, Spilsberg B, Bogožalec Košir A, Holst-Jensen A, Žel J (2015) Multiplex quantification of 12 European Union authorized genetically modified maize lines with droplet digital polymerase chain reaction. *Anal Chem* 87:8218–8226
15. Demšar T, Gruden K, Milavec M, Štebih D, Žel J (2015) Extraction of DNA from different sample types – a practical approach for GMO testing. *Acta Biol Slov* 58:61–75
16. Dobnik D, Štebih D, Blejec A, Morisset D, Žel J (2016) Multiplex quantification of four DNA targets in one reaction with Bio-Rad droplet digital PCR system for GMO detection. <https://doi.org/10.1038/srep35451>
17. Cankar K, Štebih D, Drečo T, Žel J, Gruden K (2006) Critical points of DNA quantification by real-time PCR—effects of DNA extraction method and sample matrix on quantification of genetically modified organisms. *BMC Biotechnol* 6:37
18. Codex Committee on Methods of Analysis and Sampling (2010) Guidelines on performance criteria and validation of methods for detection, identification and quantification of specific DNA sequences and specific proteins in foods
19. EUR-L-GMFF (2015) Minimum performance requirements for analytical methods of GMO testing
20. Berben G, Charels D, Demšar T, Hochegger R, Nardini E, Onori R, Schulze M, Philipp P, Weber T (2014) Guidelines for sample preparation procedures in GMO analysis. <https://doi.org/10.2788/738570>
21. Bogožalec Košir A, Spilsberg B, Holst-Jensen A, Žel J, Dobnik D (2017) Development and inter-laboratory assessment of droplet digital PCR assays for multiplex quantification of 15 genetically modified soybean lines. *Nature Sci Rep* 7:8601. <https://doi.org/10.1038/s41598-017-09377-w>
22. International Organization for Standardization (2005) ISO 21571:2005 foodstuffs – methods of analysis for the detection of genetically modified organisms and derived products – nucleic

- acid extraction ISO 21571. International Organization for Standardization, Geneva
23. International Organization for Standardization (2006) ISO 24276:2006 foodstuffs – methods of analysis for the detection of genetically modified organisms and derived products – general requirements and definitions ISO 24276. International Organization for Standardization, Geneva
24. Vallone PM, Butler JM (2004) AutoDimer: a screening tool for primer-dimer and hairpin structures. *BioTechniques* 37:226–231
25. Dreo T, Pirc M, Ramšak Ž, Pavšič J, Milavec M, Zel J, Gruden K (2014) Optimising droplet digital PCR analysis approaches for detection and quantification of bacteria: a case study of fire blight and potato brown rot. *Anal Bioanal Chem* 406:6513–6528
26. Jones M, Williams J, Gärtner K, Phillips R, Hurst J, Frater J (2014) Low copy target detection by Droplet Digital PCR through application of a novel open access bioinformatic pipeline, ‘definetherain’. *J Virol Methods* 202:46–53



# Chapter 6

## Using Droplet Digital PCR to Detect Coinfection of Human Herpesviruses 6A and 6B (HHV-6A and HHV-6B) in Clinical Samples

Ashley Vellucci, Emily C. Leibovitch, and Steven Jacobson

### Abstract

Droplet digital™ polymerase chain reaction (ddPCR™) is a unique digital PCR technique that allows for absolute quantification of nucleic acid samples. This technique operates on the basis of amplification within water–oil emulsion droplets and can detect very small quantities of target molecules, yielding extremely precise data. Here, we describe in detail a ddPCR procedure for multiplexed detection of two clinically relevant herpesviruses, HHV-6A and HHV-6B.

**Key words** ddPCR, Droplet digital, Polymerase chain reaction, HHV-6A, HHV-6B, Human herpesvirus 6, Viral detection, Absolute quantification

---

### 1 Introduction

Given the widespread use of PCR, there is an ever-growing drive for innovation and improvement. While early generations of PCR enabled qualitative and relative quantification, a third generation technology, ddPCR, enables absolute quantification. Unlike first- and second-generation PCR methods, ddPCR enables the calculation of absolute copy numbers independent of a standard curve or other external calibrator [1–3]. Positive and negative droplet PCR data are fit to a Poisson algorithm [4], resulting in highly precise quantitative data of targeted nucleic acids. To provide a brief overview of the ddPCR technology, nucleic acid samples are first digested with a restriction enzyme, and then combined with the relevant primers, fluorescent probes, and other standard PCR components. This mixture is emulsified with oil, resulting in its partitioning over approximately 20,000 nanoliter-sized droplets. Each droplet then undergoes amplification by thermal cycling and is assessed for fluorescence level, indicating the internal fluorescence

of the probe [2]. This creates positive and negative droplet populations with a readout of copies/ $\mu$ l.

Through the use of ddPCR, clinically relevant pathogens such as HHV-6 can be easily and accurately detected in clinical samples [1]. As HHV-6 is a ubiquitous virus and can be found in healthy individuals as well as those with HHV-6 associated diseases, a highly quantitative approach is required for the precise determination of HHV-6 viral loads in clinical samples. In this report, we detail two multiplex ddPCR procedures for the detection of HHV-6. A multiplex technique allows for the detection of multiple viruses, while saving money, resources, and time. HHV-6 is becoming an increasingly recognized pathogen in the clinical setting [5], and there are now two recognized species of HHV6, HHV-6A and HHV-6B. Both are associated with several CNS disorders, including epilepsy [6] and encephalitis [7], and HHV-6B is the etiologic agent of roseola, a childhood febrile illness [8]. HHV-6 reactivation with or without a symptomatic manifestation is increasingly recognized in the context of transplantation or other severely immunosuppressive regimens. HHV-6A and HHV-6B were recently reclassified as distinct viral species [9] due to differences in disease associations [8], tropism [10], and other biological and immunological properties [9]. These viruses may also respond differently to antiviral therapy, and it is therefore important in a clinical setting to accurately distinguish between HHV-6A and HHV-6B, particularly as currently available serological assays cannot distinguish between the two viral species [11].

The observation of elevated viral DNA levels in patients with HHV-6 associated diseases compared to controls is the basis for utilizing PCR for HHV-6 detection in clinical samples [8]. In addition, in approximately 1% of the human population, HHV-6 is chromosomally integrated (ciHHV-6) in subtelomeric regions of genomic DNA [5], a distinctive aspect of these viruses. It is unknown whether there are clinical consequences associated with this integration. Clinically, ciHHV-6 patients are often misdiagnosed with reactivated HHV-6, leading to inappropriate and unnecessary treatment and side effects [12]. ddPCR is a reliable tool for the rapid and accurate detection of ciHHV-6, particularly from cell-rich sources such as peripheral blood mononuclear cells (PBMC) [12, 13]. This protocol will detail the methodology for using ddPCR to distinguish between HHV-6A and HHV-6B, with a cellular housekeeping reference gene [2], to detect single or coinfection of these viruses in clinical samples, including samples from individuals with ciHHV-6 [8]. These primers and probes were newly designed for this application, and their initial characterization has been previously described [8].

---

## 2 Materials

### 2.1 DNA Extraction

1. Qiagen DNeasy Blood & Tissue Kit (for DNA extraction from whole blood or isolated PBMC, tissues, CSF cells and saliva).
2. Qiagen QIAamp UltraSens Virus Kit (for DNA extraction from CSF supernatant and serum).
3. NanoDrop 2000 UV-Vis spectrophotometer (Thermo Scientific).

### 2.2 Digest Mix Components

1. Microcentrifuge tubes (0.5–1.5 ml) (*see Note 1*).
2. HindIII restriction endonuclease; 20,000 U/ml (New England Biolabs).
3. NEB 2.1 Buffer (New England Biolabs).
4. PCR water.
5. 37 °C shaker.
6. DNA samples diluted to appropriate concentration (*see Note 2*).
7. Thermomixer.

### 2.3 Super Mix Components

1. Microcentrifuge tubes (0.5–1.5 ml) (*see Note 1*).
2. 2× ddPCR Super Mix for Probes (Biorad).
3. 10× or 20× primer–probe mixes.
  - (a) 20× is defined as a final concentration of 900 nm primers and 250 nm probe.

### 2.4 Droplet Generation and Droplet Reader Components

1. QX200 Droplet Digital PCR (ddPCR) System (Bio-Rad).
  - (a) QX100 Droplet Generator.
  - (b) QX100 Droplet Reader.
2. PCR thermocycler.
3. 96-well plate (*see Note 3*).
4. Heat sealer with block for 96-well plate.
5. DG8 Single-Use Disposable Cartridges (Bio-Rad).
6. DG8 Gaskets (Bio-Rad).
7. Droplet Generation Oil for Probes (Bio-Rad).
8. ddPCR Buffer Control for Probes (Bio-Rad).
9. Reagent reservoir, 25 ml capacity.
10. ddPCR Droplet Reader Oil (Bio-Rad).
11. Low Retention Filter Tips, 200 µl (Rainin).
12. Low Retention Filter Tips, 20 µl (Rainin).
13. Single channel manual pipette LTS lite 2–20 µl (Rainin) (*see Note 4*).

14. Eight channel manual pipette LTS 5–50 µl (Rainin) (*see Note 5*).
15. Eight channel manual pipette LTS 2–200 µl (Rainin) (*see Note 6*).
16. Pierceable sealing foil sheets (Thermo Scientific).

---

### 3 Methods

All procedures are to be carried out at room temperature in a biosafety cabinet and reagents should thaw on ice. Restriction enzyme should be removed from –20 °C immediately prior to use, and returned to –20 °C immediately after use. Fluorescent probes are light sensitive, so lights in the hood should be turned off after the DNA digestion step.

For duplexing, HHV-6A FAM 20× and HHV-6B VIC 20× are run in one well, while RPP30 VIC 20× is run in a separate well (Fig. 1). For triplexing, all primer–probe mixes are run in the same well, with different concentrations of primer–probe mixes on a given channel, i.e., RPP30 VIC 10× and HHV-6B VIC 20× (Fig. 2) (*see Notes 7 and 8*).

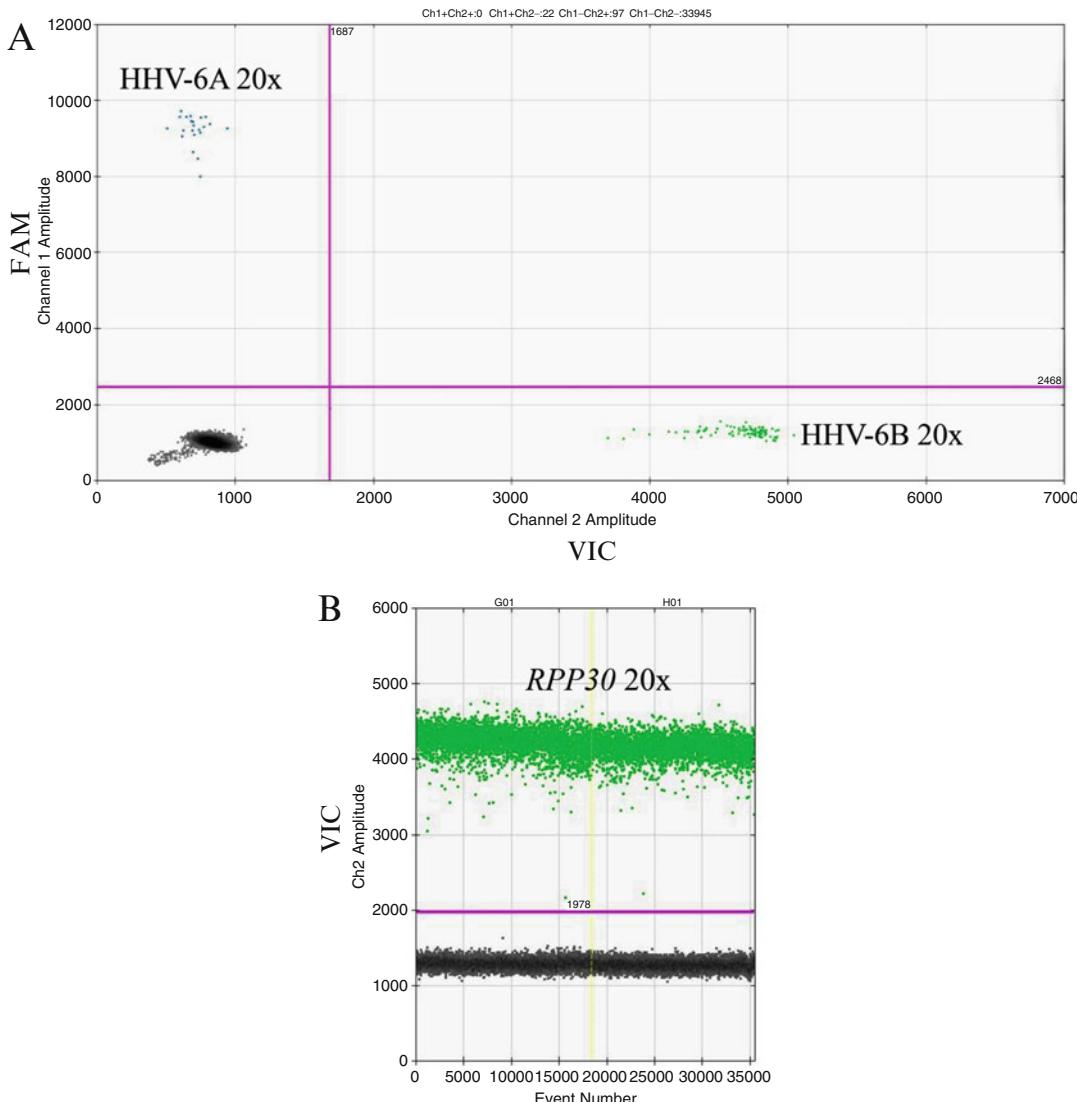
We recommend running samples in duplicate, though triplicate or quadruplicate wells may be used depending on the desired sensitivity.

#### 3.1 DNA Extraction

1. Extract DNA from the designated sample using the appropriate Qiagen kit, in accordance with the manufacturer’s instructions.
2. Obtain DNA concentration using NanoDrop 2000 UV-Vis spectrophotometer. The DNA concentration will vary with the sample type. We typically obtain values of less than 10 ng/µl for CSF cells, serum, and saliva.

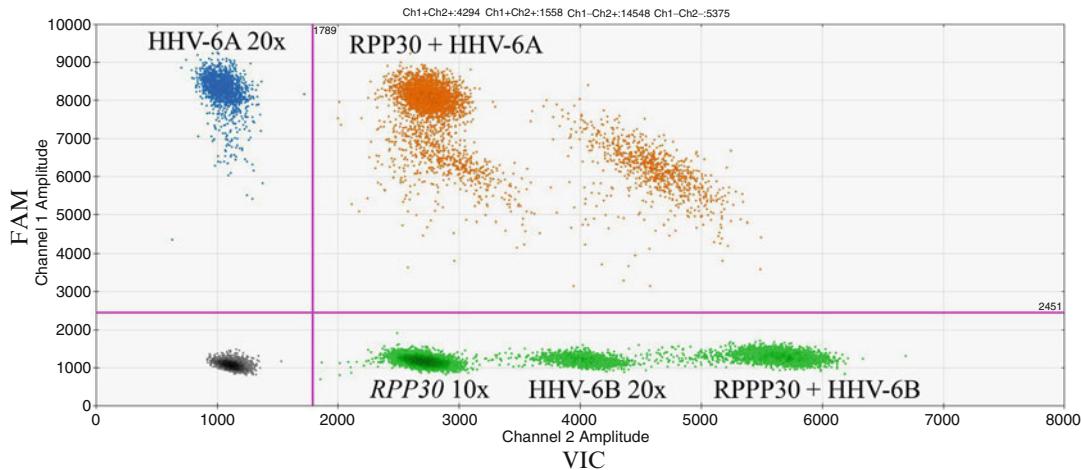
#### 3.2 Reaction Setup

1. Calculate the amount of reagents the supermix and digest mixtures will each contain (*see Note 9*).
  - (a) For one well the digest mix contains:
    - 3.65 µl water.
    - 1 µl 10× NEB Buffer.
    - 0.1 µl BSA.
    - 0.25 µl restriction endonuclease (20 U/µl).
  - (b) For one well the *duplex supermix* contains (Fig. 1):
    - 12.5 µl 2× ddPCR Supermix for Probes (*see Note 10*).
    - 1.25 µl 20× FAM primer–probe mixture.
    - 1.25 µl 20× VIC primer–probe mixture.
  - (c) For one well the *triplex supermix* contains (Fig. 2).



**Fig. 1** HHV-6A and HHV-6B duplexing. **(a)** ddPCR 2D droplet plot generated using HHV-6A 20 $\times$  and HHV-6B 20 $\times$ . HHV-6A is on the FAM channel and the positive droplet population is located in the upper left quadrant. HHV-6B is on the VIC channel and the positive droplet population is located in the lower right quadrant. **(b)** ddPCR 1D droplet plot generated using RPP30 20 $\times$

- 11.25  $\mu$ l 2 $\times$  ddPCR Supermix for Probes (see Note 10).
- 1.25  $\mu$ l 20 $\times$  FAM primer–probe mixture.
- 1.25  $\mu$ l 20 $\times$  VIC primer–probe mixture.
- 1.25  $\mu$ l 10 $\times$  VIC primer–probe mixture (see Note 11).



**Fig. 2** HHV-6A, HHV-6B, and RPP30 triplexing. By differing the primer/probe concentrations, RPP30 ( $10\times$ ) and HHV-6B ( $20\times$ ) can be used simultaneously on the VIC channel, with a clear separation between the VIC positive droplet populations. Populations double positive for RPP30 and HHV-6A or RPP30 and HHV-6B can also be appreciated. Single primer–probe sets should be used to confirm the RPP30/HHV-6A/HHV-6B specificity of the unlabeled populations in the upper right quadrant

**Table 1**  
Primer and probe sequences

	Forward primer (5'-3')	Reverse primer (5'-3')	Probe sequence <sup>a</sup> (5'-3')	Probe fluore <sup>b</sup>	Probe quencher
HHV-6A <sup>7</sup>	CCGTGGGATCGTCT AAAATTATAGATGT	CCACACTAGTC CGGACGGATAA	CTGGAACT GTATAATAGG	6FAM	MGBNFQ
HHV-6B <sup>7</sup>	CCGTGGGATCGTCT AAAATTATAGATGT	CCACACTAGTC CGGACGGCTAA	CTGGAGCT GTACAACAG	VIC	MGBNFQ
RPP30 <sup>2</sup>	GATTGGACCTGCG AGCG	GCGGCTGTCTC CACAAAGT	CTGACCT GAAGGCTCT	VIC	MGBNFQ

<sup>a</sup>See Note 20

<sup>b</sup>See Note 8

2. Calculate the amounts of primers and probe for each  $10\times$  or  $20\times$  mixture. (Sequences are listed in Table 1).
  - (a) For  $100\text{ }\mu\text{l}$  of a  $20\times$  primer–probe mixture (see Note 12).
    - $18\text{ }\mu\text{l}$  primer 1 ( $100\text{ }\mu\text{M}$ ).
    - $18\text{ }\mu\text{l}$  primer 2 ( $100\text{ }\mu\text{M}$ ).
    - $5\text{ }\mu\text{l}$  probe ( $100\text{ }\mu\text{M}$ ).
    - QS with PCR water to  $100\text{ }\mu\text{l}$ .
  - (b) For  $100\text{ }\mu\text{l}$  of a  $10\times$  primer–probe mixture (see Notes 13).
    - $9\text{ }\mu\text{l}$  primer 1 ( $100\text{ }\mu\text{M}$ ).

- 9 µl primer 2 (100 µM).
- 2.5 µl probe (100 µM).
- QS with PCR water to 100 µl.

### **3.3 Digesting the Samples**

Before beginning, turn on plate sealer and PCR thermocycler and set the thermomixer to 37 °C.

1. Dilute DNA samples as needed (*see Note 2*).
2. To a microcentrifuge tube, add 5 µl diluted DNA (30 ng/µl is a recommended DNA concentration for ddPCR) for each sample replicate. For example, add 10 µl DNA if the sample is to be run in duplicate.
3. In a separate microcentrifuge tube, prepare the digestion mixture according to the calculations in Subheading 3.1. The digestion mixture includes PCR water, NEB 2.1 buffer, and HINDIII restriction enzyme.
4. Add the digest mixture to the DNA in a 1:1 ratio. For example, add 10 µl digest mixture to 10 µl DNA. Vortex to mix.
5. Incubate at 37 °C with shaking (approximately 300 rpm) for 30 min.

### **3.4 Preparing and Aliquoting the ddPCR Supermix**

The ddPCR supermix can be prepared while the DNA samples are digesting. Ensure that the lights are off, as the probes are fluorescent.

### **3.5 Diluting the Digested DNA, and Combining with ddPCR Supermix**

1. Label one microcentrifuge tube for each sample.
2. In a separate tube, prepare the ddPCR supermix according to the calculations in Subheading 3.1. This includes the 2× ddPCR Supermix for Probes, and the appropriate 10× or 20× primer–probe mixes.
3. Aliquot 30 µl of the ddPCR supermix into each labeled tube. Place at 4 °C in the dark until use.
1. Dilute the digested DNA with PCR water in a 1:2.5 dilution. For example, add 30 µl PCR water to the 20 µl of digest reaction (containing 10 µl digest mix and 10 µl DNA added in Subheading 3.2) (*see Note 14*).
2. Add 20 µl of the diluted digested DNA to a tube containing 30 µl supermix. This volume is sufficient to run a given sample in duplicate wells. Discard any remaining digested, diluted DNA.

### **3.6 Droplet Generation**

1. Prepare a cartridge holder with a clean DG8 single-use, disposable cartridge.
2. Slowly dispense 20 µl of the DNA/ddPCR supermix mixture into the sample wells of the droplet generator cartridge.

Add 20  $\mu$ l 1 $\times$  Buffer Control for Probes to any wells without DNA-ddPCR supermix.

3. When all eight sample wells across the cartridge are full, slowly add 70  $\mu$ l Droplet Generation Oil for Probes into the cartridge wells labeled for oil.
4. Secure a DG8 gasket over the cartridge and place into the QX100 Droplet Generator (*see Note 15*) and close the lid.
5. When the green light stops flashing, remove the cartridge from the holder, and discard the gasket. Close the droplet generator system while not in use.
6. Manually ensure that 200  $\mu$ l tips are fitted tightly onto each channel of the 50  $\mu$ l multichannel pipette. From the center of the well, very slowly (over approximately ten seconds) draw up the droplets, avoiding any quick movements, and working to minimize air bubbles (*see Note 16*).
7. Dispense the droplets slowly (over approximately 5 s) into one column of the 96-well plate, with the tips lightly contacting the inner wall of each well (*see Note 17*).
8. Repeat steps 1–7 for the remaining DNA-supermix samples (*see Note 18*).
9. When all the samples have been converted to droplets, place the 96-well plate containing the droplets in the heat sealer.
10. Align a pierceable foil seal over the top, with the correct orientation according to the foil sheet instructions, and firmly press down for five seconds. Rotate the plate 180° and firmly press down again for another 5 s.

### 3.7 PCR

PCR should be performed with a 2 °C/s ramp rate to allow for a better heat transfer throughout the droplet oil mixture. Thermal cycling conditions are as follows (*see Note 19*) [2].

95 °C  $\times$  10 min (1 cycle).  
 94 °C  $\times$  30 s, 59 °C  $\times$  60 s (40 cycles).  
 98 °C  $\times$  10 min (1 cycle).  
 12 °C hold.

Once amplified, the plate can be placed into the QX100 Droplet Digital PCR Droplet Reader for analysis. Do not remove the foil seal.

### 3.8 Preparing QuantaSoft™ Template

While the plate is in the PCR thermocycler, the template can be set up using the Bio-Rad QuantaSoft software.

- (a) Select “New” plate, and double click in a well to begin the template set up.

- (b) Highlight all designated columns and select ddPCR supermix for probes in the supermix drop-down menu.
- (c) Enter primer/probe name and concentration ( $10\times$  or  $20\times$ ) for each channel. The type for each channel is “unknown.”
- (d) Click “apply” and ensure that the “U” symbol for each channel appears in each well.
- (e) Save the plate.
- (f) Following thermal cycling, place the 96-well plate into the reader. Select run, set the plate to be read in columns, and select the appropriate fluor for each channel.

### 3.9 Data Analysis

1. For calculation of HHV-6 viral load in whole blood, PBMC or tissue, the copies of HHV-6 are normalized to the copies of the reference gene RPP30, such that the final data are represented as copies/ $10^6$  cells. Before normalization, the RPP30 copies are divided by two, as there are two copies of RPP30 per diploid cell.
2. For calculation of HHV-6 viral load in saliva, CSF supernatant or serum, the copies of HHV-6 are not normalized to the copies of the reference gene RPP30, and the final data are represented as copies/ml original sample. This calculation must take into account the volume of sample used for DNA extraction and the dilution of the DNA throughout the ddPCR procedure.
3. The positive populations for each primer/probe are identified using positive and negative controls with single (i.e., not multiplexed) primer–probe sets.

---

## 4 Notes

1. We use Eppendorf DNA LoBind tubes to increase DNA recovery and reduce interplate and intraplate variability.
2. For a sample run in duplicate wells, the optimal DNA concentration is 100 ng per well. DNA input should not exceed 350 ng/assay ( $n = 2$  wells). This translates to a DNA concentration of  $<35$  ng/ $\mu$ l, if run in duplicate (5  $\mu$ l per well, or 10  $\mu$ l total input). We typically dilute DNA extracted from tissues and cells, but not from relatively acellular compartments such as saliva, CSF, serum or plasma. We typically run samples in duplicate wells, though more wells can be run (in multiples of two) if greater sensitivity is desired.
3. We use the Eppendorf twin.tec PCR 96-well plates.
4. This single channel pipette is used to transfer the DNA–ddPCR supermix mixture to the sample wells of the cartridge.
5. This eight-channel pipette is used for the transfer of droplets from the droplet wells of the cartridge to one column of a 96-well plate.

6. This eight-channel pipette is used for the transfer of droplet generation oil from the reagent reservoir to the oil wells of the cartridge.
7. In order to use multiple primers and probe mixtures on the same channel, it is necessary to vary the concentrations, i.e., 10 $\times$  and 20 $\times$ . This ensures a separation of the positive droplet populations on a given channel. It is advisable to test concentrations between 10 $\times$  and 20 $\times$  for each primer–probe set to determine optimal separation.
8. HEX can be substituted for VIC. The HEX dye is compatible with updated versions of the QuantaSoft software (v1.7.4 and above).
9. When calculating volumes for the digest and PCR supermix mixes, add 10–20% excess to account for volume loss during pipetting and tube transfers.
10. The final concentrations in the PCR reaction will be 1 $\times$  for the supermix, 1 $\times$  for the 20 $\times$  primer–probe mixture, and 0.5 $\times$  for the 10 $\times$  primer–probe mixture.
11. A primer/probe concentration less than 20 $\times$  should be determined for each assay. We determined a 10 $\times$  primer–probe mixture is optimal for separation of HHV-6B and RPP30. However, other primer–probe mixes may require different sequences and it is therefore recommended that the primer–probe mixture concentrations should be tested to determine concentrations that provide optimal separation.
12. For a 20 $\times$  primer–probe mix, the final reaction concentrations are 900 nM for each primer and 250 nM for the probe.
13. For a 10 $\times$  primer–probe mix, the final reaction concentrations are 450 nM for each primer and 125 nM for the probe.
14. The purpose of the dilution is to reduce components present in the digestion buffer that may interfere with PCR amplification.
15. The cartridge holder should lock into a secure position within the QX100 Droplet Generator. A green light indicates successful operation. An amber light may flash if there is an issue with the cartridge holder or the samples within the cartridge. Perform one or more of the following to address a flashing amber light: (a) Ensure there is oil in the wells indicated for oil; (b) Wipe the bottom of the cartridge and the QX100 Droplet Generator using an ethanol-moistened Kimwipes; (c) Place a new gasket on the current cartridge; (d) Transfer samples to a new cartridge.
16. One technique to collect droplets from the cartridge is to lower the pipette tips into the depression of the wells and draw up the contents very slowly (droplets will appear cloudy). The goal is

to minimize the uptake of air, as air bubbles may cause droplet shearing and diminish data quality.

17. We place the pipette tips at a 45° angle and slowly, over approximately ten seconds, release the droplets, with the goal of minimizing droplet shearing. If there are too few droplets per well, a concentration may not appear in Quantasoft due to insufficient droplets for autoanalysis. The accepted droplets should be greater than 10,000; however in our experience, the best data are obtained with a minimum of 15,000 droplets.
18. We use the lid of a pipette tip box to cover the 96-well plate during droplet generation, to prevent intersample contamination and the droplet generation oil from evaporating.
19. A thermal cycling gradient should be performed to find an optimal annealing temperature for new primer–probe sets.
20. One primer set is used to amplify both viruses, while the probes are specific to each virus. This SNP-like assay design enables comparable sensitivity and kinetics across amplifications, while distinguishing between HHV-6A and HHV-6B with a high degree of specificity.

## References

1. Brunetto S, Massoud R, Leibovitch E et al (2014) Digital droplet PCR (ddPCR) for the precise quantification of human T-lymphotropic virus 1 proviral loads in peripheral blood and cerebrospinal fluid of HAM/TSP patients and identification of viral mutations. *J Neurovirol* 20:341–351
2. Hindson B, Ness KD, Masquelier DA et al (2011) High-throughput droplet digital PCR system for absolute quantification of DNA copy number. *Anal Chem* 83:8604–8610
3. Pinheiro L, Coleman V, Hindson C et al (2012) Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Anal Chem* 81:1003–1011
4. Agut H, Bonnafous P, Gautheret-Dejean A (2015) Laboratory and clinical aspects of human herpesvirus 6 infections. *Clin Microbiol Rev* 28:313–335
5. Theodore W, Epstein L, Gaillard W et al (2008) Human herpes virus 6B: a possible role in epilepsy? *Epilepsia* 49:1828–1837
6. Yao K, Honarmand S, Espinosa L et al (2009) Detection of human herpesvirus-6 in cerebrospinal fluid of patients with encephalitis. *Ann Neurol* 65:257–267
7. Leibovitch E, Brunetto G, Caruso B et al (2014) Coinfection of human herpes virus 6A (HHV-6A) and HHV-6B as demonstrated by novel digital droplet PCR assay. *PLOS One*. <https://doi.org/10.1371/journal.pone.0092328>
8. Ablashi D, Agut H, Alvarez-Lafuente R et al (2014) Classification of HHV-6A and HHV-6B as distinct viruses. *Arch Virol* 5:863–870
9. De Bolle L, Van Loon J, De Clercq E et al (2005) Quantitative analysis of human herpesvirus 6 cell tropism. *J Med Virol* 75:76–85
10. Gautheret-Dejean A, Agut H (2014) Practical diagnostic procedures for HHV-6A, HHV-6B, and HHV-7. In: Flamand L, Lautenschlager I, Krueger G, Ablashi D (eds) *Human herpesviruses HHV-6A, HHV-6B & HHV-7*, 3rd edn. Elsevier, Amsterdam, Netherlands, pp 9–34
11. Sedlak R, Cook L, Huang M et al (2014) Identification of chromosomally integrated human herpesvirus 6 by droplet digital PCR. *Clin Chem* 60:765–772
12. Sedak H, Hill J, Nguyen T et al (2016) Detection of HHV-6B reactivation in hematopoietic cell transplant recipients with inherited chromosomally integrated HHV-6A by droplet digital PCR. *J Clin Microbiol* pii: JCM.03275-15



# Chapter 7

## Biomarkers in Cerebrospinal Fluid: Analysis of Cell-Free Circulating Mitochondrial DNA by Digital PCR

Petar Podlesniy and Ramon Trullas

### Abstract

Cerebrospinal fluid (CSF) contains molecules directly linked with brain function because it permeates brain tissue. The analysis of protein biomarkers in CSF is currently recommended for the diagnosis of neurodegenerative disorders, but the clinical sensitivity and specificity are still being investigated. A major drawback is that most of the currently used biomarkers of neurodegenerative diseases are proteins that are found at very low concentrations in CSF and need to be measured by immunoassays that provide relative values, which sometimes are difficult to reproduce between laboratories. In contrast, the recent availability of digital PCR platforms allows the absolute quantification of nucleic acids at single-molecule resolution, but their presence in CSF has not been characterized. CSF contains cell-free mitochondrial DNA (mtDNA) and changes in the concentration of this nucleic acid are linked to neurodegeneration. Here we describe a method to measure the concentration of cell-free circulating mtDNA directly in unpurified CSF using droplet digital PCR with either hydrolysis probes or fluorescent DNA-binding dye methods. This protocol allows the detection and absolute quantification of mtDNA content in the CSF with high analytical sensitivity, specificity, and accuracy.

**Key words** Mitochondrial DNA, Droplet digital PCR (ddPCR), Cerebrospinal fluid (CSF), Inhibitor tolerance

---

### 1 Introduction

Cerebrospinal fluid (CSF) permeates brain tissue and directly reflects the biochemical processes that occur in the brain, assisted by the limitation that the blood brain barrier imposes to the free flow of molecules from brain to other compartments and the opposite way. Consequently, CSF is considered the best fluid to measure and identify new biomarkers for brain diseases [1]. At present, the diagnostic use of CSF biomarkers, although recommended for preclinical diagnoses of neurodegenerative diseases [2], still requires further research, validation, and standardization [3, 4]. Currently, the diagnosis of the different neurocognitive disease subtypes without biomarker evidence is difficult because many of these diseases exhibit similar clinical symptoms of

dementia. Moreover, the majority of these diseases have a long preclinical phase that may last decades. Preclinical diagnosis relies on molecular biomarkers that are hypothesized to be linked with the underlying pathophysiology of the disease. However, the molecular mechanisms that underlie the majority of neurocognitive and neuropsychiatric diseases are still not well known, which hampers not only disease diagnosis but also the discovery of drugs capable of modifying the course of these diseases.

The pathophysiological biomarkers in the CSF currently used for preclinical diagnosis of neurodegenerative diseases are proteins which include amyloid beta 1-42 ( $\text{Ab}_{1-42}$ ), total levels of tau (t-tau), phosphorylated tau (p-tau), alpha-synuclein, and 14-3-3 proteins. A major difficulty for the diagnostic use of these biomarker proteins is that in body fluids some of them are found at very low concentrations, which fall near the inferior limit of the analytical sensitivity range of currently available protein detection techniques [5]. In addition, a major challenge in protein assays of body fluids in clinical analyses is the inability to establish absolute values that are reproducible and comparable between laboratories, mainly because protein quantification is based on antibody immunoassays that need to be referenced to an external standard and quantification requires a comparison with a standard calibration curve that is generally different amongst laboratories. An additional drawback is that the antibodies used for protein analysis may change significantly between different lots; for example, the epitopes recognized by the antibodies are generally unknown, the target epitopes differ markedly amongst polyclonal antibodies, and when monoclonal antibodies are used, the quality between different lots may vary. Another general source of variability for protein quantification in bodily fluids is the degree of protein integrity due to sample extraction, handling and long term storage. All these factors introduce a significant amount of variability in protein quantification assays in CSF or other bodily fluids that lessens reproducibility between laboratories.

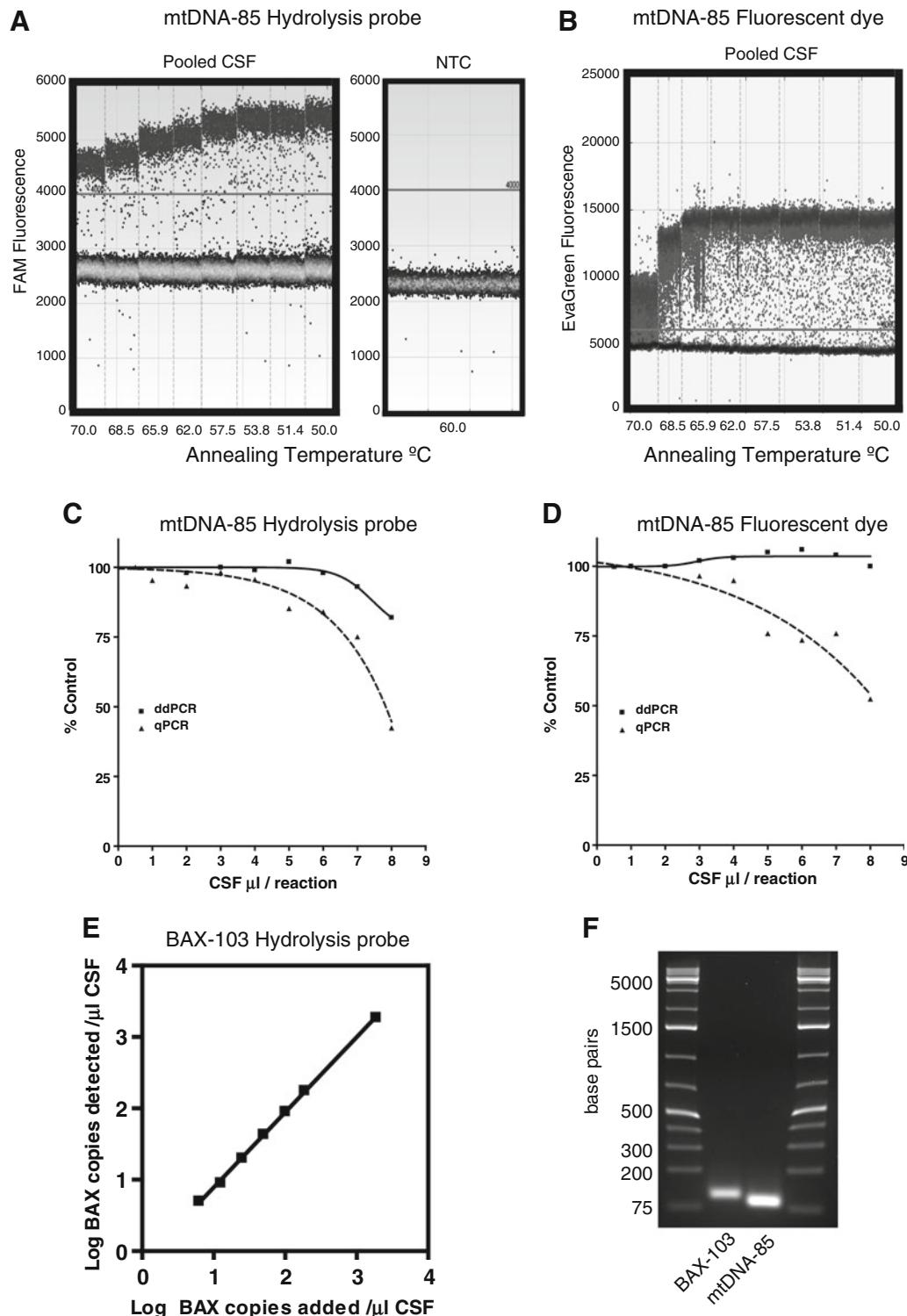
To overcome the technical difficulties associated with protein quantification in CSF samples, we decided to shift the focus from proteins to deoxyribonucleic acid (DNA) because the latter may be less susceptible to degradation due to sample manipulation or long term storage. We opted to detect mitochondrial DNA (mtDNA) because we hypothesized that the amount of mtDNA in CSF could be an index of brain injury or metabolism. Based on this hypothesis, we found that a drop in mtDNA content in CSF is a biomarker for preclinical Alzheimer's disease [6].

There are several advantages of measuring mtDNA over proteins as a biomarker in CSF. In theory, the circular nature of intact mtDNA should make it more resistant to potential degradation by endonucleases, which are likely present in CSF and other body fluids. Another advantage of mtDNA over protein in diagnostic

assays is that the former is amenable to detection and quantification with PCR amplification techniques, which allow higher analytical sensitivity and specificity than immunoassay procedures. Indeed, the launch of digital PCR has made possible the precise detection and absolute quantification of nucleic acid content at a single-molecule resolution of target sequence [7, 8]. This is achieved by performing an end-point PCR reaction in a large number of sample partitions and applying Poisson analysis to the fraction of positive partitions for quantification. The ability to perform absolute quantification offered by digital PCR, avoiding the use of external reference standards, markedly improves repeatability and reproducibility of biomarker measurements. In addition, one of the advantages of the digital PCR protocol for quantifying the amount of mtDNA in CSF described here is that it does not require previous extraction or processing of the CSF sample, because we found that the presence of inhibitory molecules in CSF does not significantly influence end-point digital PCR quantification (Fig. 1c, d). Another important benefit of digital PCR is analytical sensitivity because it allows higher precision measurements than qPCR at low concentrations of the target sequence.

Droplet digital PCR (ddPCR) incorporates all the advantages of digital PCR by partitioning the reaction in oil emulsion droplets and analyzing the fraction of droplets that exhibit positive end-point PCR amplification [9, 10]. Here we describe a protocol to measure by ddPCR the concentration of cell-free circulating mtDNA directly in unpurified CSF using either hydrolysis probe or fluorescent DNA-binding dye methods with primers that produce an amplicon of 85 base pairs (mtDNA-85). A common problem of CSF samples is that some of them may be contaminated with cells that provide an erroneous measurement of cell-free mtDNA. To circumvent this problem, the hydrolysis probe protocol includes the simultaneous measurement of cell-free mtDNA and a nuclear gene in a multiplex ddPCR assay using the mtDNA-85 primers together with primers that amplify an amplicon of 103 base pairs (BAX-103) corresponding to the apoptosis regulator BAX isoform alpha. Alternatively, other single copy nuclear genes can be used for this purpose. The presence of at least one copy of the nuclear gene per microliter of CSF indicates that the sample is contaminated with genomic DNA and is excluded from analysis.

Characterization of the ddPCR reaction amplifying mtDNA in CSF with our protocol shows that the optimum temperature to reach the maximum separation between positive and negative droplets is 60 °C for both the hydrolysis probe (Fig. 1a) and the fluorescent dye (Fig. 1b). The optimal temperature for amplification of the BAX-103 amplicon is also 60 °C (data not shown) which allows the multiplex assay of this nuclear gene and mtDNA with hydrolysis probes. The volume of unpurified CSF sample that can be added to the ddPCR reaction without inhibition of the



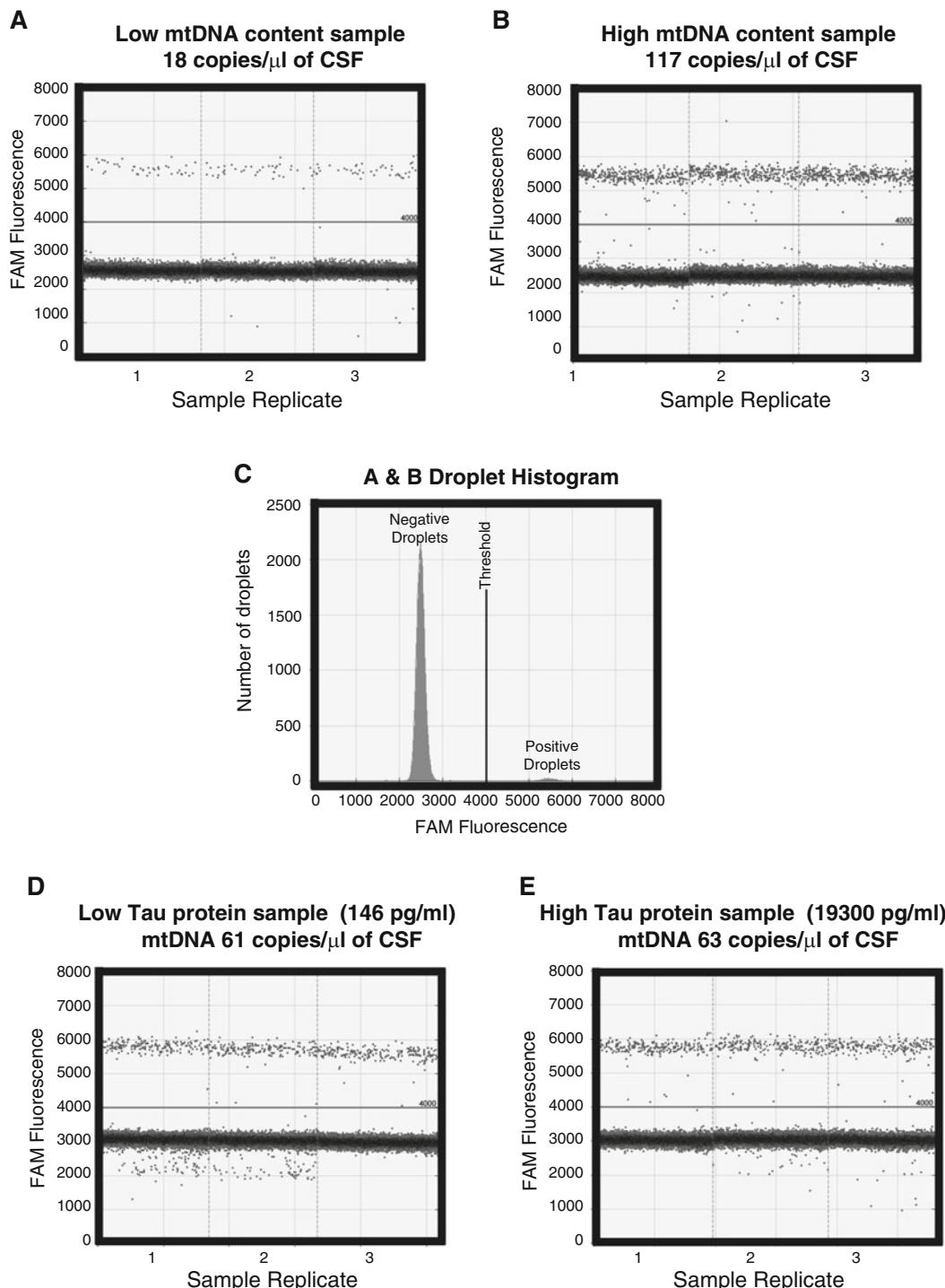
**Fig. 1** Analysis of cell-free mtDNA directly in unpurified CSF using hydrolysis probes or EvaGreen Fluorescent DNA-binding dye: Comparison between ddPCR and qPCR. Assay conditions were studied in a human CSF

amplification due to factors present in CSF can be up to 6  $\mu$ l in a total reaction volume of 20  $\mu$ l. In contrast, this volume of CSF causes significant inhibition (approximately 25%) in qPCR reactions (Fig. 1c, d). In the ddPCR protocol we choose to use 4.5  $\mu$ l of CSF because we found that this volume provides the highest signal with the lowest possibility of amplification inhibition. To check the possibility that, when individual CSF samples contain a high concentration of potential PCR inhibitors, this volume of CSF might inhibit the reaction, we measured CSF samples from subjects in an advanced stage of a rapid neurodegenerative disease that contain a very high content of neuronal death biomarkers such as tau (Fig. 2d, e). Despite a difference in approximately two orders of magnitude in the concentration of tau and possibly of many other proteins, we found no significant differences in the ddPCR reaction profile between these samples and the corresponding control samples. The absence of inhibition in the ddPCR is exemplified by a well-defined separation between the populations of positive and negative droplets in both low and high tau protein samples. In addition, the mean absolute values of FAM fluorescence of positive droplets are similar in low and high tau samples, showing that the PCR amplification reaches a similar endpoint in both reactions and that there are no inhibitory factors that differentially alter PCR in the samples (Fig. 2d, e).

The dynamic range, linearity, accuracy, and analytical sensitivity of our assay to measure the content of mtDNA and BAX in CSF were assessed by adding different known concentrations of the purified PCR amplicons to a sample of pooled CSF. Measurement of mtDNA or BAX in a sample volume of 4.5  $\mu$ l of CSF is linear up

---

**Fig. 1** (continued) sample pool from several subjects. (a and b) Temperature gradient optimization assays were performed in a ddPCR reaction amplifying the 85 base pairs mtDNA amplicon (mtDNA-85) using a hydrolysis probe (a) or the EvaGreen fluorescent DNA binding dye (b). The maximum separation between positive (upper) and negative (lower) droplets for EvaGreen is reached at 66 °C and is constant between 50 and 66 °C, whereas for the hydrolysis probe assay, maximum separation is reached at approximately 60 °C and remains constant down to 50 °C. We choose 60 °C for maximum stringency in both assays, either in duplex or singleplex forms. In nontemplate controls (NTC) there is a nonsignificant number of positive (upper) droplets. (c, d) The influence of CSF volume in the PCR reaction was studied in hydrolysis probe (c) and EvaGreen fluorescent dye (d) assays using either ddPCR (square dots, continuous line) or qPCR (triangular dots, discontinuous line). In contrast to qPCR, addition of up to 6  $\mu$ l of CSF to the PCR reaction did not inhibit ddPCR assay. (e) Accuracy and linearity of ddPCR reaction to directly measure mtDNA in CSF were assayed by adding different known concentrations of purified BAX-103 amplicons to a pooled CSF sample. Detection of mtDNA directly in CSF up to 3000 copies/ $\mu$ l fits to a linear model ( $r^2 = 0.999$ ). The detection limit under these conditions is less than 1 copy of target per microliter of CSF with a signal above the 95% confidence interval of average noise in nontemplate controls. (f) Representative image of an agarose gel showing a DNA ladder 75–20,000 base pairs (bp) and the Bax-103 (Lane 1) and mtDNA-85 (Lane 2) amplicons obtained in a CSF sample



**Fig. 2** Representative examples of results from ddPCR analyses of various CSF samples containing different levels of mtDNA or potential PCR inhibitors. The mtDNA content was measured in 4.5  $\mu$ l of CSF sample, using the mtDNA-85 mix and the ddPCR procedure for hydrolysis probes. (a and b) Representative one-dimensional scatter plots of three replicates from a CSF sample containing a low concentration of mtDNA (a) and a different

to 3000 copies/ $\mu$ l of CSF, indicating that in the conditions of our ddPCR reaction there is no influence of the inhibitory factors present in the CSF on the multiplex detection of mtDNA and BAX as exemplified by the measurement of added copies of BAX into a pooled CSF sample (Fig. 1e). The limit of detection of our ddPCR reaction was close to the theoretical limit of three copies of target per reaction, corresponding to less than one copy of target per microliter of CSF with a signal above the 95% confidence interval of average noise in nontemplate controls. Agarose gel electrophoresis analyses confirmed that the ddPCR reaction conditions used in our experiments produced single amplicons of the corresponding size, 103 and 85 base pairs for Bax-103 and mtDNA-85 primer combinations, respectively (Fig. 1f).

## 2 Materials

### 2.1 Droplet Digital PCR Using Hydrolysis Probes

1. Forward and reverse primers and hydrolysis probe targeting mtDNA (*see Note 1*). Forward: mtDNA-85F, (5'-CTCAC TCCTGGCGCCTGCC-3'); reverse: mtDNA-85R (5'-GGC GGTTGAGGCCTCTGGTG-3'); hydrolysis probe: FAM-mt DNA-85P (6-carboxyfluorescein (FAM)-5'-CCTCCAAATCA CCACAGGACTATTCTAGCCATGCA-3'-Black Hole Quencher-1(BHQ-1)).
2. Forward and reverse primers and hydrolysis probe targeting a single copy nuclear gene, e.g., BAX. Forward: BAX-103F, (5'-TTCATCCAGGATCGAGCAGG-3'); reverse: BAX-103R, (5'-TGAGACACTCGCTCAGCTTC-3') and hydrolysis probe: HEX-BAX-103P, (6-carboxy-2',4',5',7,7'-hexachlorofluorescein (HEX)-5'-CCCGAGCTGGCCCTGGACCCGGT-3'-BHQ1). (*see Notes 2 and 3*).

**Fig. 2** (continued) CSF sample containing a high concentration of mtDNA (b). Good repeatability is observed between replicates. The amount of mtDNA concentration does not influence the well-defined separation between positive (upper) and negative (lower) droplets. In addition, the proportion of positive/negative droplets is equivalent among replicates. The line drawn at 4000 fluorescence units represents the threshold used to define positive and negative droplets. (c) Histogram of total positive and negative partitions from assays shown in a and b. Threshold is determined by choosing a point between the two droplet partitions, for example, adding the mean fluorescence value of negative droplets plus the mean fluorescence value of positive droplets and dividing the result by 2. (d, e) Representative one-dimensional scatter plots of three replicates from a CSF sample containing a low concentration of tau protein (d) and a different CSF sample containing a high concentration of tau protein (e). In addition to different levels of tau, these two samples may contain also different amounts of other potential PCR inhibitors. Both samples exhibit a similar concentration of mtDNA and similar fluorescence values of positive and negative droplet populations, showing the resistance of ddPCR to the presence of potential PCR inhibitors

3. ddPCR Supermix for Probes (No-dUTP) (catalog number 186-3023) (Bio-Rad, Hercules, CA, USA).
4. ddPCR Droplet Generation Oil for probes (catalog number 186-3005) (Bio-Rad).

## **2.2 Droplet Digital PCR Using Fluorescent DNA-Binding Dyes**

1. Forward and reverse primers targeting mtDNA (*see Note 1*). Forward: mtDNA-85F, (5'-CTCACTCCTTGGCGCCTGCC-3'); reverse: mtDNA-85R, (5'-GGCGGTTGAGGCCTCTGG TG-3').
2. Forward and reverse primers targeting a single copy nuclear gene, e.g., BAX. Forward: BAX-103F, (5'-TTCATCCAGGAT CGAGCAGG-3'); reverse, BAX-103R (5'-TGAGACACTCG CTCAGCTTC-3').
3. QX200™ ddPCR EvaGreen Supermix (catalog number 186-4033) (Bio-Rad). (*see Note 3*)
4. QX200™ Droplet Generation Oil for EvaGreen (catalog number 186-4005) (Bio-Rad).

## **2.3 Common Materials to Assay Cerebrospinal Fluid Using ddPCR**

1. Cerebrospinal fluid sample (*see Note 4*).
2. Calibrated micropipettes and low retention, presterilized, aerosol barrier filter pipette tips, (Biotix, San Diego, CA, USA) (*see Note 5*).
3. Polypropylene PCR tubes, 0.2 ml.
4. Vortex mixer.
5. Microcentrifuge (5000 × *g*).
6. DG8™ Cartridges for QX200™/QX100™ Droplet Generator (catalog number 186-4008) (Bio-Rad).
7. DG8™ Cartridge Holder (catalog number 186-3051) (Bio-Rad).
8. DG8™ Gaskets for QX200™/QX100™ Droplet Generator (catalog number 186-3009) (Bio-Rad).
9. Electronic multichannel pipette with wide bore tips (*see Note 6*).
10. Eppendorf twin.tec® 96-Well PCR plates, semiskirted (catalog number 0030 128.XXX) (Eppendorf).
11. PX1™ PCR Plate Sealer (catalog number 186-4000) (Bio-Rad).
12. Pierceable Foil Heat Seal (catalog number 181-4040) (Bio-Rad) (*see Note 7*).
13. C1000 Touch™ Thermal Cycler with 96-Deep Well Reaction Module (catalog number 185-1197) (Bio-Rad).
14. ddPCR Droplet Reader Oil (catalog number 186-3004) (Bio-Rad).
15. QX200™ Droplet Digital™ PCR System (catalog number 186-4001) (Bio-Rad).
16. QuantaSoft™ Software (catalog number 186-4011) (Bio-Rad).

### 3 Methods

#### **3.1 Sample Acquisition and Processing**

CSF should be collected and processed following standard operating procedures after informed consent of the patient. In our studies, CSF was collected from subjects recruited at the Alzheimer's disease and other cognitive disorders unit of the Hospital Clinic of Barcelona. The CSF samples were obtained with informed consent at the Alzheimer's disease and other cognitive disorders Unit of the Neurology Service, following the procedure approved by the ethics committee of the Hospital Clinic of Barcelona. CSF was obtained by lumbar puncture between 9 a.m. and 12 p.m. A small sample of the first CSF volume was reserved for later measurement of cell contamination and this was followed by collection of 10 ml of CSF. The whole volume of CSF was centrifuged ( $2000 \times g/4 {^\circ}\text{C}/10 \text{ min}$ ), aliquoted after centrifugation in 500  $\mu\text{l}$  polypropylene tubes and stored at  $-80 {^\circ}\text{C}$  within two hours of the beginning of the collection. The mtDNA assay was performed in unpurified CSF that underwent only one thaw. It is convenient to avoid using CSF samples that have undergone repeated freeze-thaw cycles. However, we find that up to three freeze-thaw cycles do not significantly modify the measured concentration of mtDNA by ddPCR.

#### **3.2 Digital PCR Reaction Setup and Amplification**

Bring all reagents and samples to room temperature before use. Centrifuge briefly, mix thoroughly by vortexing and centrifuge briefly again to properly mix and collect the components to the bottom of the tubes.

1. Prepare a master mix in a polypropylene microcentrifuge tube containing ddPCR Supermix, oligonucleotide mix, and double-distilled nuclease-free water in a total volume sufficient to obtain 15.5  $\mu\text{l}$  per assay as shown in Table 1 (*see Note 8*). It is essential to include nontemplate control assays, with master mix but adding nuclease-free dH<sub>2</sub>O instead of CSF sample (*see Note 9*). In addition, for assay characterization, it is necessary to include positive control samples containing known amounts of nuclear DNA or mitochondrial DNA target amplicons cloned in vectors or purified from PCR reactions (*see Note 10*). The process of assay characterization includes assessment of accuracy, sensitivity, setting the threshold for positive droplets and the level of false positive droplets for mtDNA and nuclear DNA respectively.
  - (a) The ddPCR assay using hydrolysis probe is designed as a multiplex assay to monitor simultaneously mtDNA and BAX in the same reaction. The composition of the master mix for ddPCR using hydrolysis probes is as shown in Table 1 (*see Note 11*):

**Table 1**  
**Master mix for ddPCR using hydrolysis probes (see Note 8)**

Component	Initial concentration	Final concentration	Volume per reaction (20 µl)
ddPCR Supermix for Probes (No-dUTP) (186-3023)	2×	1×	10.0 µl
Oligonucleotide Mix in H <sub>2</sub> O:			
mtDNA-85F	18.0 µM	0.9 µM	1.0 µl
mtDNA-85R	18.0 µM	0.9 µM	
FAM-mtDNA-85P	2.0 µM	0.1 µM	
BAX-103F	18.0 µM	0.9 µM	
BAX-103R	18.0 µM	0.9 µM	
HEX-BAX-103P	6.0 µM	0.3 µM	
dH <sub>2</sub> O, nuclease free			4.5 µl
Total volume/assay			15.5 µl

**Table 2**  
**Master mix for ddPCR using EvaGreen fluorescent dye (see Note 8)**

Component	Initial concentration	Final concentration	Volume per reaction (20 µl)
QX200™ ddPCR EvaGreen Supermix (186-4033)	2×	1×	10.0 µl
Oligonucleotide mix in H <sub>2</sub> O:			1.0 µl
mtDNA-85F	2.0 µM	0.1 µM	
mtDNA-85R	2.0 µM	0.1 µM	
or			
BAX-103F	2.0 µM	0.1 µM	
BAX-103R	2.0 µM	0.1 µM	
dH <sub>2</sub> O, nuclease free			4.5 µl
Total volume/assay			15.5 µl

- (b) The ddPCR assay of mtDNA and the nuclear gene BAX using fluorescent DNA binding dyes is performed in separate reactions. The composition of the master mix for ddPCR using EvaGreen fluorescent dye is as shown in Table 2 Master mix for ddPCR using EvaGreen fluorescent dye (see Note 8).

2. Distribute 15.5  $\mu$ l of the master mix with a 20  $\mu$ l pipette to 0.2 ml Polypropylene PCR tubes (*see Note 8*).
3. Add 4.5  $\mu$ l of unpurified CSF sample to the tube (*see Note 8*).
4. Mix well by vortexing and centrifuge briefly to collect the mix to the bottom of the tube and remove bubbles. The Supermix is viscous and insufficient mixing will lead to less reproducible results.
5. Place the DG8<sup>TM</sup> Cartridge for Droplet Generator into the DG8<sup>TM</sup> Cartridge Holder.
6. Transfer each 20  $\mu$ l of reaction mixture to one of the eight sample wells in the middle of the DG8<sup>TM</sup> Cartridge (*see Note 8*).
7. Add 70  $\mu$ l of either ddPCR Droplet Generation Oil for probes (for the probe hydrolysis method) or QX200<sup>TM</sup> Droplet Generation Oil for EvaGreen (for the fluorescent DNA-binding dye method) to the oil wells of the DG8<sup>TM</sup> Cartridge adjacent to the samples (*see Note 12*).
8. Cover the cartridge with a DG8<sup>TM</sup> Gasket for QX200<sup>TM</sup>/QX100<sup>TM</sup> Droplet Generator.
9. Place the covered cartridge into the QX100/QX200<sup>TM</sup> Digital Droplet Generator to generate the droplets (*see Note 13*).
10. Transfer 40  $\mu$ l of the emulsified sample to an Eppendorf twin.tec<sup>®</sup> 96-Well semiskirted PCR plate with an electronic multi-channel pipette set on the lowest aspirating and dispensing speed using wide bore tips. Make sure to keep the pipet at a 15°–30° angle when aspirating the emulsified sample to avoid damaging the droplets.
11. Cover the 96-well PCR plate with pierceable foil heat seal and place it into PX1<sup>TM</sup> plate sealer. Seal at 170 °C for 4 s.
12. Place the Eppendorf twin.tec<sup>®</sup> semiskirted 96-Well PCR plate into a PCR thermal cycler that has the ability to adjust a ramp temperature and that is compatible with deep wells. Set ramp rate to 2 °C per second for all steps.
13. Use the following PCR profiles (*see Note 14*):
  - (a) For hydrolysis probes:

95 °C: 10 min.  
40 cycles:  
94 °C: 30 s.  
60 °C: 1 min.  
98 °C: 10 min.  
4 °C: infinite.

(b) For fluorescent dyes (*see Note 15*):

95 °C: 5 min.

40 cycles:

95 °C: 30 s.

60 °C: 1 min.

4 °C: 5 min.

90 °C: 5 min.

4 °C: infinite.

14. After PCR the plate can be left in the thermal cycler overnight or stored for a few days at 4 °C before droplet reading.

### **3.3 Digital Data Acquisition**

1. Power on the QX200™ Droplet Digital™ PCR Droplet Reader and let it warm up for at least 30 min, making certain that there is enough ddPCR Droplet Reader Oil in the reader compartment. Prime the system if the instrument has been unused for longer than a week.
2. Place the 96-well PCR plate to the QX200™ Droplet Digital™ PCR Droplet Reader.
3. Introduce the appropriate settings in the PCR Droplet reader software indicating the PCR supermix used and the fluorescence acquisition channels (*see Note 15*).

### **3.4 Data Analysis**

1. After reading all samples, open the file with the appropriate QuantaSoft software and in the Experiments window select Absolute Quantification method (ABS) and then select all samples for data analysis.
2. Click on the Analyze tab and then the multiwell thresholding icon. Choose a point between the two droplet populations shown in the fluorescence histogram (Fig. 2c) to identify a threshold value that reliably separates positive and negative droplets. For example, add the mean fluorescence value of negative droplets plus the mean fluorescence value of positive droplets and divide the result by 2. Enter the resulting value into the Set Threshold window (*see Note 16*)
3. Determine target concentrations by analyzing the proportion of negative versus positive droplets using the Absolute Quantification method (*see Note 17*). Replicates that do not exhibit equivalent mean fluorescence values of positive and negative droplets between them indicate a droplet generation failure and should be discarded. When all replicates of one sample exhibit equivalent mean fluorescence that is significantly lower than the exhibited by the replicates of another sample, this indicates inhibition of the PCR and values cannot be compared.

4. Obtain the total number of copies per 20  $\mu$ l of reaction for each target and divide by the volume (e.g., 4.5  $\mu$ l) of CSF introduced into the reaction to obtain copies per microliter of CSF.

---

## 4 Notes

1. These primers target an mtDNA sequence that is not present in the nuclear genome as a nuclear mtDNA (Numt). When designing primers targeting other mtDNA sequences it is important to check that these sequences do not have a corresponding Numt.
2. The reporter fluorophores used to label the hydrolysis probes can be FAM combined interchangeably with either HEX or VIC in any of the probes. We chose to use FAM for mtDNA and HEX for BAX, but it could also be the opposite. It is likely that in the near future there will be multichannel digital PCR platforms that will allow more fluorophore combinations for multitarget detection. We use the dark quencher BHQ-1 because it is capable of quenching both FAM and HEX with high efficiency. We found that quenchers such as TAMRA do not completely quench probe fluorescence and do not permit a clear separation between positive and negative droplets.
3. The oligomix with fluorophores or EvaGreen supermix should not be exposed to light for prolonged periods of time. We cover with aluminum foil all tubes that contain fluorophores.
4. After centrifugation of the whole volume of CSF obtained by lumbar puncture, the unpurified CSF should be placed in sterile siliconized polypropylene microcentrifuge tubes to avoid mtDNA absorption to the tube, and stored at  $-80^{\circ}\text{C}$  in small aliquots (e.g., 500  $\mu$ l) to avoid unnecessary transfer to different tubes or freezing–thawing cycles.
5. The use of low retention pipette tips is essential to obtain appropriate accuracy and repeatability of mtDNA measurements in CSF. Low retention tips are designed to sustain minimal sample loss from viscous liquids such as CSF. Regular calibration of pipettes and correct pipetting techniques with slow aspiration/delivery rates is crucial to get low coefficients of variation for repeatability and reproducibility.
6. Wide bore tips with an orifice size wider than 1.5 mm are convenient to prevent mechanical damage to the droplets when transferring from the DG8 cartridge to the PCR plate.
7. This foil seal cannot be replaced by another material because it may damage the ddPCR droplet reader.
8. We find that it is best to prepare the master mix for all the planned reaction tubes in an excess volume of 10% and

distribute to each tube 17.05 µl of master mix. To this volume, we also add a 10% excess of CSF sample (4.95 µl). This is because in Subheading 3.2.6 of the protocol it is necessary to transfer without bubbles exactly 20 µl of this mixture to the cartridge well, and we found that if the initial volume is exactly 20 µl it is very difficult not to draw bubbles in the pipette during the transfer.

9. We find it is important that each ddPCR plate contains several nontemplate control assays (NTCs) that have master mix and nuclease-free dH<sub>2</sub>O instead of CSF sample. NTCs allow identification of possible contamination of analysis materials and reagents. It is essential to perform all procedures with sterile gloves using material specially reserved for PCR in laboratory areas only dedicated to the setup of PCR reactions.
10. Care should be taken to avoid contamination of equipment and samples with DNA standards that may contain significantly higher concentrations than the samples.
11. The reaction mixture contains a different concentration of FAM versus HEX labelled probe. The reason is that we found that there is a marked bleed-through cross-emission of FAM into HEX channel. One way to reduce such fluorophore cross-emission is by lowering the concentration of FAM labelled probe. The concentration ratio of FAM/HEX in our reaction conditions is optimized for our particular probes and if using other probes or quenchers we recommend finding again the optimal conditions.
12. It is important to add the oil after the reaction mixture because if added before, it will block the microfluidic channels.
13. We find that it is important to inspect the cartridge after droplet generation to ensure that droplets have been correctly formed. A simple way when using hydrolysis probes is to slowly rotate the cartridge in front of a light source and if droplets are formed there is a change in reflection of the light source every 60° rotation. If there are no droplets in one of the wells of the cartridge (either because of air bubbles or obstruction in the microfluidic channel) it is advisable to discard the whole cartridge. When using EvaGreen fluorescent dye it is more difficult to identify whether droplets have been correctly formed prior to the analysis.
14. The PCR temperature profiles described here are the optimal for the particular primer combinations of our reactions and our laboratory conditions. We recommend to follow the digital MIQE guidelines [11] and perform annealing temperature gradients to optimize the separation between positive and negative droplets for each primer combination and laboratory setting.

15. The total time to analyze a 96-well plate is approximately 7 h. This includes: 2.5 h for sample preparation and droplet generation; 2 h for thermocycling and 2.5 h for droplet reading.
16. We recommend setting the same threshold value for all the wells in the plate (Fig. 2). Some samples may show only one of the droplet partitions and the fluorescence histogram of those samples will have only one Gaussian curve, hampering the detection of a reliable threshold value to separate positive and negative partitions. Therefore, we find convenient to select all wells for analysis because then the histogram shows the amplitude of the total events in all samples, making unlikely the possibility that there is only one droplet population in all of them. We find that when the ddPCR assay is performed in the most optimal conditions, the threshold value does not significantly influence the result.
17. Only samples that exhibit a total of more than 12,000 droplets or events should be included in the final analysis. However, for most accurate measurements, setting the limit for sample inclusion to more than 15,000 droplets provides results with significantly lower coefficient of variation among replicates. Using this ddPCR method to measure mtDNA concentration in 4.5  $\mu$ l of CSF we find that three sample replicates is sufficient to obtain precise values with repeatability levels equivalent to the sampling error expected for volumes of less than 10  $\mu$ l.

## Acknowledgment

This work is supported by the Ministerio de Economía y Competitividad of Spain (Grants: SAF2011-23550, SAF2014-56644-R) and by the Instituto Carlos III (Grant: PI2013/08-3) from Centro de Investigación Biomédica en Red de Enfermedades Neurodegenerativas to R.T.

## References

1. Lleo A, Cavedo E, Parnetti L, Vanderstichele H, Herukka SK, Andreasen N, Ghidoni R, Lewczuk P, Jeromin A, Winblad B et al (2015) Cerebrospinal fluid biomarkers in trials for Alzheimer and Parkinson diseases. *Nat Rev Neurol* 11:41–55
2. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R et al (2011) The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 7:263–269
3. Jack CR Jr, Albert MS, Knopman DS, McKhann GM, Sperling RA, Carrillo MC, Thies B, Phelps CH (2011) Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 7:257–262
4. Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL, Blennow K, DeKosky ST, Gauthier S, Selkoe D, Bateman R et al (2014)

- Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol* 13:614–629
5. Blennow K, Zetterberg H (2015) Understanding biomarkers of neurodegeneration: Ultra-sensitive detection techniques pave the way for mechanistic understanding. *Nat Med* 21:217–219
6. Podlesniy P, Figueiro-Silva J, Llado A, Antonell A, Sanchez-Valle R, Alcolea D, Lleo A, Molinuevo JL, Serra N, Trullas R (2013) Low CSF concentration of mitochondrial DNA in preclinical Alzheimer's disease. *Ann Neurol* 10:655–668
7. Sykes PJ, Neoh SH, Brisco MJ, Hughes E, Condon J, Morley AA (1992) Quantitation of targets for PCR by use of limiting dilution. *BioTechniques* 13:444–449
8. Vogelstein B, Kinzler KW (1999) Digital PCR. *Proc Natl Acad Sci U S A* 96:9236–9241
9. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC et al (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem* 83:8604–8610
10. Pinheiro LB, Coleman VA, Hindson CM, Herrmann J, Hindson BJ, Bhat S, Emslie KR (2012) Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Anal Chem* 84:1003–1011
11. Huggett JF, Foy CA, Benes V, Emslie K, Garrison JA, Haynes R, Hellemans J, Kubista M, Mueller RD, Nolan T et al (2013) The digital MIQE guidelines: Minimum Information for Publication of Quantitative Digital PCR Experiments. *Clin Chem* 59:892–902



# Chapter 8

## Testing of General and Human-Associated Fecal Contamination in Waters

Yiping Cao, Meredith R. Raith, and John F. Griffith

### Abstract

qPCR has become increasingly popular for microbial water quality testing because it is faster, more specific, and more flexible than culture-based methods. However, qPCR method limitations such as quantification bias introduced by reliance on standards and susceptibility to PCR inhibitors are major obstacles for implementation in water testing. This is because water testing requires accurate quantification of rare targets and because environmental waters often contain PCR inhibitors. Digital PCR offers the opportunity to maintain qPCR's advantages over culture-based methods while ameliorating two of qPCR's major limitations: the necessity to run standard curves and high susceptibility to inhibition. Here we describe a complete method for simultaneous testing for a general microbial water quality indicator (*Enterococcus* spp.) and a human-associated fecal marker in environmental waters. The complete method includes water sampling and filtration to capture bacteria, DNA extraction from bacteria captured on the filter, and droplet digital PCR to quantify the genetic markers from bacteria indicative of general and human-associated fecal contamination.

**Key words** Fecal contamination, Water quality, *Enterococcus*, HF183, Microbial source tracking, Duplex digital PCR

---

### 1 Introduction

Fecal contamination in water is a public health concern worldwide. In most countries, fecal indicator bacteria (FIB) such as *Enterococcus*, *E. coli*, and total and fecal coliforms are routinely tested in surface water, treated wastewater, and source and finished drinking water for public health protection. However, while FIB offer a conservative indication of general fecal contamination, they do not provide information on host sources of fecal contamination because all warm-blooded animals shed FIB in their fecal material [1]. FIB are therefore limited in their usefulness for pollution remediation because effective remediation requires knowledge of the origin of the contamination, and for accurate risk assessment because human fecal contamination generally presents a much

higher public health risk than nonhuman contamination [2]. Additionally, FIB may also originate from nonfecal sources [3], furthering the need to track fecal contamination to a specific host. Genetic markers specific to fecal contamination from different hosts are now available [4]. Known as microbial source tracking markers, the most widely used human-associated marker is HF183, which can detect human fecal contamination with high diagnostic sensitivity and specificity [5].

While FIB in water have historically been measured by growth-based methods such as membrane filtration and substrate utilization, qPCR has become a popular alternative to these methods because qPCR provides faster sample-to-result time, higher specificity, and greater flexibility [1, 6]. In fact, the revised U.S. EPA recreational water quality criteria published in 2012 allows use of qPCR for *Enterococcus* spp. in routine recreational water monitoring [7]. For the HF183 marker specific to human fecal contamination, the only available measurement methods use (q)PCR [8]. In fact, qPCR is the dominant technology for measuring host-specific fecal contamination in the field of microbial source tracking because of the same advantages qPCR affords for measuring FIB [1].

Despite wide acceptance of qPCR in the medical and food safety fields, method limitations such as quantification bias introduced by reliance on standards and susceptibility to PCR inhibitors are still major obstacles for implementing qPCR in water testing [9–11]. It is also difficult to multiplex in qPCR, making it infeasible to measure both general and human-associated fecal contamination in one analysis [12]. However, digital PCR provides direct quantification without standard curves, is more resistance to inhibition, and much easier to multiplex [13]. A recently developed duplex digital PCR assay (EntHF183 duplex ddPCR) enabled unbiased, precise, sensitive, specific, highly inhibition resistant, and highly reproducible, simultaneous quantification of *Enterococcus* spp. and the HF183 marker. This is the first assay to simultaneously measure concentrations of a general microbial water quality indicator used for regulatory purposes and provide information regarding human-associated fecal contamination [12].

Here we describe the complete protocol for simultaneous testing of general and human-associated fecal contamination in water. The protocol consists of three major procedures: (1) water sampling and filtration to capture bacteria, (2) DNA extraction from bacteria captured on the filter, and (3) droplet digital PCR to quantify DNA representing the general microbial water quality indicator *Enterococcus* spp. and the HF183 human-associated fecal marker.

---

## 2 Materials

Carry out all procedures at room temperature unless otherwise specified. Aseptic technique must be used with all portions of the following procedures and gloves, a clean laboratory coat and other personal protective equipment should be worn for safety and to avoid contamination. Workspaces and equipment must be cleaned (using DNAAway or a 10% solution of chlorine bleach) prior to starting all laboratory procedures. The water sampling procedure described is specific to recreational water quality sampling at marine or freshwater beaches. Sampling of other water bodies may require different materials and follow different protocols. Unless bottles are new and certified to be nucleic acid free, it is recommended that sampling bottles be treated by adding a small volume (ca. 10 mL) of 10% HCl to conveniently destroy potential residual DNA/RNA and stored with the acid inside until use.

### 2.1 Water Sampling and Filtration

#### 2.1.1 Sample Bottle Preparation

1. 10% hydrochloric acid.
2. Deionized water.

#### 2.1.2 Water Sampling

1. New or acid treated polypropylene or polyethylene sample collection bottle.
2. Cooler with ice packs or ice.
3. Sampling field sheet.

#### 2.1.3 Water Filtration

1. Polycarbonate membrane (47 mm Isopore™ or Nuclepore™, 0.4 µm).
2. Disposable filtration funnel (such as Nalgene® Analytical Test Filter Funnels).
3. 2 mL screw-cap bead beating tube preloaded with glass beads (GeneRite S0205-50 or equivalent).
4. Sterile PBS rinsing buffer: pH 7.4, sodium dihydrogen phosphate at 4 mM, sodium monohydrogen phosphate at 21 mM, sodium chloride at 145 mM in 1 L of molecular grade water.
5. Anhydrous ethanol (for alcohol lamp).
6. Liquid nitrogen.

### 2.2 DNA Extraction

1. GeneRite Extraction kit (K200-02C-50).
2. Bead beater (Biospec 607 or equivalent).
3. Low binding 1.7 mL microcentrifuge tubes.

**Table 1**  
**Primer and probe sequences**

Target	Primer and probe sequences	Reference
<i>Enterococcus</i> spp.	EnteroF1A: GAGAAATTCCAAACGAACCTTG EnteroR1: CAGTGCTCTACCTCCATCATT GPLQ813TQ: [6-FAM]- TGGTTCTCTCCGAAATAGCTTAGGGCTA-[BHQ1]	[15]
Human fecal-associated Bacteroidales	HF183ND: ATCATGAGTTCACATGTCCG BthetaR1: CGTAGGAGTTGGACCGTGT BthetaP1: [HEX]-CTGAGAGGAAGGTCCCCACATTGGA-[BHQ1]	[5]

### 2.3 Droplet Digital™ PCR

#### 2.3.1 Making Assay Master Mixes

1. *Enterococcus* and HF183 primers and probe (Table 1).
2. TE pH 8 buffer.
3. ddPCR™ Supermix for Probes (no dUTP) (Bio-Rad, 186-3024).
4. Nuclease-free PCR grade water.
5. Low binding 1.7 mL microcentrifuge tubes.
6. Aluminum Sealing Film.
7. Hard-shell 96-well Plate.

#### 2.3.2 Making Droplets

1. Droplet Generation Oil for Probes (Bio-Rad, 186-3005).
2. DG8™ Cartridges for Droplet Generator (Bio-Rad, 186-4008).
3. DG8 Gaskets for Droplet Generator (Bio-Rad, 186-3009).
4. DG8 Cartridge Holders (Bio-Rad, 186-3051).
5. Droplet Generator (Bio-Rad).
6. Rainin 20 µL multichannel pipet/tips (Rainin 20 µL XLT and GPL20F).
7. Rainin 50 µL multichannel pipet/tips (Rainin 50 µL XLT and GPL200F).
8. Rainin 200 µL multichannel pipet/tips (Rainin 200 µL XLT and GPL200F).

#### 2.3.3 ddPCR Thermal Cycling, Detection, and Data Analysis

1. Hard-shell 96-well PCR plate (Such as the Eppendorf Twin Tec PCR plate).
2. Pierceable Heat Seal Foil (Bio-Rad).
3. PX1™ PCR Plate Sealer (Bio-Rad).
4. CFX96™ Thermalcycler (Bio-Rad).
5. QX100™ Droplet Reader (Bio-Rad).
6. Droplet Reader Oil (Bio-Rad, 186-3004).
7. QuantaSoft™ Software.

---

### 3 Methods

#### 3.1 Water Sampling and Filtration

##### 3.1.1 Preparing Sampling Bottles

1. Rinse the bottles three times with deionized water (DI) in order to remove any large particles.
2. Pour all the water out and add about 10 mL of 10% hydrochloric acid (10% HCl). Securely cap the bottle and swirl acid to coat inner bottle surface. Store bottle capped with acid inside until use.

##### 3.1.2 Beach Water Sampling

1. Dispose of acid in sampling bottle by making a small divot in the sand with your toe and pouring acid into divot. Avoid getting acid on your skin or clothes.
2. Fill the sampling bottle  $<\frac{1}{4}$  full with sample water. Close bottle, shake vigorously, and discard rinse water away from collection area. Repeat this process twice more.
3. Collect full volume water sample at ankle depth on an incoming wave.
4. Place water sample in cooler, transport sample on ice pack/ice to lab for filtration within 6 h.

##### 3.1.3 Water Filtration

1. Setup vacuum filtration system (*see Note 1*). Light the alcohol lamp or Bunsen burner (for flame sterilization). Soak the tips of two pairs of stainless steel forceps in a beaker with 95% ethanol.
2. Prepare the disposable filtration funnel for water filtration by replacing the manufacturer's preloaded filter (generally not the one specified in this method) with the specified polycarbonate membrane (*see Note 2*).
3. Filter 100 mL water sample (*see Note 3*), rinse funnel wall with approximately 10 mL of PBS so all materials are captured on the membrane filter. After all liquid passes through the membrane filter, close vacuum valve and remove the funnel.
4. Using flame sterilized (*and cooled*) forceps (*see Note 4*), carefully roll up the polycarbonate membrane filter on the backing plate. First, fold one edge of the membrane onto itself (about  $\frac{1}{4}$  of the diameter of the membrane) and hold in place with second forceps. Then, using each pair of forceps alternately, roll the membrane to create a tube. Place into the corresponding collection tube (i.e., 2 mL screw-cap microtube suitable for  $-80^{\circ}\text{C}$  storage and the bead beating step during DNA extraction), screw the cap securely (*see Note 5*).
5. Flash-freeze the tube containing the filter in liquid nitrogen and transfer to  $-80^{\circ}\text{C}$  freezer for storage until DNA extraction (*see Note 6*).

### 3.2 DNA Extraction

#### 3.2.1 Bead Beating

This protocol employs the published bead beating DNA extraction procedure using the GeneRite DNA EZ kit (commonly employed for molecular testing of recreational waters [1, 9, 11]) (*see Note 7*).

1. Remove filter tubes to be analyzed from  $-80^{\circ}\text{C}$  freezer and pipet 600  $\mu\text{L}$  lysis buffer into each. Bead beat the tubes for 1 min (*see Note 8*).
2. Centrifuge the bead beaten tubes for 3 min at  $12,000 \times g$ . Pull out all liquid from tubes and place into a new 1.7 mL low binding microtube. Centrifuge the supernatants in the microtubes for 1 min at  $12,000 \times g$ .

#### 3.2.2 Binding DNA to Column

1. Pull out 380  $\mu\text{L}$  of supernatant from the centrifuged tube and place liquid into a new 1.7 mL low binding tube containing 760  $\mu\text{L}$  of binding buffer. Vortex and spin the mixture briefly.
2. Add  $\sim$ 650  $\mu\text{L}$  of mixture to the provided column (the column is already sitting in the provided collection tube) and spin for 1 min at  $12,000 \times g$ . Discard the collection tube containing the flow through and place the column in a new collection tube. Repeat **step 4** until all of the mixture from **step 3** is put through the column.

#### 3.2.3 Column Washing

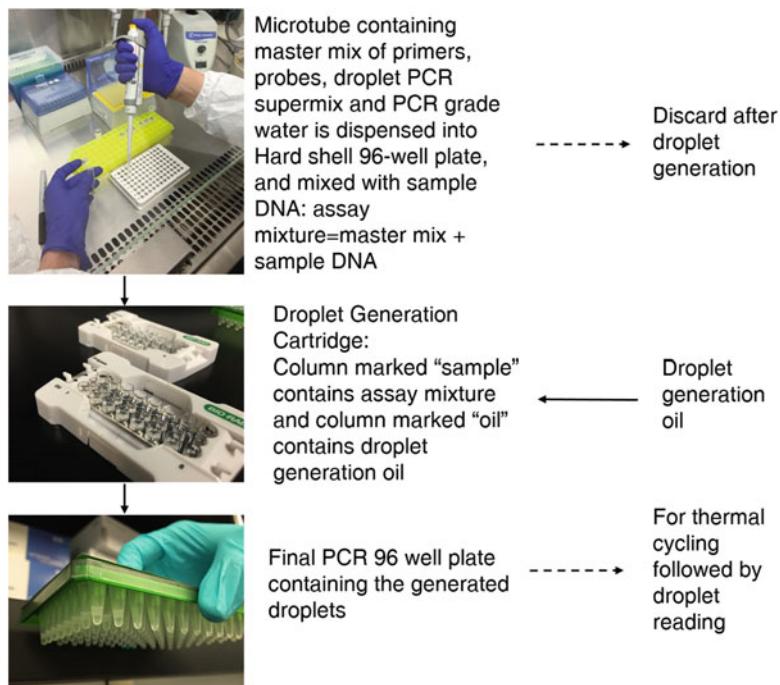
1. After placing the column in a new collection tube add 500  $\mu\text{L}$  of wash buffer to the column and spin for 1 min at  $12,000 \times g$ . Discard the collection tube containing the flow through and place column in new collection tube. Repeat this step.

#### 3.2.4 Column Elution

1. Place column in new 1.7 mL low binding microtube. Add 50  $\mu\text{L}$  elution buffer to the column and let sit 1 min. Spin the column for 1 min at  $12,000 \times g$ . Repeat this step.
2. After 100  $\mu\text{L}$  of DNA extract has been collected, dispose of the column and aliquot the extract if needed.
3. Quantify total DNA in each sample by spectrophotometer (*see Note 9*)

### 3.3 Droplet Digital PCR

Two hard-shell 96-well plates are used in this procedure: one for preparing the assay mixture and the other as the final PCR plate. The former can be any plate that is easy to work with, while the latter should be a PCR plate with thermal conductivity, plate height compatible with the PCR thermal cycling parameters, the thermal cycler, and the Droplet Reader. However, currently, it is recommended to use an Eppendorf Twin Tec 96-well plate (Bio-Rad, personal communication). Figure 1 depicts the transfer of liquid (reagents, sampled DNA, oil, and droplets) in this section. The steps described in this section are also demonstrated in a video article [14].



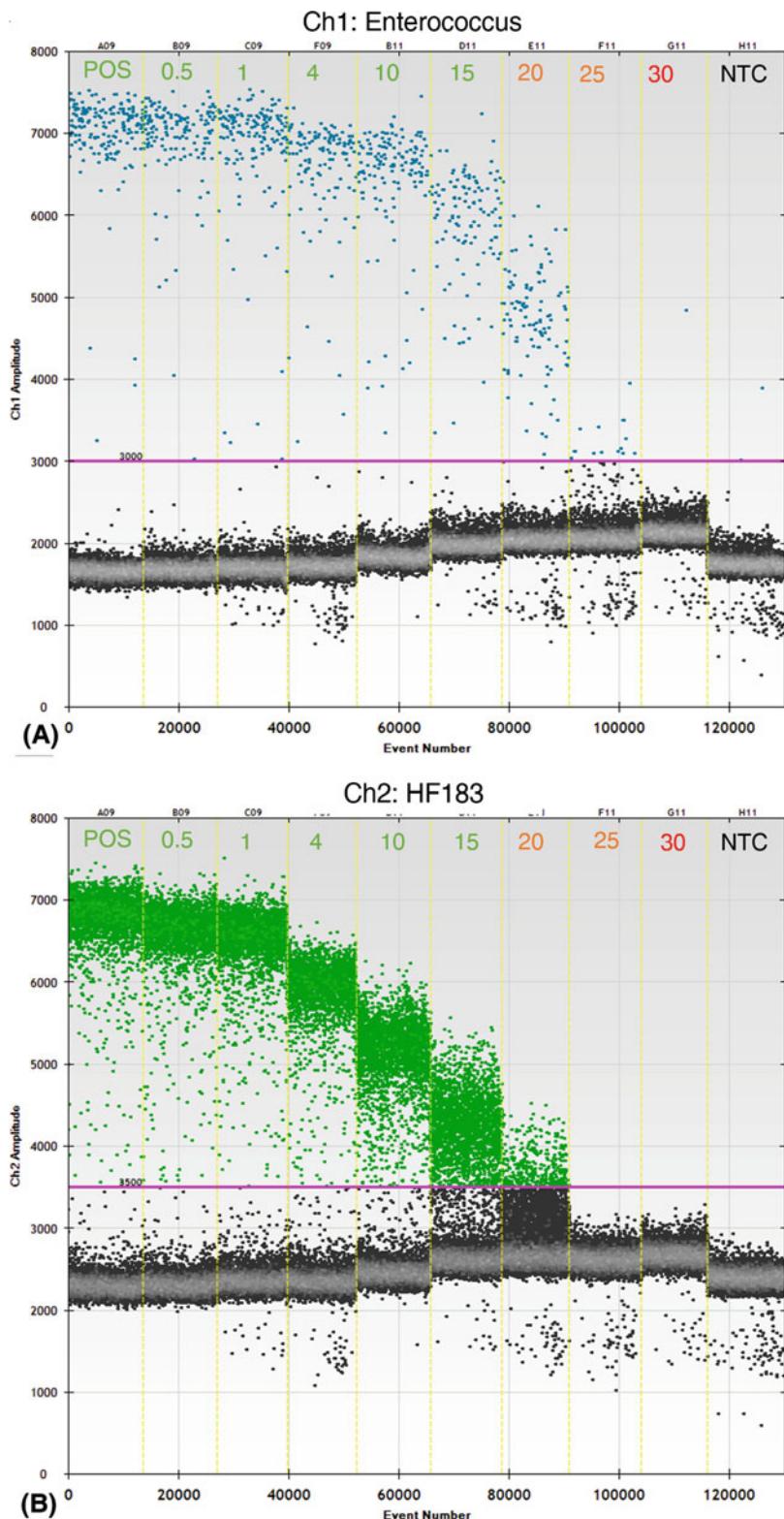
**Fig. 1** Liquid transfer flow diagram for Subheading 3.3

1. Make 100  $\mu\text{mol/L}$  stock concentrations for all *Enterococcus* and HF183 primers and probes. Use molecular grade water to hydrate and dilute primers and TE Buffer pH 8 to hydrate probes (see Table 1).
2. Use the primers and probes to prepare the master mix by adding 12  $\mu\text{L}$  Droplet PCR Supermix ( $2\times$  stock), 0.216  $\mu\text{L}$  each forward and reverse primer, 0.06  $\mu\text{L}$  each probe, and 5.016  $\mu\text{L}$  nuclease free PCR grade water, per reaction well. Prepare master mix in bulk for all reactions. Ensure that the master mix is well vortexed and spun down before pipetting into the hard-shell 96-well plate (no specific requirement on plate height, thermal conductivity, or supplier).
3. Pipet 18  $\mu\text{L}$  master mix into each well on the 96-well reaction mix plate and add 6  $\mu\text{L}$  DNA template. If running samples in duplicate, pipet 36  $\mu\text{L}$  of master mix per well and 12  $\mu\text{L}$  DNA template into each well and leave the corresponding replicate wells (e.g., in the adjacent plate column) empty on the plate (see Note 10).
4. Pipet the assay mixture up and down with the 20  $\mu\text{L}$  multi-channel pipet a minimum of 12 times to ensure adequate mixing. Avoid making excess bubbles within the mixture (see Note 11).

5. Place a cartridge in the cartridge holder, click shut, and pipet 20  $\mu$ L of the assay mixture using the multichannel pipet into the section of the cartridge marked “Sample” (*see Note 12*). Use the 200  $\mu$ L pipet to transfer 70  $\mu$ L of droplet generation oil to the section of the cartridge marked “Oil.” Cover cartridge with a gasket and gently place on droplet generator.
6. While droplet generation is occurring prepare the next cartridge in the same manner as described in **step 5** to save total setup time (*see Note 13*).
7. When the droplet generator is done, take off the first cartridge, set aside, and gently place the second cartridge onto the droplet generator.
8. Remove the gasket from the first cartridge and discard. Do not unclick the cartridge from the cartridge holder as this may break the newly generated droplets. Gently transfer the total volume of generated droplets into the final PCR plate for thermal cycling (*see Note 14*). Repeat droplet generation with the remaining samples.
9. When all the droplets are in the final PCR plate, place plate on plate sealer with pierceable foil on top and heat seal at 180 °C for 10 s.
10. Remove the plate from sealer and place on a thermal cycler compatible with the final PCR plate and a temperature ramping speed of 2.0 °C/s. Run thermal program as follows: 10 min at 95 °C, followed by 40 cycles of 30 s at 94 °C and 60 s at 60 °C, followed by a 10 min hold at 98 °C (optional: final hold at 4 °C, or 12 °C if left overnight).
11. After cycling, transfer the plate to the Droplet Reader (*see Note 15*). The PCR plate may also be stored at 4 °C for up to 3 days before droplet reading.
12. Once the plate is on the droplet reader, configure the software to read RED (rare event detection) for each well. Define all sample names and select Channel 1 (FAM) and Channel 2 (HEX or VIC) for data collection. Select Run and choose FAM-HEX (or FAM-VIC) to start the run (*see Note 16*).

### 3.4 Data Analysis

1. When the run is complete, open the data file and check the events number (i.e., number of acceptable droplets in each well). All wells containing less than 10,000 droplets should be excluded.
2. The fluorescence threshold should then be set approximately one standard deviation (500–700 fluorescence units, Fig. 2) above the negative droplets in NTC wells (*see Note 17*). Fluorescence values of unknown samples should also be compared



to that of positive controls to check for signs of PCR inhibition (*see Note 18*).

3. Results may be exported in.csv file for further data analysis (*see Note 19*).

---

## 4 Notes

1. Test whether the vacuum assembly is working properly before moving forward. Turn on the pump and close all valves. Then adjust the vacuum pressure using the knurled knob on the vacuum inlet of the pump ( $<20 \text{ in.Hg}$ , or  $<0.6 \text{ atm}$ ). Ensure that suction is carried from the pump to the filtration manifold.
2. If the commercially available prepacked disposal filtration funnels are preloaded with a filter other than the polycarbonate membrane filter, the following steps can be followed to aseptically replace the preloaded filter:
  - (a) Mount the filtration funnel (i.e., housing) on to the adaptor. Remove the funnel from the base and place it upside down on the bench on top of its lid.
  - (b) Take one pair of forceps that have been soaking in 100% ethanol, flame to burn off the ethanol, allow forceps to cool, then pick up the preloaded filter from the filtration housing being careful not to damage the filter support underneath the grid filter. Discard the gridded filter.
  - (c) Use the same forceps to carefully pick up one isopore polycarbonate membrane (*clear grayish membrane in-between paper separators, do not mistake the paper separator for the membrane filter*), and carefully place it onto the center of the filter support in place of the gridded filter.
  - (d) Snap the filtration funnel back onto the housing to secure the polycarbonate membrane (*make sure that there is no gap between the edge of polycarbonate membrane and the bottom of the filtration funnel and that the edge of the membrane is not folded, i.e., no liquid should go through the housing/support without passing through the polycarbonate membrane first*).

---

**Fig. 2** Example 1-D plots for the two targets (A: *Enterococcus*, B: HF183) of the duplex ddPCR assays. Positive control, no-template controls, and samples with no, low, and high levels of inhibitors are labeled on each well on top of each plot as POS, NTC, and numeric numbers, respectively. Numeric numbers denote humic acid concentration (spiked into the same sample) in ng per  $\mu\text{L}$  reaction. Green, orange, red, and black colors of the label indicate no effect on quantification, underestimation, failure to detect, and NTC, respectively

3. Depending on turbidity of the water, smaller volumes may need to be added incrementally (e.g., 20 mL) to the funnel until the filter clogs or 100 mL successfully filters through the membrane. The filtered volume must be recorded for data analysis downstream.
4. Forceps must be used to move the filter to the final microtube. The forceps must be flame sterilized but cannot be used right away. If the forceps are too hot then portions of the sample and/or filter will be destroyed. After flame sterilization, let forceps cool for ~15 s.
5. The membrane can be folded twice before placing into the 2 mL screw-cap microtube. Keep in mind that the material collected on the filter should not be touched or damaged during folding as this compromises the integrity of the sample. Other tubes compatible with -80 °C storage and downstream DNA extraction protocol can also be used.
6. If liquid nitrogen is not available, alternatives (such as ethanol–dry ice bath) suitable to flash-freezing a sample can also be used. It is important to flash freeze samples to minimize enzymatic activity causing sample degradation on the membrane filter.
7. DNA extraction using other procedures/kits can be used but should be evaluated for recovery, purity, and the type of elution buffer used. Recovery and purity of the DNA are necessary to ensure the integrity of the sample. The elution buffer should not contain chemicals that might interfere with downstream analysis (e.g., ddPCR), either with enzyme function nor with droplet formation and stability.
8. Bead beating is necessary to break open cells and free DNA. However, too much bead beating may lead to DNA fragmentation and sample degradation, which will alter the concentrations of measurable target DNA downstream. One minute of bead beating is sufficient for environmental samples using the filters described in this protocol.
9. Total DNA concentration should be measured to ensure that the ddPCR reaction is not overloaded. If DNA concentration is greater than 13 ng/µL (i.e., 66 ng per ddPCR reaction), a dilution should be considered when prepping ddPCR.
10. When setting up ddPCR reaction mix plates, it is convenient to set up samples for each DG8 cartridge vertically in a plate column.. All wells in a cartridge, and thus in a column if using an eight-channel pipettor to load the samples, must be filled in order for successful droplet generation. If all eight unique samples on a chip are run with two technical replicates, then both chips can be loaded from a single plate column containing a double volume ( $2 \times 24 \mu\text{L}$ ) of reaction mix.

11. New pipet tips must be used for each column when mixing the samples. During mixing, pipet less than 20 µL and keep the tips in the liquid, as this avoids the formation of bubbles in the sample. Ensure that all the sample mixture is pushed out of the pipet tip before disposing.
12. When transferring the sample assay mixture from the hard-shell 96-well plate to the cartridge, ensure that all 20 µL of volume has been taken up in each of the tips. Then placing the pipet tip into the cartridge, hold the pipet perpendicular to the bench top and ensure that each pipet tip touches the side near but not touching the bottom of the cartridge. Once touching, slowly depress the plunger to release ~20% of the volume. Continue releasing volume slowly but work the pipet tip up the side of the cartridge walls. This action ensures minimal bubbles are introduced into cartridge. If excess bubbles are present droplet generation will end prematurely.
13. The second cartridge may be prepped when the first cartridge has been placed on the droplet generator. If the second cartridge corresponds to the replicate column of the first cartridge then sample mixture is taken from the same column in the hard-shell 96-well plate as when prepping cartridge one. Remember that less volume is now present in the once-used reaction mix column and pipetting must be done slowly to ensure that a full 20 µL is put onto the second cartridge.
14. Transferring the droplets from the cartridge to the final PCR 96-well plate (such as the Eppendorf Twin Tec PCR plate) must be done slowly and precisely. Using the 50 µL multichannel pipet place the tips into the droplets at a 45° angle. The tips should not touch any surface but the droplets. Slowly pipet the droplets into the tips all the while following the liquid level as it decreases. At this point, the tips will touch the hard surface of the cartridge, continue to slowly pipet the droplets. Once the maximum amount of volume has been pipetted place the tips containing the droplets to the side of the final PCR plate at a 45° angle, close to the bottom of the well. Slowly depress the pipet while transferring the droplets to the well of the droplet PCR plate. Check that all the droplets have been recovered from the cartridge. If not, pipet the remaining droplets out and place them into the final plate in the same manner as just described. Note that pipetting too quickly at this step will shear the droplets and reduce the accuracy of the final concentration.
15. The plate can only be placed on the droplet reader with well A1 in the top left corner. If the plate is placed rotated 180°, the components will not sit correctly and reading cannot commence. When the plate is placed on the reader, it is recommended that the samples are at room temperature.

16. The QX100 droplet reader reads two fluorophores. FAM is read on channel 1 and HEX or VIC are read on channel 2. In this protocol FAM is used for the *Enterococcus* target and HEX is used for human HF183 target.
17. A single manual fluorescence threshold applied to all samples is recommended for the following reasons. First, quantification of target concentrations have been shown to be relatively insensitive to threshold position. Second, a single threshold facilitates method implementation among water testing practitioners.
18. This duplex ddPCR assay is much more resistant to PCR inhibition than its qPCR counterparts [12]. The ddPCR continued to provide quantification for both *Enterococcus* and HF183 even at humic acid (HA) concentration of 15 ng per  $\mu\text{L}$  reaction. Nevertheless, extremely high inhibitor concentration did lead to underestimation (20–25 ng HA per  $\mu\text{L}$  reaction) or false negatives (30 ng HA per  $\mu\text{L}$  reaction) (Fig. 2). Typically, a lower fluorescence value for the positive droplets is observed before underestimation occurs (Fig. 2).
19. Once the data is exported, multiply the exported target concentration by 4 to convert it from copies of target (23S gene of *Enterococcus* spp. or the HF183 marker) per  $\mu\text{L}$  reaction to copies of target per  $\mu\text{L}$  DNA template. Additional calculations may be needed in order to account for filtration volume and to compute the target copies per unit volume of sample water.

## References

1. Boehm AB, Van De Werfhorst LC, Griffith JF et al (2013) Performance of forty-one microbial source tracking methods: a twenty-seven lab evaluation study. *Water Res* 47:6812–6828
2. Soller JA, Schoen ME, Bartrand T et al (2010) Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. *Water Res* 44:4674–4691
3. Whitman RL, Nevers MB (2003) Foreshore sand as a source of *Escherichia coli* in nearshore water of a lake Michigan beach. *Appl Environ Microbiol* 69:5555–5562
4. Field KG, Samadpour M (2007) Fecal source tracking, the indicator paradigm, and managing water quality. *Water Res* 41:3517–3538
5. Layton BA, Cao Y, Ebentier DL et al (2013) Performance of human fecal anaerobe-associated PCR-based assays in a multi-laboratory method evaluation study. *Water Res* 47:6897–6908
6. Griffith JF, Weisberg SB (2011) Challenges in implementing new technology for beach water quality monitoring: lessons from a California demonstration project. *Mar Technol Soc J* 45:65–73
7. U.S. EPA (2012) Recreational water quality criteria. 820-F-12-058. Office of Water, Washington, DC
8. Bernhard AE, Field KG (2000) A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. *Appl Environ Microbiol* 66:4571–4574
9. Cao Y, Griffith JF, Dorevitch S et al (2012) Effectiveness of qPCR permutations, internal controls and dilution as means for minimizing the impact of inhibition while measuring *Enterococcus* in environmental waters. *J Appl Microbiol* 113:66–75
10. Cao Y, Sivaganesan M, Kinzelman J et al (2013) Effect of platform, reference material, and quantification model on enumeration of

- Enterococcus by quantitative PCR methods. Water Res 47:233–241
11. Shanks OC, Sivaganesan M, Peed L et al (2012) Inter-laboratory general fecal indicator quantitative real-time PCR methods comparison study. Environ Sci Technol 46:945–953
  12. Cao Y, Raith MR, Griffith JF (2015) Droplet digital PCR for simultaneous quantification of general and human-associated fecal indicators for water quality assessment. Water Res 70:337–349
  13. Hindson BJ, Ness KD, Masquelier DA et al (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Anal Chem 83:8604–8610
  14. Cao Y, Raith MR, Griffith JF (2016) A duplex digital PCR assay for simultaneous quantification of the *Enterococcus* spp. and the human fecal-associated HF183 marker in waters. J Vis Exp:e53611. <https://doi.org/10.3791/53611>
  15. U.S. EPA (2012) Method 1611: enterococci in water by TaqMan® quantitative polymerase chain reaction (qPCR) assay. EPA-821-R-12-008. Office of Water, Washington, DC

# **Part III**

## **Copy Number Variation**



# Chapter 9

## Analyzing Copy Number Variation with Droplet Digital PCR

Avery Davis Bell, Christina L. Usher, and Steven A. McCarroll

### Abstract

Many genomic segments vary in copy number among individuals of the same species, or between cancer and normal cells within the same person. Correctly measuring this copy number variation is critical for studying its genetic properties, its distribution in populations and its relationship to phenotypes. Droplet digital PCR (ddPCR) enables accurate measurement of copy number by partitioning a PCR reaction into thousands of nanoliter-scale droplets, so that a genomic sequence of interest—whose presence or absence in a droplet is determined by end-point fluorescence—can be digitally counted. Here, we describe how we analyze copy number variants using ddPCR and review the design of effective assays, the performance of ddPCR with those assays, the optimization of reactions, and the interpretation of data.

**Key words** Copy number variants, Genomic structural variation, Droplet digital PCR, Digital PCR, Genotyping, Genotyping assay design

---

### 1 Introduction

Even within a single species, such as humans, thousands of genomic segments vary in copy number from individual to individual. In the context of cancer and other proliferative disorders, substantial parts of the genome can also differ in copy number between disease and healthy cells from the same person. Precisely measuring such differences is key to ascertaining their biological import.

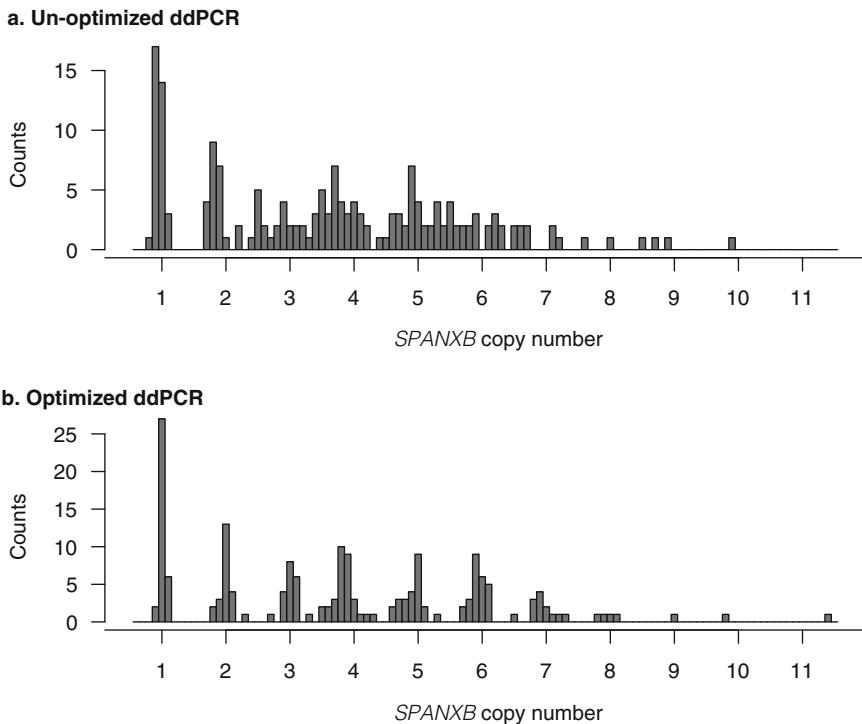
Though precise and accurate measurement is critical in all research, many research contexts present particular challenges for accurate copy number determination. In cancer cells, many oncogenes become amplified to high copy numbers. Many inherited copy number variants (CNVs) are also present in a wide range of copy numbers (e.g., from two to ten) within different individuals' diploid genomes, due to multiallelism. In humans, such CNVs appear to generate most inherited gene-dosage variation and make a substantial contribution to gene-expression variation [1],

suggesting that they may contribute to variation in phenotypes. To understand how copy number variation contributes to phenotypes, how alleles are distributed within and across populations, and how CNVs relate to SNPs and haplotypes, it is crucial to accurately measure (or “genotype”) such CNVs.

Droplet digital PCR (ddPCR) obtains precise and accurate measurements of copy number by partitioning the reagents for two fluorescence assays (one detecting the CNV of interest and one detecting a control reference locus of known copy number) into thousands of droplets—creating thousands of individual reactions—and determining whether each droplet contained either DNA molecule by measuring the fluorescence of each droplet after PCR [2, 3]. Copy number is calculated by comparing the number of molecules arising from the CNV segment of interest (calculated from the number of positive droplets) to the number of molecules arising from the reference genomic locus. Because the fluorescence measurement is taken after (rather than during) PCR, its accuracy relies only on distinguishing the fluorescence-positive from the fluorescence-negative droplets, not on the quantitative PCR kinetics that classical real-time PCR attempts to measure. This yields a powerful improvement in the precision of analysis, allowing a precise determination of integer copy number at loci where rtPCR has been unable to do so [2–4]. For example, at the highly copy-number variable sperm gene *SPANXB* [5], studies using qPCR have only estimated the copy numbers that are present in each genome [6, 7], whereas ddPCR can measure the precise, integer level in each individual’s genome (Fig. 1b).

Here, we share a detailed protocol for analyzing copy number variation with ddPCR including (1) designing successful assays targeting genomic segments of interest, (2) using those assays in ddPCR and optimizing reaction conditions, and (3) improving the ddPCR analysis results after data generation. We pay particular attention to assay design and optimization, which can greatly affect data quality (Fig. 1). We have used this method for deep interrogation of particular genomic regions, including characterization and phenotype association analyses [4, 8], as well as for confirmation, validation, and population-based analysis of copy number variants [1].

While this protocol includes many details that are most helpful for typing germline copy number variants in stable euploid genomes, the protocol is readily adapted for analyzing copy-number-variable segments in cancer genomes. A key difference is that because cancer samples are often mosaic (a mixture of clones with different genomes), analysis results for cancer samples may involve noninteger copy-number levels that represent an average across the cells in a sample. Another useful application of ddPCR involves quantifying the copy number of transgenes.



**Fig. 1** ddPCR-generated copy numbers for 179 individuals at the *SPANXB* locus before (a) and after (b) the assay and reaction optimization techniques outlined in this protocol. The optimized copy numbers were generated by combining data from replicates run with two separate X chromosome-located replication-timing matched control assays (see Notes 26–28)

## 2 Materials

Prepare all solutions with ultrapure, molecular biology-grade water. Protect all solutions containing fluorescently labeled probes from light. Mix all reagents by briefly vortexing and centrifuging them before use.

### 2.1 Locus-Specific Reagents

1. Assay targeting CNV region of interest (20× target mix): 18 μM forward primer, 18 μM reverse primer, and 5 μM 5' FAM-labeled, 3' ZEN or Black Hole-quenched probe designed to genomic region of interest. To make, combine 25.2 μL of 100 μM forward primer, 25.2 μL of 100 μM reverse primer, and 7 μL of 100 μM probe with 82.6 μL water. Store at –20 °C (see Note 1).
2. Assay targeting control region (20× control mix): 18 μM forward primer, 18 μM reverse primer, and 5 μM 5' HEX-labeled, 3' ZEN or Black Hole-quenched probe designed to non-copy number variable genomic region (see Notes 2 and 3). To make, combine 25.2 μL of 100 μM forward primer, 25.2 μL of 100 μM reverse primer, and 7 μL of 100 μM probe with 82.6 μL water. Store at –20 °C (see Note 1).

## **2.2 ddPCR Components and Equipment**

1. Genomic DNA at a concentration of 5 ng/µL or higher, totaling at least 50 ng (*see Note 4*).
2. Restriction enzyme and associated buffer for digesting genomic DNA, potentially AluI with 10× CutSmart® buffer (New England Biolabs) (*see Note 5*).
3. 2× ddPCR™ Supermix for Probes, with or without dUTP (Bio-Rad).
4. DG8™ Cartridges for droplet generation (Bio-Rad).
5. DG8™ Gaskets for droplet generation (Bio-Rad).
6. Droplet Generation Oil for Probes (Bio-Rad).
7. Droplet Reader Oil (Bio-Rad).
8. QX200™ Droplet Digital PCR System: droplet generator and cartridge holders, droplet reader, and QuantaSoft reader software (Bio-Rad).
9. Rainin multichannel pipettors and corresponding tips for pipetting 20 µL and 40 µL volumes (*see Note 6*).
10. Half-skirted Eppendorf twin.tec 96-well plates for droplet thermal cycling and reading.
11. Pierceable, heat-sealable foil seals (Bio-Rad Pierceable Foil Heat Seal).
12. Plate sealer capable of sealing for 5 s at 180 °C (e.g., Bio-Rad PX1™ Plate Sealer).
13. Thermal cycler.

## **2.3 Web Resources**

1. UCSC genome browser (hg19): <http://genome.ucsc.edu/cgi-bin/hgGateway>.
2. Primer3 primer design tool: <http://bioinfo.ut.ee/primer3/> [9, 10].
3. SNP masking tool: <http://bioinfo.ut.ee/snpmasker/> [11].
4. NEB cutter: <http://nc2.neb.com/NEBCutter2/> [12].
5. IDT oligoanalyzer: <http://www.idtdna.com/calc/analyzer>.
6. Multiple primer heterodimer analyzer: <http://www.thermoscientificbio.com/webtools/multipleprimer/>.

## **3 Methods**

### **3.1 Assay Design**

1. The first step of assay design is to determine the best region for assay placement in order to optimize detection of the genomic segment of interest and ddPCR performance. To begin, obtain the DNA sequence for the copy-number-variable region of interest by entering the coordinates spanning the region into the UCSC genome browser. Determine whether the region is

present once or more than once in the reference genome by displaying segmental duplications. Select “dense” from the “Segmental Dups” pull-down menu under the “Repeats” section at the bottom of the page (*see Note 7*). Many copy number variants are found more than once in the reference genome, raising special considerations; if this is the case for the region of interest, *see Note 8*.

2. When the specific region of interest is identified and displayed in the genome browser, set the “RepeatMasker” track (under “Repeats”) pull-down menu to “dense” and reload the page. Get the sequence for the visualized region by selecting “DNA” under the “View” menu at the top of the page. In order to prevent the assay from being designed to target repeat regions, check the box next to “Mask repeats” and select “to N,” then click the “get DNA” button (*see Note 9*).
3. Design the primers and probe to assay this region using the Primer3 primer design tool. Enter the DNA sequence obtained in **step 1** into the box at the top of the webpage. Check “Pick hybridization probe (internal oligo)” under the input sequence.
  - (a) From the “Mispriming library (repeat library)” pull-down menu above the sequence box, choose “HUMAN.”
  - (b) Under “General Primer Picking Conditions,” set the optimal primer length (“Primer size”) to 22 bp, “Primer Tm” Min to 59, Opt to 60, and Max to 61. Set the product size range to 60–90 bp (which can be relaxed to 60–150 bp if no assays are found) (*see Note 10*).
  - (c) Under “Internal Oligo (Hyb Oligo) General Conditions,” set “Internal Oligo Tm” Min to 68, Opt to 69, and Max to 70, and choose “HUMAN” from the “Internal Oligo Mishyb Library” pull-down menu.
- Leave the remaining options unchanged and click “Pick primers.” The temperatures can be adjusted if no suitable assays are found, as long as the internal oligo (probe) melting temperature is still higher than the primer temperature.
4. Choose an assay from the results of **step 3** that contains any necessary sequences and is likely to perform well. Avoid probe sequences that start with G (*see Note 11*). Use the UCSC BLAT tool (under the “Tools” menu at the top of the page) to check that the forward and reverse primers match the region of interest uniquely and perfectly (*see Note 12*). View the region with the “Common SNPs” track displayed to ensure the primers and probe do not bind over a SNP. In addition, check whether the primers or probes are likely to bind each other or the control assay by using the “Hetero-Dimer” option in the right menu bar of IDT’s oligoanalyzer; delta-Gs lower

than  $-7$  should be avoided because their heterodimerization may interfere with the PCR (*see Note 13*).

5. Ensure that the amplicon generated by the primers does not contain a cut site for the restriction enzyme that will be used to digest the DNA prior to ddPCR. Obtain the amplicon sequence from UCSC genome browser. Do not mask repeats this time. Copy and paste this sequence into the NEB cutter webpage. Select “All commercially available specificities” to the right of “Enzymes to use,” then click the “Submit” button to the right of the box containing the DNA sequence. Under the resulting graphic in the “List” box, click “0 Cutters” and make sure the enzyme of interest is included (*see Note 14*).
6. Order the primers and probe from your usual oligo supplier. We prefer the FAM and HEX probe fluorophores, both with the ZEN quencher (Integrated DNA Technologies), though other combinations of fluorophores, quenchers, and suppliers also perform well. When making or ordering the  $20\times$  assay mix, please note that the proportion of primers to probes is different than in qPCR.

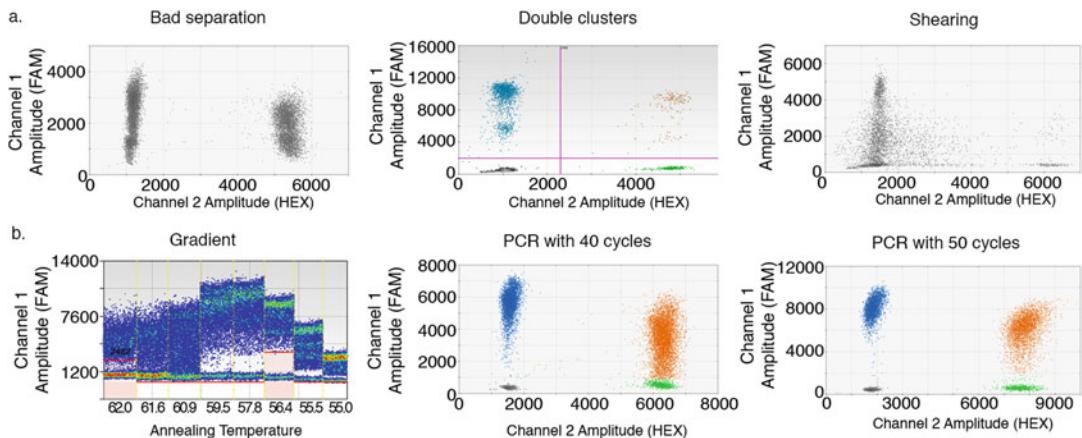
### **3.2 ddPCR for Copy Number Determination**

1. Digest the genomic DNA with a restriction enzyme to separate the copies of the CNV. For each sample, make an enzyme master mix consisting of 0.2 units/ $\mu\text{L}$  AluI and  $2\times$  CutSmart buffer (New England Biolabs). Add 10  $\mu\text{L}$  of this master mix to 50 ng DNA in 10  $\mu\text{L}$ , for a total reaction volume of 20  $\mu\text{L}$ . Mix by pipetting up and down. Do not vortex the enzyme or enzyme solution.
2. Incubate the enzyme-DNA mixture at  $37^\circ\text{C}$  for 1 h.
3. Dilute the digested DNA twofold by adding 20  $\mu\text{L}$  of water to each sample, yielding a DNA concentration of 1.25 ng/ $\mu\text{L}$ . Keep digested DNA at  $4^\circ\text{C}$  or on ice for immediate use, or at  $-20^\circ\text{C}$  for long-term storage. (*See Note 15* for an alternative restriction digestion strategy.)
4. For each sample add:
  - (a) 12.5  $\mu\text{L}$  of  $2\times$  ddPCR Supermix for Probes (Bio-Rad).
  - (b) 1.25  $\mu\text{L}$  of  $20\times$  assay targeting the CNV region.
  - (c) 1.25  $\mu\text{L}$  of  $20\times$  assay targeting control region (these first three reagents can be combined to form a master mix.).
  - (d) 10.0  $\mu\text{L}$  of the digested, diluted DNA (*see Notes 16 and 17*).
5. Mix well by pipetting up and down ten times. Proper mixing is critical. Spin the plate to collect the liquid at the bottom of wells. Keep the plate protected from light until droplet generation, and allow the reactions to equilibrate to room temperature for 3 min prior to droplet generation.

6. Place a DG8™ cartridge into the QX200 droplet generation cartridge holder and snap the holder closed. Pour Droplet Generation Oil for Probes into a reservoir for ease of multi-channel pipetting.
  - (a) Pipette 20 µL of the PCR mix into the middle row of the cartridge (the smallest wells) (*see Note 18*). Only push down to the first stop when ejecting liquid, and ensure there are no air bubbles in the sample (*see Note 19*). Using a Rainin multichannel pipettor with Rainin tips is preferred at this stage (*see Note 6*).
  - (b) Pipette 70 µL of oil into the bottom row of wells in the cartridge. Always be sure to pipette the oil after the samples. The top row is left empty.
  - (c) Place a DG8™ rubber gasket over the cartridge by hooking the prongs of the cartridge holder through the gasket's four holes.
7. Place the cartridge holder with cartridge and gasket into the QX200 droplet generator. Close the generator; droplets will be formed. Prepare the next cartridge while the first set of droplets is being generated.
8. When the triangles on the button on the lid of the droplet generator return to being lit solid green and the generator stops making noise, remove the cartridge. Carefully discard its gasket and transfer the droplets in the top row to a clean, half-skirted Eppendorf plate. The output sample has greater volume than the input, so set the Rainin pipette to 40 µL. It is important that the pipetting at this stage is slow and careful, with the pipette oriented at 45°, otherwise the droplets may shear. Afterward, discard the gasket and cartridge (*see Note 20*).
9. After all droplets are made, seal the droplet plate with a foil seal by heating the seal on the plate to 180 °C for 5 s.
10. Thermal cycle the plate as follows:
  - (a) 95 °C for 10 min
  - (b) 40 cycles of 94 °C for 30 s followed by 60 °C for 1 min (*see Notes 21 and 22*)
  - (c) 98 °C for 10 min
  - (d) 8 °C hold
11. Set up a template on the QX200 droplet reader computer. Open QuantaSoft. Under “Template” in the top left corner, select “New” in order to fill in a new plate map. To fill in the information for each sample, double-click on the first

non-empty well. In the “Sample” box, under “Experiment,” select any of the “CNV” experiments, and then, under “Supermix,” select “ddPCR Supermix for Probes” (*see Note 23*). In the “Target 1” box, enter the name of the FAM assay in the “Name” field. From the “Type” menu, select “Ch1 Unknown” if this is the target assay or “Ch1 Reference” if this is the control assay. In the “Target 2” box, enter the name of the HEX or VIC assay in the “Name” field. From the “Type” menu, select “Ch2 Reference” if this is the control assay or “Ch2 Unknown” if this is the target assay. Without closing this menu or double clicking, select all wells of the plate that will contain samples that are using the same assays. Click the blue “Apply” button in the top window to set the assays and experiment for all these wells. Once finished, click “OK” and save the template.

12. Read the droplets on the QX200 droplet reader. Put the plate into the plate holder in the QX200 compartment under the door, then place the black plate holder on top and click the silver tabs on either side down into place, making sure the A1 well is in the top left corner. Close the lid of the QX200. In QuantaSoft on the QX200 computer, make sure the template created in **step 10** is loaded, then click “Run” in the column of options to the left of the plate map. On the popup menu that appears, select “FAM/HEX” or “FAM/VIC,” depending on the pair of fluorophores used, then click “OK.”
- 3.3 Data Finalization and Quality Control**
  1. While the initial output from QuantaSoft can be sufficient for downstream data analysis, careful quality control and optimization of this data often yields more accurate, more reliable copy number calls. So when all the wells containing samples have been run, perform a well-by-well visual inspection of droplet clusters in QuantaSoft by clicking “Analyze” on the leftmost menu followed by “2D Amplitude.” Ensure that there is clear separation between the positive and negative clusters for both the target and reference assay channels. Some bleeding of droplets between the positive and negative channels (sometimes referred to as “rain”) is acceptable, but a substantial amount can cause inaccuracy (Fig. 2). If only a few samples show poor cluster separation, exclude these from analysis. If all wells have bad cluster separation, *see Notes 17, 21, and 22* or redesign the assay according to Subheading [3.1](#).
  2. Determine that the software has made the correct call for each droplet cluster in each well. Make sure that all droplets are correctly labeled by the software: droplets in the top left corner of the 2D amplitude plot are FAM positive only, droplets in the bottom right corner are HEX/VIC positive only, droplets in the top right corner are positive for both fluorophores, and

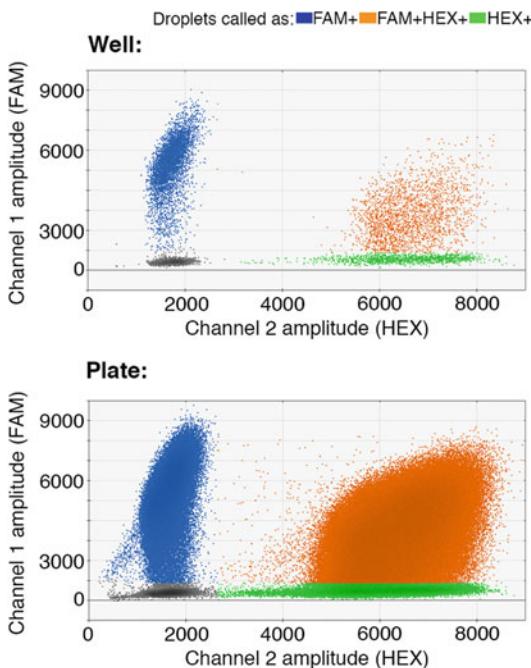


**Fig. 2** Common assay issues and solutions. **(a)** Examples of common issues. A poor separation of clusters (left) can be resolved by optimizing the thermocycling conditions or assay design. Two positive clusters (center) likely result from a SNP being in the assay-binding region or the amplification of a secondary genomic region. Droplet shearing or excess rain (right) can result from not handling the droplets properly. Assays displaying these characteristics should be redesigned or optimized following the suggestions in the protocol. **(b)** Examples of PCR reaction optimization. A temperature gradient (left) can be used to determine the optimal annealing temperature for the PCR, as shown by greatest cluster separation (here, 56.4 °C yields the cleanest clusters). Increasing the number of PCR cycles from 40 (center) to 50 (right) can increase cluster separation to an acceptable amount

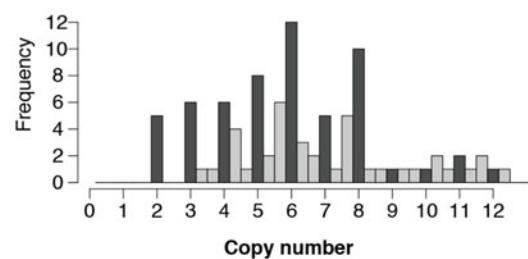
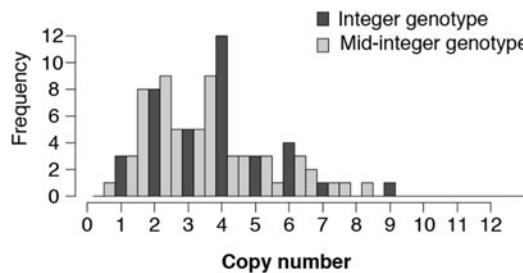
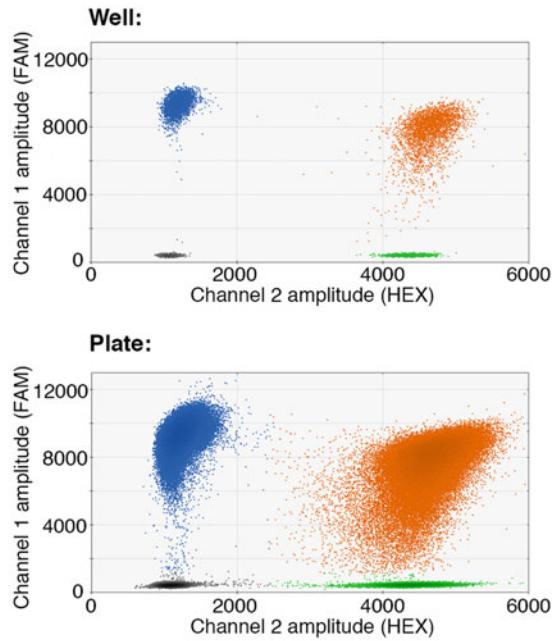
droplets in the bottom left corner are negative for both fluorophores. If a well has some droplets called incorrectly, manually assign them to clusters. (Using QuantaSoft version 1.6.6, this is accomplished by designating the groupings with the “Threshold” or “Lasso” tools). For wells where droplets have been correctly assigned to clusters, make sure the “Status” column is set to “OK”—if it says “Check,” click anywhere in the amplitude plot to get the software to recognize the data (*see Note 24*).

3. Export data for all wells that passed visual inspection. Select these wells and click “Export CSV.”
4. Perform further sample-level quality control (*see Note 25*). Exclude from analysis samples with data drawn from fewer than 5000 droplets (“AcceptedDroplets” column of the exported CSV). Mark samples with mid-integer CNV calls (those 0.35–0.65 away from an integer number). Samples with CNV confidence intervals wider than 1 and samples with fewer than 10% double negative droplets as unreliable; optimize and rerun them (*see Note 26*).
5. Repeat the ddPCR for the rare individual samples that failed visual inspection (**step 1**) or quality control (**step 4**) (*see Note 27*). If many samples failed a given test, there may be a systemic issue that needs to be remedied before repeating (*see Notes 28–30*) (Figs. 2 and 3).

**Un-optimized assay for a difficult CNV locus (AMY1):**



**More optimized assay:**



**Fig. 3** An example of assay optimization for an exceptionally difficult CNV locus (*AMY1* in the amylase locus). The cluster plots (top and middle) and final copy number calls (bottom) both improve after assay and PCR-reaction optimization. The reaction was optimized by designing an assay that conformed to the assay design guidelines in the Methods, running a melting temperature gradient to determine the optimal melting temperature (see Note 20), adding ten extra cycles to the PCR (see Note 19), using a replication-matched control assay (see Note 26), and using the optimal amount of DNA input (see Note 24). It can be improved further by averaging replicates

6. For germline CNV studies where integer copy numbers are expected, round the copy number calls to the nearest integer for all wells that pass visual inspection and quality control if copy numbers generally cluster around integers. If not, a systemic issue may need to be remedied (see Notes 28 and 30) (Figs. 2 and 3).

---

## 4 Notes

1. 20× assay mixes should be kept at –20 °C for long-term storage, but avoid repeated freezing and thawing. We have found 20× mixes to be stable at 4 °C for at least 1 month.
2. Any region of the genome that is known or strongly expected to be invariant in copy number can be used as a control. For human genomes, a particularly well-validated assay targets the *RPP30* gene and is a useful control to use as a starting point for target assay testing. The sequences for this assay are: forward primer, 5'-GATTGGACCTGCGAGCG-3'; reverse primer, 5'-GCGGCTGTCTCCACAAGT-3'; probe, 5'-CTGACCTGAAGGCTCT-3'. When working with aneuploid samples (e.g., those from cancers), it may be necessary to either use multiple control assays or empirically determine the copy number and copy number stability of the control locus.
3. FAM and HEX probes can be ordered through Integrated DNA Technologies. VIC-labeled probes can be ordered through Life Technologies. HEX and VIC probes are read in the same channel in ddPCR, so either a HEX or a VIC probe can be paired with a FAM probe. To minimize costs, we use the more-expensive HEX or VIC fluorophores for control assays, and FAM for the (more numerous and diverse) target locus assays.
4. This protocol is suitable for high-quality, nondegraded DNA from any source (e.g., from cell lines, PBMCs, and fresh tissues). We have had success using ddPCR to type copy number variation in DNA extracted using Qiagen's DNeasy DNA extraction kits. When using DNA from sources, such as FFPE tissue and urine, in which the DNA may be degraded, special assay design considerations should be taken into account (*see Note 10*).
5. To obtain accurate copy number calls, it is crucial that the DNA is digested. In particular, it is important that the restriction enzyme cuts between the assay sites (and not within them). This ensures that intact, individual copies of the region of interest segregate independently into droplets. Any restriction enzyme that accomplishes this goal can be used; AluI is preferred because its recognition site occurs frequently in DNA.
6. Using low-quality pipette tips can cause droplet shredding during droplet generation, likely because they shed tiny pieces of plastic into the reaction. Rainin pipettors and tips perform extremely well when working with droplets but are not strictly necessary in all applications. The use of Rainin pipettors and tips at all stages of DNA extraction, preparation, and droplet

generation ensures that plastic particles are not present in the ddPCR reaction and is considered as the best practice.

7. The segmental duplication track on the UCSC browser includes only regions larger than 1 kilobase; shorter regions of high identity will be missed. For advanced assay design, to determine whether the region of interest is within one of these shorter regions, display the “Mapping” track under the “Mapping and Sequencing” menu. Regions with high (dark) uniqueness values are present only once on the reference, while regions with lower uniqueness values are present more than once.
8. If the region of interest is within a segmental duplication (i.e., present more than once in the reference genome), there are likely differences between the two (or more) copies of the region present (“paralogs”). In this situation, an assay can be designed to target a specific paralog or all copies of the region, depending on the goal of the experiment. For example, a specific paralog might be targeted if its particular function is of interest, while total copy number might be desired if differences between the paralogs are not expected to impact the biological question of interest.

If one paralog is to be targeted, design the assay to exploit differences between the two copies. To find these differences, use the segmental duplication track within the UCSC genome browser to identify the locations of the duplicated regions, then align their sequences and search for sites with several nucleotide differences (termed paralogous sequence variants or PSVs). Design the assays to include these PSVs, particularly by placing the PSVs in the probe-binding region or in the 3' end of primers.

If total copy number is to be targeted, the assay should be designed to avoid differences between the copies. Perform the alignment of the paralogs (segmental duplications) as above, but find regions that do not contain PSVs. Restrict the region used for assay targeting to these PSV-free regions.

9. SNPs in the primer or probe binding sites can prevent or hinder assay binding, so the sites of common SNPs must be excluded from the sequence used to design assays. If the region of interest is duplicated on the reference, make sure to avoid SNPs that occur in any copies of the sequence. One way to avoid SNPs is to turn on the “CommonSNPs(138)” or “CommonSNPs(141)” track under the “Variation” section at the bottom of the USCS genome browser page and subsequently narrow the region of sequence to avoid any common SNPs. This is workable for small regions, but can be tedious for larger ones. Another option is to use the SNP masking tool website: after masking the repeat regions using the UCSC genome

browser, feed that sequence into the SNP masking tool to create a sequence where all SNPs and repeat regions are masked to “N.”

10. When working with high-quality DNA, it is generally unnecessary to match or restrict amplicon sizes beyond the guidelines presented in the Methods. However, if the DNA is composed of short fragments due to degradation or shearing, having PCR amplicons of different lengths can result in a bias, as longer stretches of DNA are less likely to be intact than shorter stretches. Designing target and reference assays that have similar, preferably short, amplicon lengths can maximize and match amplification efficiency between the target and reference regions.
11. G nucleotides at the beginning of probes can quench nearby fluorophores. If Primer3 suggests a probe beginning in G, use the probe’s reverse complement sequence as the assay probe. (Redesign assays if the reverse complement also begins with a G.)
12. If the primers are too short to BLAT, use the in silico PCR function (“In-Silico PCR” under the UCSC genome browser “Tools menu”) instead, though BLAT is preferable. If the region of interest is in a segmental duplication and the assay is designed to target one copy specifically, one unique and perfect match may not be possible; make sure that the best match is the duplication of interest. If the assay is designed to capture all copies of a region, make sure all of these regions are present in BLAT’s output.
13. Multiple pairs of oligos can be checked for heterodimerization using ThermoScientific’s multiple primer tool. This tool is quite sensitive, so use it for preliminary screening and then check any proposed heterodimers with the IDT oligoanalyzer. Paste the named oligo sequences (tab delimited) into the box at the top of the ThermoScientific multiple primer web page to get results.
14. If the amplicon does contain a restriction site for the selected enzyme, change either the enzyme or the assay, making sure that any new enzyme is compatible with the control assay. Generally, it is simpler to redesign the assay, unless the genomic context restricts the assay to a very specific sequence.
15. It is possible to digest the DNA in the ddPCR reaction mixture, rather than predigesting the DNA as explained in Sub-heading 3.2, steps 1–3. This in-Supermix digestion is useful when the sample is limited, as a lower total amount of sample can be used. To perform this in-Supermix digestion, include 2–5 units of restriction enzyme diluted to a volume of 1 µL in the enzyme’s buffer in the ddPCR reaction described in Sub-heading 3.2, step 4, and decrease the volume of the

DNA–water mixture commensurately. Undigested DNA of high concentration, totaling 10 ng, should be substituted for the digested, diluted DNA.

16. ddPCR droplets are made in sets of eight samples at a time and read in 96-well plate format, so it is easiest to set up the PCR in a 96-well plate.
17. Adding 10  $\mu$ L of DNA equates to using 10 ng of DNA in the assay, since only 20  $\mu$ L of this PCR mix is used for droplet generation. Ten nanograms of DNA is generally a good starting amount for ddPCR, but often individuals with high copy numbers need to be regenotyped using half this much DNA to avoid overwhelming the droplets with CNV-containing molecules. In general, if double-negative droplets constitute less than 10% of the droplets generated, decrease the input concentration of DNA; if the error bars on the CNV estimate are too large, increase the input concentration of DNA. In all cases, keep the volume of DNA and water added constant at 10  $\mu$ L. We often genotype all samples using 10  $\mu$ L digested DNA input, then regenotype individuals with high copy numbers or low numbers of double negative droplets using 5  $\mu$ L digested DNA and 5  $\mu$ L water.
18. Though only 20  $\mu$ L of the 25  $\mu$ L PCR mix is used for droplet generation, the 5  $\mu$ L excess prevents air bubbles from being pipetted into the droplet generation reaction, ensuring that the full 20  $\mu$ L is converted into droplets. If the sample is limited, however, a PCR mix with final volume of 22  $\mu$ L can be substituted.
19. Pushing the pipette down to the final stop introduces air bubbles, which compromises the number and quality of droplets. Pipette only until the first stop. If air bubbles are introduced into the sample chamber, manually pop them with a clean pipette tip to improve droplet generation.
20. An automatic droplet generator (Bio-Rad QX200 AutoDG) and associated consumables can be used instead of the manual droplet generation described in Subheading 3.2, steps 6–8. The AutoDG can be run using increments of eight samples, though some reagents are partitioned into 32-sample sets. The AutoDG is best suited for use with full 96-well plates.
21. If clusters are too close together on the scatterplot, the intensity, and thus the separation of the fluorescent signals, may be increased by adding ten extra cycles to the PCR (Figs. 2b and 3b).
22. Although assays are designed to work best at 60 °C, certain assays—or combinations of assays—may give cleaner data at another temperature. We find it best to run a temperature gradient (55–65 °C) on one sample to determine which temperature yields the cleanest, most clearly separated clusters (Fig. 2b).

23. We like to leave the “Name” field blank and merge the final data with a plate map using Excel or another statistical program. Otherwise the name of each and every sample must be entered by hand.
24. The software only calls wells with 10,000 or more droplets and sets the “Status” column to “Check” for wells with fewer droplets, but we have found that CNV calls are reliable down to 5000 droplets, at least for individuals carrying 0–3 copies.
25. This quality control step can be performed in any software for quantitative or statistical analysis. Doing it in R is convenient for automation and repetition, but it can be done in Excel either manually or with formulas.
26. Wells with fewer than 5000 accepted droplets may not contain enough droplets to accurately determine copy number, especially when copy number is above four. Wells with mid-integer CNV calls are not informative (e.g., it is not clear whether a called copy number of 3.5 corresponds to an actual copy number of 3 or 4). Wide confidence intervals suggest the DNA concentration was too low to make a definitive call. Reactions in which fewer than 10% of droplets are negative typically involve situations in which the reaction is too close to saturation with DNA template. In these situations, the Poisson statistics used to estimate the number of droplets with more than one locus copy may be inaccurate, and it may be preferable to rerun the reaction with a lower concentration of genomic DNA. If using a lower concentration of input DNA results in a reference concentration too low to be reliable, multiple wells can be run for each sample, with the resulting data merged during analysis to increase precision.
27. Increasing input DNA concentration typically decreases confidence interval size; samples that failed quality control because of CNV confidence interval size should be repeated with higher DNA input. Mid-integer CNV calls for high copy numbers (above six or so) can often be resolved by repeating the assay using a lower amount of input DNA. For mid-integer calls with lower copy numbers, *see Note 25*.
28. An over-abundance of mid-integer copy number calls can be caused by degraded DNA, undigested DNA, or an incompatibility between the target and control assays. (For cancer samples, it can also reflect clonal mosaicism or mixtures of tumor and stromal cells, and therefore will not benefit from the corrections proposed here.) When DNA is derived from replicating cells (such as a cell line), another cause of mid-integer calls is a difference in the replication timing of the control and target loci. DNA replication occurs in different stages across the genome; this timing is heritable, visible in sequencing data,

and largely the same across individuals [13–15]. DNA from genomic regions that replicate early in the cell cycle is more abundant in asynchronous cell culture, because these regions exist in a duplicated state for much of the cells' lives. Most genes (and as a consequence, most popular control assays) are in these early-replicating regions.

Mid-integer copy number calls are often observed in regions of the genome that replicate late but were paired with an early-replicating control region for ddPCR. In our experience, this discrepancy results in copy number calls that are about 10% under the true call, though this amount varies from sample to sample depending on the proportion of replicating cells at the time of DNA extraction. This can have a large impact, especially on samples with high copy number. Designing a control assay to a region of the genome that replicates at the same time as the region of interest improves copy number analysis. Ideally this control assay could be located very close to (but still genetically outside) the CNV. Replication profiles for lymphoblastoid cell lines can be found using the data from a recent study of replication timing in humans [15]. (Figures 1 and 3b demonstrate using replication-matched controls as well as other optimizations.)

29. In germline-CNV analyses that use DNA derived from proliferating cells (such as the HapMap and 1000 Genomes Project DNAs, widely used as controls), we have found that copy number calls for CNVs on the X chromosome can be improved by using a control assay targeted to nearby X chromosome sequence. The late and unstructured replication of the inactive X chromosome in females and the resulting varying number of X chromosomal regions in asynchronous cell culture may explain this phenomenon [16]. (Figure 1 demonstrates using an X chromosomal control as well as other optimizations.) When using an X chromosomal control assay, make sure to divide the CNV estimates for males given by QuantaSoft by two, as QuantaSoft assumes a diploid control is used and males are haploid for the X chromosome.
30. For germline CNVs, for which integer copy numbers are expected, we have also found that using two different control assays in separate reactions (or two slightly different target assays) and pooling the data to obtain a final copy number increases the proportion of samples with clear integer copy number calls, especially for samples with high copy number. If the same input DNA concentration is used for both repetitions, data can be pooled at the droplet level and then reanalyzed with Poisson statistics. However, if the DNA input was changed between the replicates, data should be pooled by averaging the copy number calls. (Figure 1 demonstrates

pooling the data from two control assays as well as other optimizations.) Alternatively, two control assays with the same fluorophore can be used in the same reaction with the target assay, creating a synthetic four-copy reference. This method may increase the precision of calls made from a single reaction for genomic segments that are present at high copy numbers.

For a few loci, ddPCR may tend to slightly undercount or overcount a genomic locus for an unknown reason; this effect is usually very small at low copy numbers but can become more visible at high copy numbers. If all copy number measurements trend away from integers in the same direction (e.g., if all copy numbers tend to be below integer values), applying a plate-wide multiplicative correction factor that moves all measurements closer to the corresponding integer value appears to be a legitimate correction (as validated by correspondence to sequencing-based measurements of copy number). If attempting this, optimize this correction factor by multiplying the copy numbers by a series of factors between 0.9 and 1.1 (in increments of 0.001) and choose the factor that gives the lowest overall deviation from the closest integer (summing the absolute values of the deviations). Generally, a correction factor within 3% is optimal.

## Acknowledgment

Our understanding of CNVs and assays has benefited greatly from interactions with our colleagues Robert Handsaker, Aswin Sekar, and Linda Boettger. We also thank Katherine Tooley for helpful discussions of this protocol. This work was supported by a grant from the National Human Genome Research Institute (R01 HG006855, to S.A.M.).

## References

- Handsaker RE, Van Doren V, Berman JR, Genovese G, Kashin S, Boettger LM, McCarroll SA (2015) Large multiallelic copy number variations in humans. *Nat Genet* 47(3):296–303. <https://doi.org/10.1038/ng.3200>
- Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC, Kitano TK, Hodel MR, Petersen JF, Wyatt PW, Steenblock ER, Shah PH, Bousse LJ, Troup CB, Mellen JC, Wittmann DK, Erndt NG, Cauley TH, Koehler RT, So AP, Dube S, Rose KA, Montesclaros L, Wang S, Stumbo DP, Hodges SP, Romine S, Milanovich FP, White HE, Regan JF, Karlin-Neumann GA, Hindson CM, Saxonov S, Colston BW (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem* 83(22):8604–8610. <https://doi.org/10.1021/ac202028g>
- Pinheiro LB, Coleman VA, Hindson CM, Herrmann J, Hindson BJ, Bhat S, Emslie KR (2012) Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Anal Chem* 84(2):1003–1011. <https://doi.org/10.1021/ac202578x>
- Usher CL, Handsaker RE, Esko T, Tuke MA, Weedon MN, Hastie AR, Cao H, Moon JE,

- Kashin S, Fuchsberger C, Metspalu A, Pato CN, Pato MT, McCarthy MI, Boehnke M, Altshuler DM, Frayling TM, Hirschhorn JN, McCarroll SA (2015) Structural forms of the human amylase locus and their relationships to SNPs, haplotypes and obesity. *Nat Genet* 47(8):921–925. <https://doi.org/10.1038/ng.3340>
5. Kouprina N, Pavlicek A, Noskov VN, Solomon G, Oststot J, Isaacs W, Carpten JD, Trent JM, Schleutker J, Barrett JC, Jurka J, Larionov V (2005) Dynamic structure of the SPANX gene cluster mapped to the prostate cancer susceptibility locus HPCX at Xq27. *Genome Res* 15(11):1477–1486. <https://doi.org/10.1101/gr.4212705>
6. Salemi M, Bosco P, Cali F, Calogero AE, Soma PF, Galia A, Lanzafame M, Romano C, Vicari E, Grasso G, Sirago P, Rappazzo G (2008) SPANX-B and SPANX-C (Xq27 region) gene dosage analysis in Sicilian patients with melanoma. *Melanoma Res* 18(4):295–299. <https://doi.org/10.1097/CMR.0b013e32830aaa90>
7. Hansen S, Eichler EE, Fullerton SM, Carrell D (2010) SPANX gene variation in fertile and infertile males. *Syst Biol Reprod Med* 55:18–26. <https://doi.org/10.3109/19396360903312015>
8. Boettger LM, Handsaker RE, Zody MC, McCarroll SA (2012) Structural haplotypes and recent evolution of the human 17q21.31 region. *Nat Genet* 44(8):881–885. <https://doi.org/10.1038/ng.2334>
9. Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. *Bioinformatics* 23(10):1289–1291. <https://doi.org/10.1093/bioinformatics/btm091>
10. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. *Nucleic Acids Res* 40(15):e115. <https://doi.org/10.1093/nar/gks596>
11. Andreson R, Puurand T, Remm M (2006) SNPmasker: automatic masking of SNPs and repeats across eukaryotic genomes. *Nucleic Acids Res* 34(Web Server):W651–W655. <https://doi.org/10.1093/nar/gkl125>
12. Vincze T, Posfai J, Roberts RJ (2003) NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic Acids Res* 31(13):3688–3691
13. Hiratani I, Takebayashi S, Lu J, Gilbert DM (2009) Replication timing and transcriptional control: beyond cause and effect—part II. *Curr Opin Genet Dev* 19(2):142–149. <https://doi.org/10.1016/j.gde.2009.02.002>
14. Koren A, Handsaker RE, Kamitaki N, Karlic R, Ghosh S, Polak P, Eggan K, McCarroll SA (2014) Genetic variation in human DNA replication timing. *Cell* 159(5):1015–1026. <https://doi.org/10.1016/j.cell.2014.10.025>
15. Koren A, Polak P, Nemesh J, Michaelson JJ, Sebat J, Sunyaev SR, McCarroll SA (2012) Differential relationship of DNA replication timing to different forms of human mutation and variation. *Am J Hum Genet* 91(6):1033–1040. <https://doi.org/10.1016/j.ajhg.2012.10.018>
16. Koren A, McCarroll SA (2014) Random replication of the inactive X chromosome. *Genome Res* 24(1):64–69. <https://doi.org/10.1101/gr.161828.113>



# Chapter 10

## Assessing *HER2* Amplification in Plasma cfDNA

Isaac Garcia-Murillas and Nicholas C. Turner

### Abstract

Digital PCR (dPCR) is a highly accurate method to determine DNA concentration. In dPCR, DNA is portioned into many discrete single entities, and these are analyzed individually for the presence or absence of a target molecule of interest. Here we describe how digital PCR can be employed to determine the presence of oncogenic amplification through noninvasive analysis of circulating free DNA (cfDNA), and exemplify this approach by developing a plasma circulating free DNA dPCR assay for *HER2* copy number.

**Key words** Breast cancer, *HER2*, Circulating free DNA, Plasma, Digital PCR

---

### 1 Introduction

Genomic amplifications are important therapeutic targets. In routine clinical practice, the presence of amplifications is determined by analysis of a tumor biopsy at initial diagnosis. To optimally deliver targeted therapy, repeated sampling of a tumor is required to determine whether the genetic profile of a cancer has changed following prior therapy. This would require repeated biopsies of recurrent and metastatic cancers, yet this approach has limitations. Biopsy has associated risks, may be technically challenging depending on the site (s) of relapsed disease, and is often costly. A biopsy usually samples only a single area of tumor and in heterogeneous tumors may underestimate the array of genetic aberrations present [1]. Ideally, to overcome these limitations, and to allow repeated sampling, the presence of genomic amplifications could be diagnosed noninvasively.

DNA arising from tumor cells is found in the plasma of patients with cancer and this circulating DNA represents a potential source to noninvasively analyze tumor DNA [2]. High sensitivity assays of coding mutations on circulating free DNA (cfDNA) have reported high concordance with cancer mutational status [3, 4]. Analysis of cfDNA is noninvasive, can be repeated at multiple occasions throughout the disease course, and potentially may assess the full heterogeneity of genomic aberrations present. Analysis of cfDNA

requires an assay of high sensitivity as DNA is frequently present at only low concentration in plasma and tumor cell derived DNA may be only a small fraction of the total plasma DNA [2, 5] with the remainder being derived from non-tumorous cells.

Digital PCR (dPCR) has the potential to accurately quantify the concentration of nucleic acids in a sample, to a much greater degree than traditional quantitative PCR (qPCR), by counting individual DNA molecules [6]. To examine the potential of dPCR for amplification detection we developed an assay to test for *HER2* amplification on cfDNA extracted from plasma [7]. This assay can also be used to test for *HER2* amplification on formalin fixed paraffin embedded (FFPE) tissue and Fresh Frozen Tissue [8]. Chromosomal aneuploidy complicates copy number assessment in plasma, and the appropriate selection of a control probe. Here, we employ a copy number reference gene on the same chromosomal arm as *HER2* that is very rarely coamplified with *HER2*, but is of stable neutral relative copy number in nonamplified cancers.

---

## 2 Materials

Ensure that the work is divided into clearly separated plasma processing, pre-dPCR, and post-dPCR areas and that no reagents or equipment is shared between spaces. This includes lab coats, reagents, plasticware, lab books, notebooks, computers, and any other material.

### 2.1 Blood Collection, Plasma Separation, and Storage

Blood should be collected into Vacutainer Plastic K2 EDTA (BD Biosciences) blood collection tubes as per local guidelines. Blood should be separated using a horizontal rotor swing out head centrifuge ideally within 2 h of collection, aliquoted into cryogenic vials and stored at  $-80^{\circ}\text{C}$ . Avoid repeated freeze-thaw cycles of plasma aliquots.

### 2.2 Circulating Free DNA (cfDNA) Extraction from Plasma

Circulating free DNA (cfDNA) should be extracted from plasma using Qiagen's QIAamp circulating nucleic acid kit on a dedicated plasma processing area using a QIAVac 24 Plus vacuum as instructed by the manufacturer. cfDNA should be eluted into nuclease-free nonstick 1.5-ml microcentrifuge tubes. Extracted cfDNA can be stored short-term at  $-20^{\circ}\text{C}$ . For long-term storage, consider storing cfDNA at  $-80^{\circ}\text{C}$ . Avoid repeated thaw-freeze cycles of extracted cfDNA and consider aliquoting it into smaller volumes.

### 2.3 Quantification and Quality Control of cfDNA Extracted from Plasma

cfDNA should be quantified using a fluorescence-based method like Quant-iT PicoGreen dsDNA assay (Thermo Fisher Scientific) or Qubit dsDNA assays (Thermo Fisher Scientific) on 1  $\mu\text{l}$  of eluate as per manufacturer's instructions. Do not quantify cfDNA using a spectrophotometric method (260/280) as it will not be accurate enough. Alternatively cfDNA can be quantified using Agilent's

**Table 1**  
**Sequences for the primers and probes used for the HER2 amplification assay**

<b>Sequence 5'-3'</b>	
<i>HER2 assay</i>	
Forward primer sequence	ACAACCAAGTGAGGCAGGTC
Reverse primer sequence	GTATTGTTCAGCGGGCTCC
Probe sequence	6-FAM/CCCAGCTCTTGAGGACAAC/MGBNFQ
<i>EFTUD2 assay</i>	
Forward primer sequence	GGTCTTGCCAGACACCAAAG
Reverse primer sequence	TGAGAGGACACACGCAAAAC
Probe sequence	VIC/GGACATCCITTGGCTTTGA/MGBNFQ

**Table 2**  
**Constituents for the HER2 Primer-Probe cocktails**

<b>20× HER2 cocktail</b>	<b>20× EFTUD2 cocktail</b>
90 µl 100 µM primer forward	90 µl 100 µM primer forward
90 µl 100 µM primer reverse	90 µl 100 µM primer reverse
25 µl 100 µM HER2 FAM/MGBNFQ probe	25 µl 100 µM EFTUD2 VIC/MGBNFQ probe
295 µl nuclease-free H <sub>2</sub> O	295 µl nuclease-free H <sub>2</sub> O

Bioanalyzer/Tapestation (or similar) when checking its quality. cfDNA quality should be checked using Agilent's Bioanalyzer/Tapestation (or similar). These procedures should be carried out on a dedicated pre-PCR area and with dedicated pre-PCR equipment.

#### **2.4 HER2 Droplet Digital PCR**

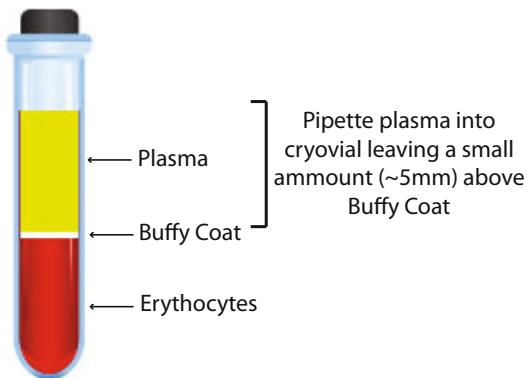
Droplet digital<sup>TM</sup> PCR (ddPCR<sup>TM</sup>) should be performed on Bio-Rad's QX100/QX200, as described in Subheading 3, in a pre-PCR area. Reactions should be performed using 2× ddPCR Supermix for probes (Bio-Rad). Primers and probes (Table 1) should be used at a final concentration of 900 nM and 250 nM respectively on a final volume of 20 µl PCR mix. Prepare Primer-Probe cocktails as shown in Table 2.

---

### **3 Methods**

#### **3.1 Venous Blood Collection, Isolation of Plasma from Blood and Storage**

1. Collect blood via vein puncture directly into an EDTA blood tube.<sup>a</sup>
2. Centrifuge at 1,600 × g for 20 min at room temperature on a horizontal rotor, swing out head centrifuge.<sup>b</sup>



**Fig. 1** When separating plasma for extraction of cfDNA, only the top layer should be taken. The buffy coat layer containing white blood cells should be left undisturbed

3. Remove plasma carefully without disturbing the buffy coat layer (white blood cells and platelets) and transfer into a clean conical centrifuge tube (Fig. 1).<sup>c</sup>
4. Centrifuge plasma at 16,000 rpm for 10 min to pellet cellular debris.
5. Pipette off all the plasma into labeled Cryogenic vials.<sup>d</sup>
6. Place the cryogenic vials containing plasma upright into a freezer storage box in a -80 °C freezer.<sup>e</sup>

### 3.2 Circulating Free DNA Extraction from Plasma Using a QIAvac

Extract cfDNA from 2 ml of plasma using the QIAamp circulating nucleic acid kit from Qiagen. Ensure that the plasma has been thawed on ice before starting the procedure.

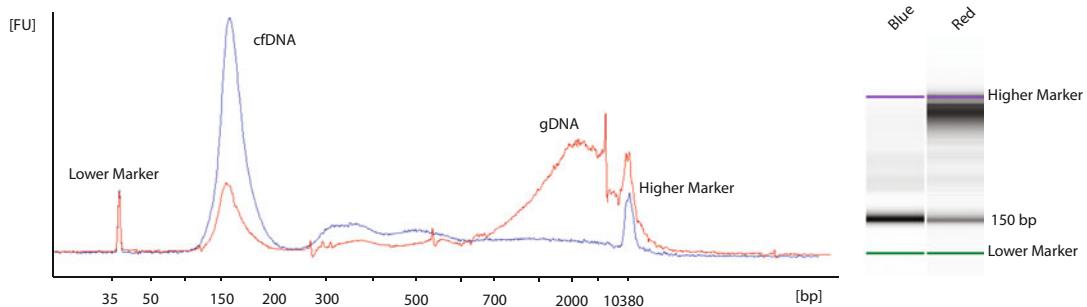
1. Pipet 200 µl QIAGEN Proteinase K into a 50 ml centrifuge tube.
2. Add 2 ml of plasma to the 50 ml tube.
3. Add 1.6 ml Buffer ACL (containing 1.0 µg carrier RNA). Close the cap and mix by pulse-vortexing for 30 s.<sup>a</sup>
4. Incubate at 60 °C for 30 min.
5. Place the tube back on the lab bench and unscrew the cap.
6. Add 3.6 ml Buffer ACB to the lysate in the tube. Close the cap and mix thoroughly by pulse-vortexing for 15–30 s.
7. Incubate the lysate–Buffer ACB mixture in the tube for 5 min on ice.
8. Insert the QIAamp Mini column into the VacConnector on the QIAvac 24 Plus. Insert a 20 ml tube extender into the open QIAamp Mini column.<sup>b</sup>
9. Carefully apply the lysate–Buffer ACB mixture from step 7 into the tube extender of the QIAamp Mini column. Switch on the

vacuum pump. When all lysate have been drawn through the column completely, switch off the vacuum pump and release the pressure. Carefully remove and discard the tube extender.<sup>c</sup>

10. Apply 600  $\mu$ l Buffer ACW1 to the QIAamp Mini column. Leave the lid of the column open, and switch on the vacuum pump. After all of Buffer ACW1 has been drawn through the QIAamp Mini column, switch off the vacuum pump and release the pressure.
11. Apply 750  $\mu$ l Buffer ACW2 to the QIAamp Mini column. Leave the lid of the column open, and switch on the vacuum pump. After all of Buffer ACW2 has been drawn through the QIAamp Mini column, switch off the vacuum pump and release the pressure.
12. Apply 750  $\mu$ l of ethanol (96–100%) to the QIAamp Mini column. Leave the lid of the column open, and switch on the vacuum pump. After all of ethanol has been drawn through the spin column, switch off the vacuum pump and release the pressure.
13. Close the lid of the QIAamp Mini column. Remove it from the vacuum manifold, and discard the VacConnector. Place the QIAamp Mini column in a clean 2 ml collection tube, and centrifuge at full speed (14,000 rpm) for 3 min.
14. Place the QIAamp Mini column into a new 2 ml collection tube. Open the lid, and incubate the assembly at 56 °C for 10 min to dry the membrane completely.
15. Place the QIAamp Mini column in a clean 1.5 ml elution tube and discard the 2 ml collection tube from **step 14**. Carefully apply 25  $\mu$ l of Buffer AVE to the center of the QIAamp Mini column membrane. Close the lid and incubate at room temperature for 3 min.<sup>d</sup>
16. Centrifuge in a microcentrifuge at full speed (14,000 rpm) for 1 min to elute the nucleic acids.
17. Repeat **steps 15** and **16** again eluting into the same tube. Final volume ~50  $\mu$ l, labeled as eluate #1.
18. Repeat **steps 15** and **16** (2  $\times$  25  $\mu$ l) eluting into a fresh 1.5 ml elution tube. Label as eluate #2.
19. Store extracted cfDNA at –20 °C.

### **3.3 Quantification of Extracted cfDNA**

1. Take the Qubit dsDNA Assay Kit out of the freezer and equilibrate at room temperature for 30 min.
2. Make the Qubit working solution by diluting the Qubit dsDNA BR reagent 1:200 in Qubit dsDNA BR buffer. Do not mix the working solution in a glass container. Label two tubes for the standards.



**Fig. 2** cfDNA extracted from plasma free from contaminant gDNA should exhibit a peak at around 150–300 bp when analyzed on the Bioanalyzer/Tapestation (blue trace), while cfDNA contaminated with gDNA will show an extra peak/s above 1000 bp (red trace)

3. Load 190  $\mu$ l of Qubit working solution into each of the labeled standard tubes.
4. Add 10  $\mu$ l of each Qubit standard to the appropriate tube and mix by vortexing 2–3 s. The final volume in each tube should be 200  $\mu$ l.
5. Load 199  $\mu$ l of Qubit working solution into individual assay tubes.
6. Add 1  $\mu$ l of plasma eluate to the tubes containing Qubit working solution. The final volume in each tube should be 200  $\mu$ l.
7. Incubate tubes at room temperature for 2 min.
8. Read the standards and plasma samples on a Qubit 2.0 Fluorometer as instructed by the apparatus.
9. Calculate the concentration of the plasma samples taking into account the dilution step on [6].

### 3.4 Checking cfDNA Quality Using Agilent's Bioanalyzer

Run 1  $\mu$ l of eluate on a High Sensitivity DNA chip as per manufacturer's instructions. Assess the quality of the extracted cfDNA based on the presence of a peak at around 150–180 bp and the absence of a high molecular weight peaks of more than 1000 bp (Fig. 2, blue trace). Genomic DNA (gDNA) contaminated cfDNA will exhibit high molecular weight peaks of more than 1000 bp (Fig. 2, red trace).

### 3.5 Droplet Digital PCR (ddPCR) for HER2 Amplification

1. For each sample prepare a reaction mix, minus the DNA, into nuclease-free nonstick 1.5-ml microcentrifuge tubes as shown in Table 3.
2. Aliquot reaction mix into 0.2 ml PCR plate wells (or tubes), add DNA to be analyzed to tube, vortex and spin down.
3. Assemble droplet generator cartridge into cartridge holder.
4. Aliquot 20  $\mu$ l of PCR reaction mix onto central row reservoirs of droplet generator cartridge.<sup>a</sup>

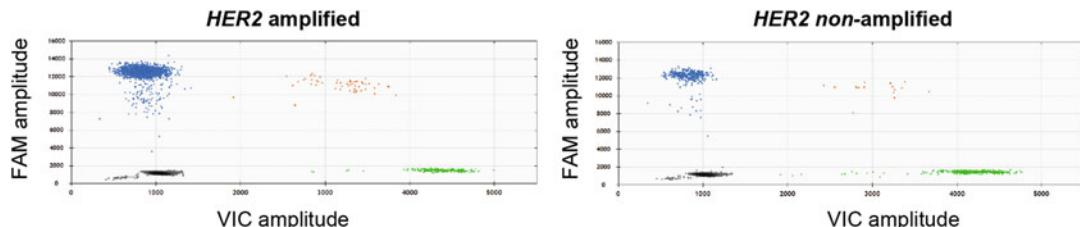
**Table 3**  
**PCR master mix components and volumes used for HER2 amplification testing on cfDNA**

	$\times 1 \text{ (\mu l)}$
DNA	x
20× HER2 cocktail	1
20× EFTUD2 cocktail	1
Supermix for probes	10
Nuclease-free H <sub>2</sub> O	to 20

**Table 4**  
**PCR conditions for HER2 amplification testing on cfDNA extracted from plasma**

Step	Temperature (°C)	Time (min)	# Cycles	Ramp temp increment
Heated lid	105			
Denaturing	95	10:00	1	2.5°C/s
Denaturing	95	00:15	40	2.5°C/s
Anneling/ extension	60	01:00	1	2.5°C/s
Completion	98	10:00	1	2.5°C/s
Hold	10	Indefinite	1	

5. Aliquot 70  $\mu\text{l}$  of droplet generator oil (using same tip for all wells) onto bottom row reservoirs of droplet generator cartridge.<sup>b</sup>
6. Cover cartridge with gasket making sure ends fit over plastic overhangs. Place into droplet generator.<sup>c</sup> When droplets have been generated, remove cartridge from generator, remove gasket, dispose and collect 40  $\mu\text{l}$  of droplets from top row to dispense into a PCR plate (Fisher Scientific). Collect droplets using an eight channel multichannel air displacement pipette by tipping pipette at 20–25° angle and slowly (up to 15 s) collect the droplets.<sup>d</sup> Dispense droplets on the PCR plate by tipping the pipette 20–25° angle and touching the wall of the plate well, slowly (up to 15 s) release droplets by letting them slide along the wall of the well. Dispose of cartridge.<sup>e</sup>
7. Repeat as many times as needed, covering wells between droplet generation steps.
8. Seal Plate and run samples on thermal cycler as shown in Table 4.



**Fig. 3** Representative droplet digital plots from a sample with high level amplification (a), and a non-amplified sample (b). The four quadrants represent top left—droplets with *HER2* DNA only (blue population), top right—droplets with both *HER2* and *EFTUD2* DNA (brown population), bottom right—droplets with *EFTUD2* DNA only (green population), and bottom left—droplets with no DNA (black population)

9. When PCR cycling is complete, read droplets on a QX100/QX200 Droplet Reader as per manufacturer's instructions. Read the plate using Absolute Quantification.
10. Samples that do not have at least 400 WT droplets should be enhanced by running more sample until at least 400 WT droplets are analyzed.

### 3.6 Droplet Digital™ PCR (ddPCR) Analysis

Data produced on the QX100/QX200 reader may be analyzed using Bio-Rad QuantaSoft™ software. Gate the four distinct populations produced by the assay as shown in Fig. 3 and use the built-in analysis tool to calculate the ratio of *HER2* to *EFTUD2* reference assay. Values above a ratio of 1.25 should be considered amplified, while ratios below that should be considered non-amplified.

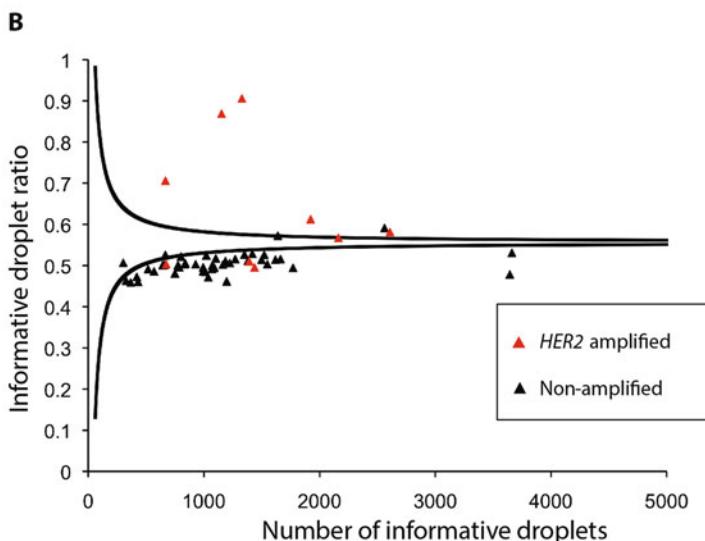
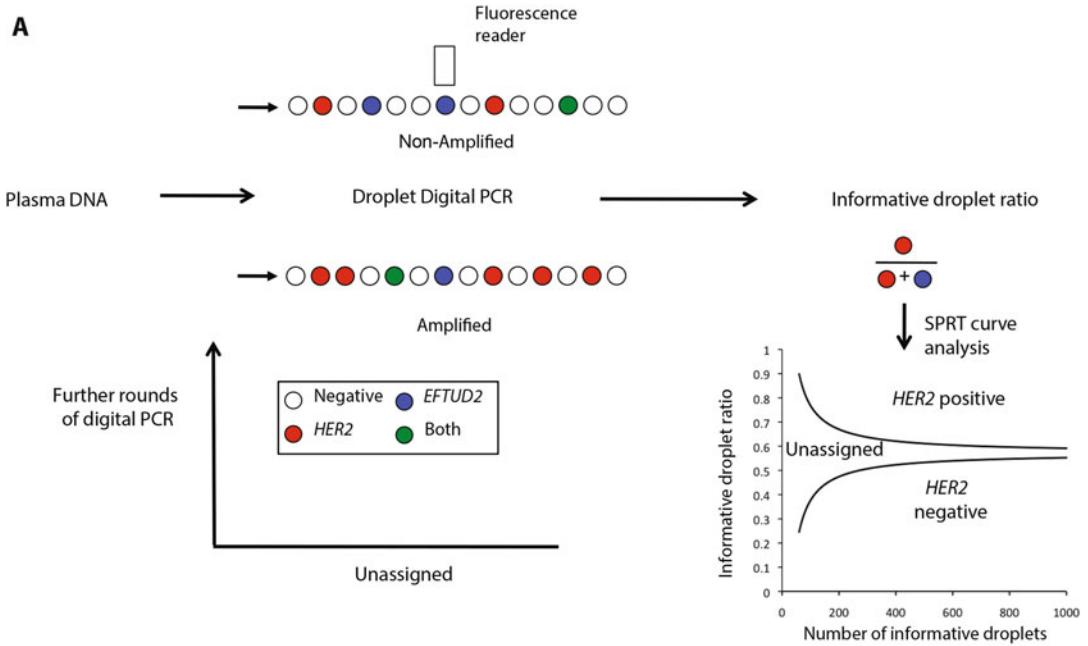
For optimal results in samples, and in particular in samples with a copy number ratio near to the cutoff, the sequential probability ratio test (SPRT) (Fig. 4a) should be used to determine whether the copy number is elevated, not-elevated, or indeterminate. The SPRT uses the maximum likelihood method to determine whether the observed ratio is above or below the threshold within acceptable bounds of error, or whether further digital PCR runs are required to determine the copy number ratio.

For data analysis using the SPRT test, a threshold likelihood ratio of 8, as previously reported with modifications [9, 10], should be used. Only informative droplets should be analyzed, i.e., those droplets positive for either *HER2* alone or *EFTUD2* alone. The proportion of informative droplets positive for *HER2* is calculated as  $P_{HER2} = N_{HER2}/N$  where  $N = (N_{HER2} + N_{EFTUD2})$  the number of informative droplets,  $N_{HER2}$ , is the number of droplets positive for *HER2* alone, and  $N_{EFTUD2}$ , is the number of droplets positive for *EFTUD2* alone.

The boundaries of the SPRT test are calculated as follows:

$$\text{Upper} = ((\ln 8)/N - \ln d)/\ln \gamma$$

$$\text{Lower} = ((\ln 1/8)/N - \ln d)/\ln \gamma$$



**Fig. 4 (a)** Droplet digital PCR with a FAM labeled *HER2* probe and VIC labeled *EFTUD2* probe. DNA is partitioned into droplets, and after PCR, droplets are assessed by a fluorescent reader. The concentration of DNA in each sample can be quantified from the number of wells positive using the Poisson distribution. This is further analyzed with the SPRT using informative droplets, those droplets positive for *HER2* or *EFTUD2* alone, and not those positive for both or neither. The SPRT assesses whether the proportion of informative wells positive for *HER2*, informative wells ratio, is elevated as data accumulates. SPRT defines two boundaries, with a ratio above the upper boundary being considered *HER2*-positive and below the lower boundary considered *HER2*-negative. A ratio between the two boundaries is considered as unassigned, and the sample is subjected to further rounds of digital PCR until the result is above or below the boundaries. **(b)** Analysis of digital PCR with Sequential Probability Ratio Test (SPRT) using informative droplets on a cohort of 58 metastatic breast cancer patients, 11 patients with *HER2*-amplified cancers and 47 patients with *HER2*-nonamplified cancers. Red triangles indicate patients with *HER2*-amplified tumors and black triangle *HER2*-nonamplified tumors. Adapted from H. Gevensleben et al., Noninvasive detection of *HER2* amplification with plasma DNA digital PCR. *Clinical cancer research: an official journal of the American Association for Cancer Research* **19**, 3276–3284 (2013) with permission from The American Association of Cancer Research (AACR)

where

$$d = (1 - q_1)/(1 - q_0)$$

$$g = q_1(1 - q_0)/q_0(1 - q_1)$$

and

$q_1$  = the proportion of informative droplets positive for *HER2* if alternative hypothesis is accepted (plasma sample is from a patient with *HER2* amplification).

$q_0$  = the proportion of informative droplets positive for *HER2* if null hypothesis is accepted (plasma sample is from a patient without *HER2* amplification).

$q_1$  is calculated from *HER2/EFTUD2* copy number ratio ( $T_{AMP} = 1.3$ ) for assigning a sample as *HER2* positive sample, and varies according to  $M_{EFTUD2}$ .

$$q_1 = (X_{AMP} - X_{AMP}n_{EFTUD2}/n) /$$

$$(X_{AMP} + n_{EFTUD2}/n - 2X_{AMP}n_{EFTUD2}/n)$$

where  $X_{AMP} = 1 - \exp(-T_{AMP}M_{EFTUD2})$  the expected proportion of informative droplets positive for *HER2* at the threshold ratio  $T_{AMP}$ .

$q_0$  is similarly calculated from *HER2/EFTUD2* copy number ratio ( $T_{NONAMP} = 1.2$ ) for assigning a sample as *HER2* positive or negative, and also varies according to  $M_{EFTUD2}$ .

$$q_0 = (X_{NONAMP} - X_{NONAMP}n_{EFTUD2}/n) /$$

$$(X_{NONAMP} + n_{EFTUD2}/n - 2X_{NONAMP}n_{EFTUD2}/n),$$

where  $X_{NONAMP} = 1 - \exp(-T_{NONAMP}M_{EFTUD2})$ .

Using these equations for any given  $N$ , the upper and lower boundaries of the SPRT curve are calculated. If  $P_{HER2}$  is greater than the upper boundary then the test result is *HER2* positive. If  $P_{HER2}$  is less than the lower boundary then the test result is *HER2* negative. If  $P_{HER2}$  lies between the two boundaries, then further round(s) of digital PCR are required until a sample is above, or below, the boundaries.

The SPRT uses a likelihood ratio of 8, which corresponds roughly to two sided 95% confidence for differentiating between ratios of 1.30 and 1.20 [9]. With these parameters a sample with an actual *HER2:EFTUD2* ratio of, for example, 1.15 would have a very high probability (>99.9%) of being correctly called as negative, and likewise a sample with an actual ratio of 1.35 would have a very high probability of being correctly called as positive.

Figure 4b illustrates the use of the SPRT test on a cohort of 58 metastatic breast cancer patients (11 with *HER2*-amplified cancers and 47 with *HER2*-nonamplified cancers). Red triangles indicate patients with *HER2*-amplified tumors and black triangle *HER2*-non-amplified tumors. The displayed SPRT decision

boundaries are for illustrative purposes only, as the exact level varies according to the *EFTUD2* control probe concentration ( $M_{EFTUD2}$ ), with the displayed boundaries calculated with  $M_{EFTUD2} = 0.025$ . Cases with a number of informative droplets >5000 are not displayed [7].

---

## 4 Notes

1. Venous blood collection, isolation of plasma from blood and storage (Section 3.1)
  - (a) When collecting blood, hemolysis must be avoided.
  - (b) EDTA BCTs should be stored in an upright position at room temperature and centrifuged within 2 h of sample collection.
  - (c) Invert the ETDA tube gently 8–10 times immediately prior to centrifugation.
  - (d) Use a piston-driven air displacement pipette (avoid using a Pasteur pipette) to remove the plasma.
  - (e) When freezing plasma do not use polystyrene boxes.
2. Circulating free DNA extraction from plasma using a QIAvac (Section 3.2)
  - (a) Make sure that a visible vortex forms in the tube. In order to ensure efficient lysis, it is essential that the sample and Buffer ACL are mixed thoroughly to yield a homogeneous solution. Do not interrupt the procedure at this time. Proceed immediately to the lysis incubation step.
  - (b) Make sure that the tube extender is firmly inserted into the QIAamp Mini column in order to avoid leakage of sample. Keep the collection tube for the dry spin in **step 13**.
  - (c) Note that large sample lysate volumes may need up to 10 min to pass through the QIAamp Mini column membrane by vacuum force. To avoid cross-contamination, be careful not to move the tube extenders over neighboring QIAamp Mini columns.
  - (d) Ensure that the elution buffer AVE has been warmed to 42 °C before use. Elution volume is flexible and can be adapted according to the requirements of downstream applications. The recovered eluate volume will be up to 5 µl less than the elution volume applied to the QIAamp Mini column.
3. Droplet digital PCR (ddPCR) for HER2 amplification (Section 3.5)
  - (a) Carefully load the cartridge and do not introduce bubbles onto the bottom of the reservoirs. Always fill all wells on a

generator cartridge. If less than eight samples then fill empty wells with 20 µl of 2× Buffer Control kit.

- (b) Droplet generator oil is volatile, so avoid leaving the tube or well open for very long time to prevent evaporation. Always add the oil after having added the sample to the central row of the reservoir to avert oil filling up the microchannels connecting the reservoirs.
- (c) When loading the cartridge onto the droplet generator, be careful as to not spill oil (hold by central part). Droplets will be generated in around 2 min.
- (d) Collecting droplets slowly will prevent droplets from breaking up and sticking together.
- (e) Once droplets are generated and deposited onto the PCR plate cover filled wells with tape to prevent evaporation.

## Acknowledgments

Prof. Nicholas C. Turner is a CRUK Clinician Scientist. We acknowledge NHS funding to the NIHR Biomedical Research Centre. Figure 4 is adapted from H. Gevensleben et al., Noninvasive detection of HER2 amplification with plasma DNA digital PCR. *Clinical cancer research: an official journal of the American Association for Cancer Research* **19**, 3276–3284 (2013) with permission from The American Association of Cancer Research (AACR).

## References

1. Ding L et al (2010) Genome remodelling in a basal-like breast cancer metastasis and xeno-graft. *Nature* **464**:999–1005
2. Johnson PJ, Lo YM (2002) Plasma nucleic acids in the diagnosis and management of malignant disease. *Clin Chem* **48**:1186–1193
3. Li M, Diehl F, Dressman D, Vogelstein B, Kinzler KW (2006) BEAMing up for detection and quantification of rare sequence variants. *Nat Methods* **3**:95–97
4. Board RE et al (2010) Detection of PIK3CA mutations in circulating free DNA in patients with breast cancer. *Breast Cancer Res Treat* **120**:461–467
5. Diehl F et al (2008) Circulating mutant DNA to assess tumor dynamics. *Nat Med* **14**:985–990
6. Vogelstein B, Kinzler KW (1999) Digital PCR. *Proc Natl Acad Sci U S A* **96**:9236–9241
7. Gevensleben H et al (2013) Noninvasive detection of HER2 amplification with plasma DNA digital PCR. *Clin Cancer Res* **19**:3276–3284
8. Garcia-Murillas I, Lambros M, Turner NC (2013) Determination of HER2 amplification status on tumour DNA by digital PCR. *PLoS One* **8**:e83409
9. Wei Zhou GG, Goodman SN, Romans KE, Kinzler KW, Vogelstein B, Choti MA, Montgomery EA (2001) Counting alleles reveals a connection between chromosome 18q loss and vascular invasion. *Nat Biotechnol* **19**:78–81
10. Lo YM et al (2007) Digital PCR for the molecular detection of fetal chromosomal aneuploidy. *Proc Natl Acad Sci U S A* **104**:13116–13121



# Chapter 11

## Detection and Quantification of Mosaic Genomic DNA Variation in Primary Somatic Tissues Using ddPCR: Analysis of Mosaic Transposable-Element Insertions, Copy-Number Variants, and Single-Nucleotide Variants

Bo Zhou, Michael S. Haney, Xiaowei Zhu, Reenal Pattni, Alexej Abyzov, and Alexander E. Urban

### Abstract

Here, we describe approaches using droplet digital polymerase chain reaction (ddPCR) to validate and quantify somatic mosaic events contributed by transposable-element insertions, copy-number variants, and single-nucleotide variants. In the ddPCR assay, sample or template DNA is partitioned into tens of thousands of individual droplets such that when DNA input is low, the vast majority of droplets contains no more than one copy of template DNA. PCR takes place in each individual droplet and produces a fluorescent readout to indicate the presence or absence of the target of interest allowing for the accurate “counting” of the number of copies present in the sample. The number of partitions is large enough to assay somatic mosaic events with frequencies down to less than 1%.

**Key words** Droplet digital PCR (ddPCR), Somatic mosaicism, Mobile elements, Copy number variations (CNVs), Single nucleotide variations (SNVs)

---

### 1 Introduction

Droplet digital polymerase chain reaction (ddPCR) is a powerful approach to validate and to quantify somatic genomic mosaicism in primary tissues [1–3]. Somatic genomic mosaicism is the phenomenon where variations in the genomic DNA sequence are present in only a subset of cells comprising a given tissue. This type of genomic variation is presumed to arise within an individual due to postzygotic mutation [4]. Somatic genomic variants will typically have been detected first using a variety of “next-generation” high-throughput DNA sequencing-based approaches [5]. In addition to

---

Bo Zhou and Michael S. Haney contributed equally to this work.

validating such sequencing-based variant-detections, ddPCR can also more accurately quantify the frequency at which the variant occurs in a given tissue [1, 3]. Furthermore, ddPCR is one of only a few approaches that can be used for such a purpose (validation *and* quantification of mosaic genomic variation). Since it is a nonsequencing based method, it serves as an orthogonal approach to the typical method of discovery in somatic mosaicism studies [5]. This makes validation and quantification of somatic variants with ddPCR particularly useful, especially in the context of a topic where extensive and specific experimental validation of initial discoveries is essential in order to avoid the reporting of sequencing artifacts. The approaches described here have been tested in DNA samples extracted either directly from noncancerous primary tissue or non-cancer cell lines. We do however expect these same approaches to be equally applicable to cancer tissues or cell lines, respectively.

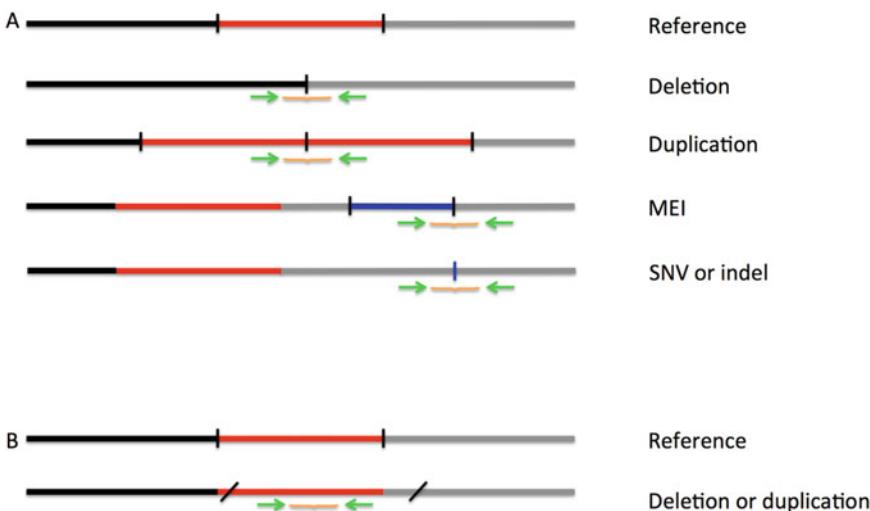
In this context of somatic genomic mosaicism, ddPCR can be used in the analysis of genomic variants spanning the entire spectrum of sizes, from the smallest (single-nucleotide variants, SNVs) over medium-sized (e.g., insertions of retrotransposons and mobile genomic elements in general—i.e., mobile element insertions, MEIs) to large copy number variants (CNVs) [1, 3, 6]. There are only minor adjustments needed to the general ddPCR protocol and assay design to allow for the analysis of each of these types of variants. Here, we describe how to employ ddPCR, specifically the Bio-Rad QX200 Droplet Digital PCR system [7], in the analysis of three types of somatic mosaic genome variants: SNVs, MEIs and CNVs. By combining high throughput microfluidics with quantitative (digital) fluorescent readouts, using either TaqMan style probes or an intercalating fluorescent dye (EvaGreen), ddPCR assays achieve unprecedented levels of sensitivity and accuracy in PCR-based detection [7, 8].

### 1.1 Principle of Approach

Briefly, a standard ddPCR reaction mixture, containing template DNA, target specific primers and fluorescent probes (or intercalating fluorescent dye, *see* below), is partitioned into approximately 20,000 oil-in-water droplets [7, 8]. The most crucial aspect of understanding the experimental approach is to keep in mind that while all other reagents are in excess, the template DNA is partitioned into the droplets following a Poisson manner [7], and as a result, target occupancy in droplets or copies per droplet (cpd) increases with input DNA concentration. At <10% target occupancy or <0.1 cpd the vast majority of droplets contain either 0 or 1 template copy of your PCR target of interest. While maximal precision for target copy number quantification is achieved at 1.6 cpd or 32,000 human genome equivalents per 20 µL reaction, accurate quantification can be achieved as long as negative droplet partitions are present in the reaction (<5 cpd) [8]. Overloading of template DNA (>5 cpd) results in the absence of null droplets

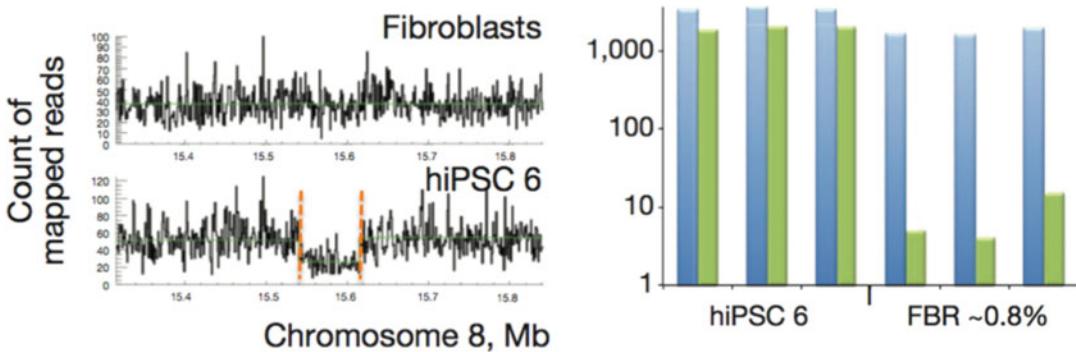
(droplets that contain zero copies of template target), thus Poisson correction cannot be applied rendering accurate quantification impossible [9, 10]. It is also important to note that the high degree of sample DNA partition in ddPCR allows for the detection and quantification of low-frequency sequence variants by giving a binary-like readout for the presence or absence of the genetic variant in each droplet [8]. The partitioning effect is perhaps most beneficial (compared to unpartitioned qPCR) for the detection and quantification of SNVs where the effect of imperfect probe specificity is “exploited” by partitioning allowing for the discernment between reference and variant signals (*see* Subheading 3.5).

For each individual DNA sequence variant, a set of PCR primers is designed, as well as a specific fluorophore-conjugated, TaqMan-style probe that will bind to the template DNA during the primer extension step. Alternatively, when using the EvaGreen assay (i.e., an intercalating fluorescent dye), only PCR primers are required. There are two basic principles of probe/primer design. Primers can be designed to “bracket” a predicted mosaic SNV, CNV-deletion junction, CNV-duplication junction or MEI-insertion point or designed to bind within the boundaries of a predicted mosaic deletion or duplication CNV. These principles are illustrated in Fig. 1.



**Fig. 1** The basic principles of primer and probe design for ddPCR in the analysis of somatic genomic mosaics. Panel **a**: if the breakpoint junction of a suspected mosaic CNV, the insertion point of a mosaic MEI event, the coordinate of a mosaic SNV, or the location of a small indel is exactly known, then a primer-and-probe set can be designed to cover these exact coordinates (green arrows: forward and reverse primers, orange bracket: specific TaqMan style fluorophore conjugated oligomer or EvaGreen dye intercalating into amplicon). Panel **b**: if the exact coordinates of a suspected mosaic CNV are not known then primer-and-probe sets can be designed to produce an amplicon from a locus somewhere between the suspected approximate breakpoint coordinates (green arrows: forward and reverse primers, orange bracket: specific TaqMan style fluorophore conjugated oligomer or EvaGreen dye intercalating into amplicon). In this latter scenario somatic mosaic MEIs and SNVs/small indels are not accessible to ddPCR analysis by approaches covered in this chapter. However, these may be accessible with alternative EvaGreen approaches [6]

Under the *first* principle of probe-and-primer design (Fig. 1a), mosaic CNVs and SNVs can be detected because the custom fluorescent probes are designed such that it spans the breakpoint junction of the mosaic deletion-CNV or the insertion-point sequence of the mosaic duplication-CNV or MEI, or sits right on top of the mosaic SNV. If a given droplet with the partitioned sample DNA contains a DNA fragment with this specific mosaic breakpoint or insertion point junction or SNV, there will be a discrete digital readout. In droplets that do not contain such a fragment (because the frequency of its occurrence in the sample DNA is low or it may even be completely absent) there will simply be no signal at all. With this binary, digital readout of the proportion of cells harboring the mosaic variant in the sample tissue can be calculated with sensitivity as low as 1% mosaic frequency for CNVs and less than 0.1% for SNVs and MEIs. For example, consider a deletion-CNV that had been detected in a clonally expanded iPSC line by shallow whole-genome sequencing, but not in the fibroblast culture from which the iPSC line hiPSC 6 was derived (Fig. 2, left panel) [1]. Using ddPCR, it was established that this deletion-CNV was present as a heterozygous mosaic variant in ~0.8% of the cells from the original fibroblast culture and present in ~100% of the cells in the iPSC culture (Fig. 2, right panel), suggesting that the iPSC line was derived and clonally expanded from a fibroblast cell harboring



**Fig. 2** Detection and quantification of a lineage-manifested CNV (LM-CNV). Left panel: low-coverage whole-genome sequencing (using CNVnator [11]) could not detect a CNV that was present in mosaic fashion in a fibroblast sample. After clonal expansion of a single cell from the fibroblast sample in the process of producing human induced pluripotent stem cells (hiPSCs) the CNV became unmasked in low-coverage whole-genome sequencing (dashed orange lines). Right panel: ddPCR analysis could confirm the presence of the LM-CNV detected in hiPSC line 6 in the fibroblast sample from which hiPSC 6 was created as well as quantify its mosaic allele frequency. y-axis: number of ddPCR events per well. Blue bars: reference probe. Green bars: LM-CNV specific probe. The frequency of CNV events detected in hiPSC 6 indicates that this is a heterozygous deletion present in almost all cells while the frequency of CNV events detected in fibroblast samples indicate this heterozygous deletion is present in ~0.8% of cells. ddPCR was carried out in three replicates in each of the hiPSC and fibroblast genomic DNA samples (Figure adapted with permissions from [1])

this specific deletion-CNV in the mosaic cell population [1]. The *second* principle of probe-and-primer design (Fig. 1b) can be applied when the genomic coordinates of the breakpoint junction of a CNV is ambiguous or unknown. Primers and probes can be designed to target the predicted region of duplication or deletion rather than spanning the breakpoint junction. The sensitivity of this approach is lower than that of bracketing a given variant, and it cannot be applied to analyze a specific MEI or SNV. On the other hand, however, this second approach can be used to still achieve validation and at least limited quantification of such CNVs that have been detected by sequencing but for which the exact breakpoint sequences have not been resolved. In all scenarios, it is recommended to pair the mosaic variant assay of interest with an assay for a reference gene that simultaneously quantifies the number of genome equivalents in the input sample (*see Note 1*).

---

## 2 Materials

In order to prevent contamination, carefully follow general PCR preparation guidelines and store ddPCR reagents and consumables away (ideally in separate rooms) from where ddPCR assays are prepared. All oils, cartridges, and gaskets should be stored at room temperature unless otherwise noted. When preparing assays, primers and probes should be thawed and kept on ice (or stored at 4 °C for up to 1 year); probes should be kept in the dark or wrapped in aluminum foil; all other reagents should be prepared at room temperature. Follow all waste disposal regulations when disposing of waste materials.

### 2.1 Instruments

1. QX200 Droplet Generator (Bio-Rad).
2. PX1 PCR Plate Sealer (Bio-Rad).
3. 96-well thermal cycler with gradient temperature feature (e.g., Bio-Rad C1000 or Applied Biosystems Veriti).
4. QX200 Droplet Reader (Bio-Rad).
5. PC with QuantasSoft™ Software (Bio-Rad).
6. Eight-channel P100 or P200 Pipette.
7. Single channel P2, P20, and P200 pipettes for single reaction preparations.

### 2.2 ddPCR Consumables (All Bio-Rad)

1. DG8 Cartridges for QX200 Droplet Generator (Cat# 186-4008).
2. DG8 Gaskets for QX200 Droplet Generator (Cat# 186-3009).
3. Pierceable Foil Heat Seal (Cat# 1814040).
4. QX200 Droplet Reader Oil (Cat# 186-3004).

### **2.3 ddPCR Reagents for Probe Assay**

1. ddPCR Supermix for Probes (no dUTP) (Bio-Rad Cat# 186-3024).
2. ddPCR 2× Buffer Control Kit for Probe (Bio-Rad Cat# 186-3052).
3. Droplet Generation Oil for Probes (Bio-Rad Cat#186-3005).
4. Primers for target and control amplicons (e.g., IDT).
5. FAM-VIC or FAM-HEX probes for target and reference amplicons (e.g., IDT 6'Fluro FAM/HEX, 3'Quencher ZEN).

### **2.4 ddPCR Reagents for EvaGreen Assay**

1. QX200 ddPCR EvaGreen Supermix (Bio-Rad Cat# 186-4034).
2. QX200 2× Buffer Control for EvaGreen (Bio-Rad Cat# 186-4052).
3. Droplet Generation Oil for EvaGreen store at room temperature (Bio-Rad Cat# 186-4005).
4. Primers for target and reference control amplicons (e.g., IDT).

### **2.5 Other Consumables**

1. Twin Tec semiskirted 96-well plate (Eppendorf Cat# 951020362).
2. 25 mL Reagent trough (Thermo Fisher).
3. P20 barrier pipette tips (Rainin Only, gentler for pipetting droplets Cat# GP-20F).
4. P200 barrier pipette tips (Rainin Only, gentler for pipetting droplets Cat# GP-200F).
5. DNA restriction enzyme compatible with assay (e.g., New England Biolabs).
6. Molecular biology grade water.

### **2.6 Genomic DNA Isolation**

---

## **3 Methods**

### **3.1 Primer and Probe Design**

#### **3.1.1 General Considerations**

The design of the ddPCR primers and probes should follow these general guidelines. For primers,  $T_m$  should be 50–65 °C with  $\leq 5$  °C difference between both primers. The primer length should be 9–40 bp (ideally ~ 20 bp) and the amplicon size should be 60–200 bp. The primer GC percentage should be between 50% and 60%. In addition, these features should be avoided in primer design:

1. Stretches of 4 or more C or G;
2. Hairpin structures ( $T_m$  of the hairpin should be lower than the annealing temperature);

3. Primer dimers ( $\Delta G$  for dimer analysis should be between 0 and  $-9\text{ kcal/mol}$ );
4. Nonspecific priming (check with *in silico* PCR available on UCSC Genome Browser).

To design TaqMan probes for ddPCR probe assays, the probe should be 15–30 bp in length with GC content between 30% and 80%. Probe  $T_m$  (noninclusive of quencher effect) should be  $\sim 10\text{ }^\circ\text{C}$  higher than the primer  $T_m$ . If using a 3'-minor groove binder (MGB) the MGB increases probe  $T_m$  by 15–30  $^\circ\text{C}$ , which allows for the use of shorter probes with high specificity. MGB  $T_m$  enhancer can be obtained from ABI or Thermo Fisher. Target and reference probe 5' fluorophores should be either FAM and VIC (“ABI”) respectively or FAM and HEX respectively (IDT). Probes should also have a quencher group such as ZEN (IDT), NFQ or black hole (on the 3'-end). The probe sequence can be based off of either the Watson or Crick strand; it is preferable to use the strand that places more Cs than Gs in the probe sequence, as long as the sequence does not start with a G at the 5'-end.

### 3.1.2 Reference Primer and Probe Design

Primers that target an amplicon region within the human *RPP30* gene, for which two copies should be present in each cell, can serve as reference primers. VIC or HEX fluorophore-conjugated Taq-Man probe hybridizing this amplicon can serve as the reference probe (reference *RPP30* probes and primers can be also be supplied by Bio-Rad (<https://www.bio-rad.com/digital-assays/#/>)).

### 3.1.3 Mosaic Variant Primer and Probe Design

Primers should be designed to target a variant such that the amplicon contains the SNV or the breakpoint-junction or insertion-point sequence of the CNV or MEI, respectively. FAM probes should be designed to hybridize directly onto the predicted SNV (variant sequence) or the breakpoint-junction or insertion-point sequence of the predicted CNV/MEI (Fig. 1) and assayed with *RPP30* as the reference target. When using single-color EvaGreen intercalating dye for the analysis of somatic variants, it is important to consider that the signal amplitude differentiation between different targets comes from the differences in the lengths of the amplicons (i.e., reference and target amplicon); therefore, target amplicons and reference amplicons must be sufficiently different in length to yield clear signal differentiation [6, 12]. For instance, mixing primers that target a mosaic CNV yielding amplicons 200 bp in size and primers that target a reference region (60 bp) within *RPP30* in the same reaction may result in clear signal separation, which is not expected if both target and reference amplicons are similar in lengths. In certain instances, clusters may be sufficiently discriminated by as little as 6 bp [6]. The frequency of the target CNV can

be reliably quantified by comparing the groups of droplets (target, reference, and null) with distinct EvaGreen intensities. The optimal annealing and extension temperature should be examined for each target and reference primer set by testing the separation of EvaGreen signal intensities from the target-variant and reference-*RPP30* amplicons. If the mosaic event is novel and rare, one can perform positive and negative control experiments to ensure accurate interpretation of ddPCR results. DNA (synthesized, cloned, or isolated) containing the variant of interest may be diluted to desired frequency to mimic mosaic events and mixed with template DNA to serve as positive control. Synthesized DNA fragments may for example be purchased as gBlocks from Integrated DNA Technologies (IDT).

### **3.2 Genomic DNA Isolation**

Use nondegraded genomic DNA for all ddPCR assays. Perform genomic DNA extraction using QIAGEN DNeasy Blood & Tissue kit (Cat. 69504) following manufacturer guidelines or any other genomic DNA isolation method of choice as long as genomic DNA of high quality (nondegraded) and high purity can be obtained.

### **3.3 DNA Digestion**

It is essential to load restriction enzyme in the reaction mix if more than 66 ng of genomic DNA per reaction is used. However, it can also be very important for assaying a genomic target region that is poorly accessible, thus restriction enzyme at 10–20 U/1 µg can be loaded for such cases. A four-cutter enzyme such as *MseI* will fragment the DNA every 250 base pairs, on average. Choose a digestion enzyme that does not cut in either reference or target amplicon by entering the amplicon sequence into NEBCutter (<http://nc2.neb.com/NEBcutter2/>). Restriction enzyme may be directly loaded into the ddPCR reaction. One may also choose to digest template genomic DNA prior to loading into ddPCR reaction. However, the user should be aware that the predigested sample must be diluted at least tenfold in the final ddPCR reaction to avoid any inhibition of the reaction. This approach requires significantly more DNA in situ restriction digest.

### **3.4 ddPCR Experimental Setup**

1. For each targeted SNV, CNV, or MEI ddPCR reaction the following should be prepared at room temperature.

Probe assay:

- (a) 10 µL 2× ddPCR Probe Master Mix.
- (b) 1.8 µL 10 µM Forward Target Primer.
- (c) 1.8 µL 10 µM Reverse Target Primer.
- (d) 0.5 µL 10 µM Target Probe.
- (e) 1.8 µL 10 µM Forward Reference Primer.

- (f) 1.8  $\mu$ L 10  $\mu$ M Reverse Reference Primer.
- (g) 0.5  $\mu$ L 10  $\mu$ M Reference Probe.
- (h) 20 ng (5–66 ng) of template DNA (volume variable).
- (i) 0.5  $\mu$ L DNA restriction enzyme (if more than 66 ng of template DNA is used).
- (j) Nuclease-free water to make total volume 20  $\mu$ L).

EvaGreen assay (*see Note 4*):

- (a) 10  $\mu$ L 2 $\times$  ddPCR EvaGreen Master Mix.
- (b) 0.2  $\mu$ L 10  $\mu$ M Forward Target Primer.
- (c) 0.2  $\mu$ L 10  $\mu$ M Reverse Target Primer.
- (d) 0.2  $\mu$ L 10  $\mu$ M Forward Reference Primer.
- (e) 0.2  $\mu$ L 10  $\mu$ M Reverse Reference Primer.
- (f) 20 ng (5–66 ng) of template DNA (volume variable).
- (g) 0.5  $\mu$ L DNA restriction enzyme (if more than 66 ng of template DNA is used).
- (h) Nuclease-free water to make total volume 20  $\mu$ L).

2. Mix ddPCR solution by slowly pipetting up and down 10 times.
3. Place 8-well DG8 cartridge into cartridge holder.
4. Slowly load 20  $\mu$ L of PCR solution into the middle row of wells in the DG8 cartridge using eight-channel multichannel pipette. Ensure that there are no air bubbles in the bottom of the wells. If air bubbles are introduced, remove with an extra clean pipette tip.
5. Load 70  $\mu$ L of droplet generation oil in the indicated wells in the DG8 cartridge, again using an eight-channel multipipette. Droplet generation should begin within 2' after oil addition.
6. Attach gasket to the cartridge holder.
7. Place cartridge in droplet generator and press button to close the droplet generator lid and begin droplet generation.
8. Wait until droplet generation is complete (~2').
9. After droplet generation is complete, the droplet mixture should appear cloudy.
10. Carefully transfer each ddPCR droplet mixture (~38  $\mu$ L) to 96-well plate using an eight-channel multipipette.
11. Seal plate with foil using plate sealer.
12. Place 96-well plate in thermocycler and run the following protocols. (Note: annealing temperatures may vary depending on primer–probe sets used).

Probe assay:

95 °C 10 min.

94 °C 30 s, 60 °C 60 s (40 cycles).

98 °C 10 min.

4 °C hold.

EvaGreen assay:

95 °C 5 min.

95 °C 30 s, 60 °C 60 s (40 cycles).

4 °C 5 min.

90 °C 5 min.

4 °C hold.

13. Once thermocycler protocol is completed, place 96-well plate in droplet reader.
14. Read assay and analyze results using the QuantaSoft Software, choosing appropriate droplet reading parameters for probes or EvaGreen chemistries.

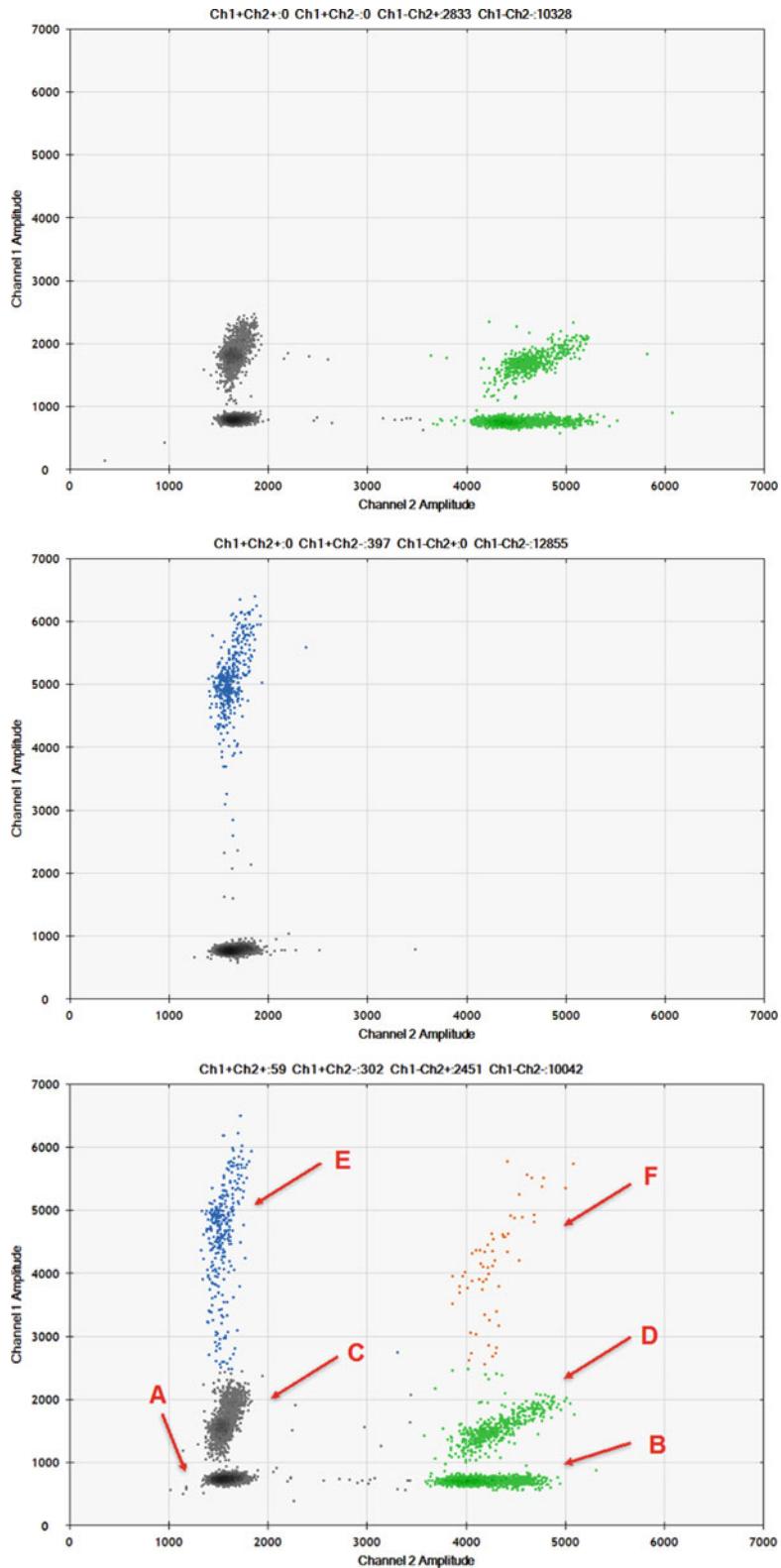
### **3.5 Mosaic SNV Analysis**

The equations (Table 1) used for estimating the SNV allele frequency ( $\hat{r}$ ) and its 95% confidence interval ( $\hat{r}_{\text{Low}}$  and  $\hat{r}_{\text{High}}$ ) are adopted from [9] where the principles are discussed in more detail. In theory, these principles can be applied for both probe-based and EvaGreen assays. Three input values are needed: the total number of droplets generated ( $C$ ), the number of droplets that contain the SNV of interest ( $H_1$ ), and the number of droplets that contain the reference allele ( $H_2$ ). All three input values can be obtained from the Bio-Rad QuantaSoft software readout of the ddPCR reaction:  $C$  = “Accepted Droplets,”  $H_1$  = “Positives” under Chanel 1,  $H_2$  = “Positives” under Chanel 2. If replicates are performed, then droplet counts should be combined. It is generally encouraged to perform more ddPCR replicates for low-frequency SNVs to achieve more precise quantification (see Note 2). In the example shown in Fig. 3,  $C$  is the sum of all droplets which is 12,854;  $H_1$  is be the sum of the droplets contained in clusters E and F which is 367.  $H_2$  is be the sum of the droplets contained in the green cluster and clusters B, D, and F which is 2516. These three values will subsequently be used as input values for equations listed in Table 1 to calculate allele frequency. Since SNVs only contain a single nucleotide difference from the nonvariant allele, the TaqMan probe designed against the variant allele of interest will also hybridize with the nonvariant allele albeit at a lower amplitude (Fig. 3 bottom, cluster C). It is strongly encouraged to perform the SNV ddPCR assay on positive and negative control templates to ascertain the optimal ddPCR conditions, primer/probe concentrations,

**Table 1**  
**Variant frequency calculation**

<i>Input values</i>
Total number of droplets: $C$
Total number of droplets containing the variant: $H_1$
Total number of droplets containing the reference: $H_2$
<i>Output values</i>
Variant frequency: $\hat{r}$
Lower bound of the 95% confidence interval of $\hat{r}$ : $\hat{r}_{\text{Low}}$
Upper bound of the 95% confidence interval of $\hat{r}$ : $\hat{r}_{\text{High}}$
Positive fractions for variant $\hat{p}_1$ and reference $\hat{p}_2$ :
$\hat{p}_1 = \frac{H_1}{C} \quad \hat{p}_2 = \frac{H_2}{C}$
Standard deviations for variant $S_1$ and reference $S_2$ :
$S_1 = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{C}} \quad S_2 = \sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{C}}$
$\hat{p}_{1,\text{Low}} = \hat{p}_1 - 1.96S_1 \quad \hat{p}_{1,\text{High}} = \hat{p}_1 + 1.96S_1$
$\hat{p}_{2,\text{Low}} = \hat{p}_2 - 1.96S_2 \quad \hat{p}_{2,\text{High}} = \hat{p}_2 + 1.96S_2$
Average number of positive events per droplet for variant $\hat{\lambda}_1$ and reference $\hat{\lambda}_2$ :
$\hat{\lambda}_1 = -\ln(1 - \hat{p}_1) \quad \hat{\lambda}_{1,\text{Low}} = -\ln(1 - \hat{p}_{1,\text{Low}}) \quad \hat{\lambda}_{1,\text{High}} = -\ln(1 - \hat{p}_{1,\text{High}})$
$\hat{\lambda}_2 = -\ln(1 - \hat{p}_2) \quad \hat{\lambda}_{2,\text{Low}} = -\ln(1 - \hat{p}_{2,\text{Low}}) \quad \hat{\lambda}_{2,\text{High}} = -\ln(1 - \hat{p}_{2,\text{High}})$
$H_{\text{Top}} = \hat{\lambda}_{1,\text{High}} - \hat{\lambda}_1 \quad H_{\text{Bottom}} = \hat{\lambda}_1 - \hat{\lambda}_{1,\text{Low}}$
$W_{\text{Right}} = \hat{\lambda}_{2,\text{High}} - \hat{\lambda}_2 \quad W_{\text{Left}} = \hat{\lambda}_2 - \hat{\lambda}_{2,\text{Low}}$
$\hat{r} = \frac{\hat{\lambda}_1}{\hat{\lambda}_2}$
$\hat{r}_{\text{Low}} = \frac{\hat{\lambda}_1 \hat{\lambda}_2 - \sqrt{\hat{\lambda}_1^2 \hat{\lambda}_2^2 - (H_{\text{Bottom}}^2 - \hat{\lambda}_1^2)(W_{\text{Right}}^2 - \hat{\lambda}_2^2)}}{\hat{\lambda}_2^2 - W_{\text{Right}}^2}$
$\hat{r}_{\text{High}} = \frac{\hat{\lambda}_1 \hat{\lambda}_2 - \sqrt{\hat{\lambda}_1^2 \hat{\lambda}_2^2 - (H_{\text{Top}}^2 - \hat{\lambda}_1^2)(W_{\text{Left}}^2 - \hat{\lambda}_2^2)}}{\hat{\lambda}_2^2 - W_{\text{Left}}^2}$

and number of ddPCR cycles before performing the assay on the sample. As positive control, template DNA that contains the variant of interest may be diluted to desired allele frequency to mimic mosaic events (*see Note 3*). The optimal ddPCR conditions for SNV assays are generally those in which the droplets that contain



the SNV allele has the widest signal (amplitude) separation from those that contain the nonvariant allele only (cluster C, in the example shown in Fig. 3).

### **3.6 Mosaic MEI Analysis**

The mosaic allele quantification principles as outlined for mosaic SNV analysis under Subheading 3.4 may be applied for mosaic MEI analysis, where  $H_1$  is the total number of droplets that contain the MEI allele of interest in the sample and, as above,  $H_2$  is the total number of droplets that contain the reference allele. An example of the sensitivity of ddPCR in the detection and quantification of Line1 mobile genomic elements using different fluorophore options is illustrated in Fig. 4.

### **3.7 Mosaic CNV Analysis**

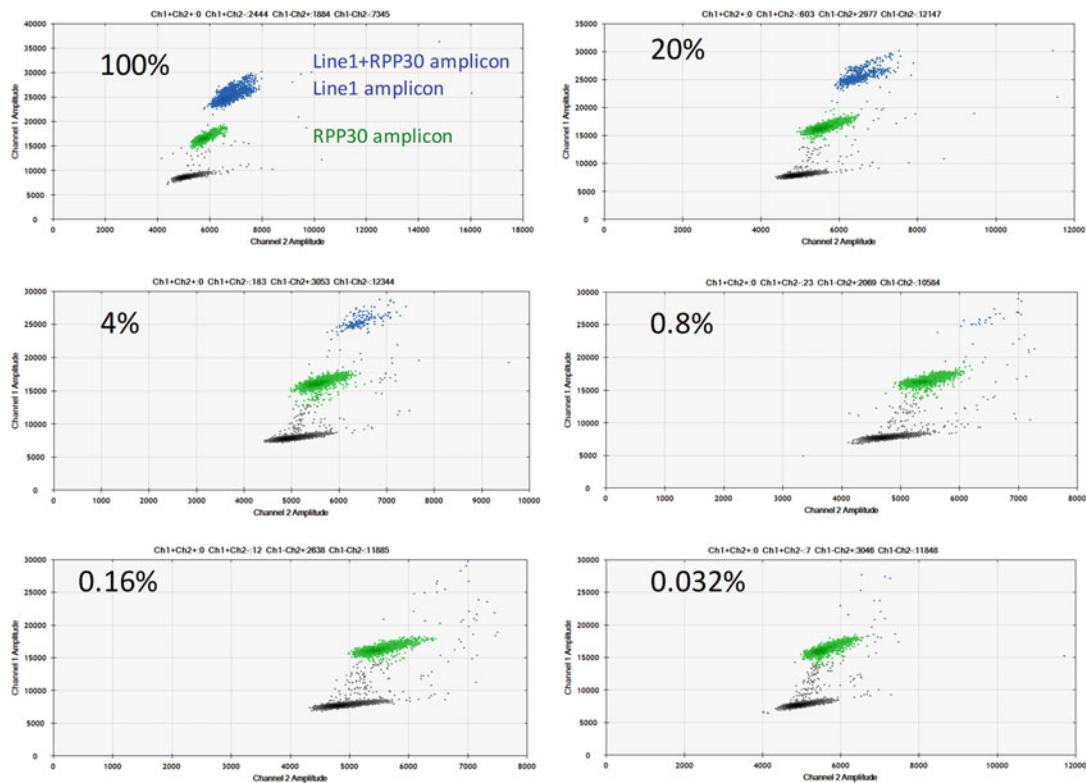
In general, assuming that the ddPCR primers and/or TaqMan probe span the breakpoint of the CNV of interest, the mosaic allele quantification principles as outlined for mosaic SNV analysis under Subheading 3.4 may be applied for mosaic CNV analysis (Fig. 5), where  $H_1$  is the total number of droplets that contain the target CNV allele in the sample, and  $H_2$  is the total number of droplets with the reference allele.

## **4 Notes**

1. The accuracy of mosaic frequency quantification depends mostly on the sample size (or abundance) of the target of interest as well as the total number of droplets. As long as the ddPCR assay readout gives clear signal separation, EvaGreen and probe assays should be equally reliable. EvaGreen assays are significantly cheaper per reaction as no specific fluorophore-conjugated TaqMan probe has to be synthesized, but they may require longer time to optimize for each primer pair as three or more clusters of droplets need to clearly separate in a single channel amplitude in order for quantification to proceed.

**Fig. 3** Quantifying the frequency of SNV in human genomic DNA using TaqMan Probe Assay. In order to mimic template DNA that contains a mosaic SNV on chromosome 10, copies of synthesized double-stranded DNA (500 bp) that contain the SNV at the locus of interest were added to human genomic DNA at ~16% allele frequency. Data shown are from a single ddPCR well. The target TaqMan probe (Channel 1) designed for this assay has a perfect sequence match to the variant allele and contains one base pair mismatch to the nonvariant allele. A probe targeting the RNaseP gene *RPP30* was used as the reference allele (Channel 2). (Top) Negative control assay with NA12878 genomic DNA as template DNA in which only *RPP30* is present. (Middle) Positive control assay with only synthesized double-stranded DNA as template, and Channel 1 positive droplets appear at a higher amplitude cluster compared to the negative control assay. (Bottom) Six separate characteristic clusters of droplets were observed in this assay: no template (**A**), *RPP30* only (**B**), nonvariant allele only (**C**), *RPP30* + nonvariant allele (**D**), SNV only (**E**), and SNV + *RPP30* (**F**)

## Detection of Line-1 insertion with Evagreen

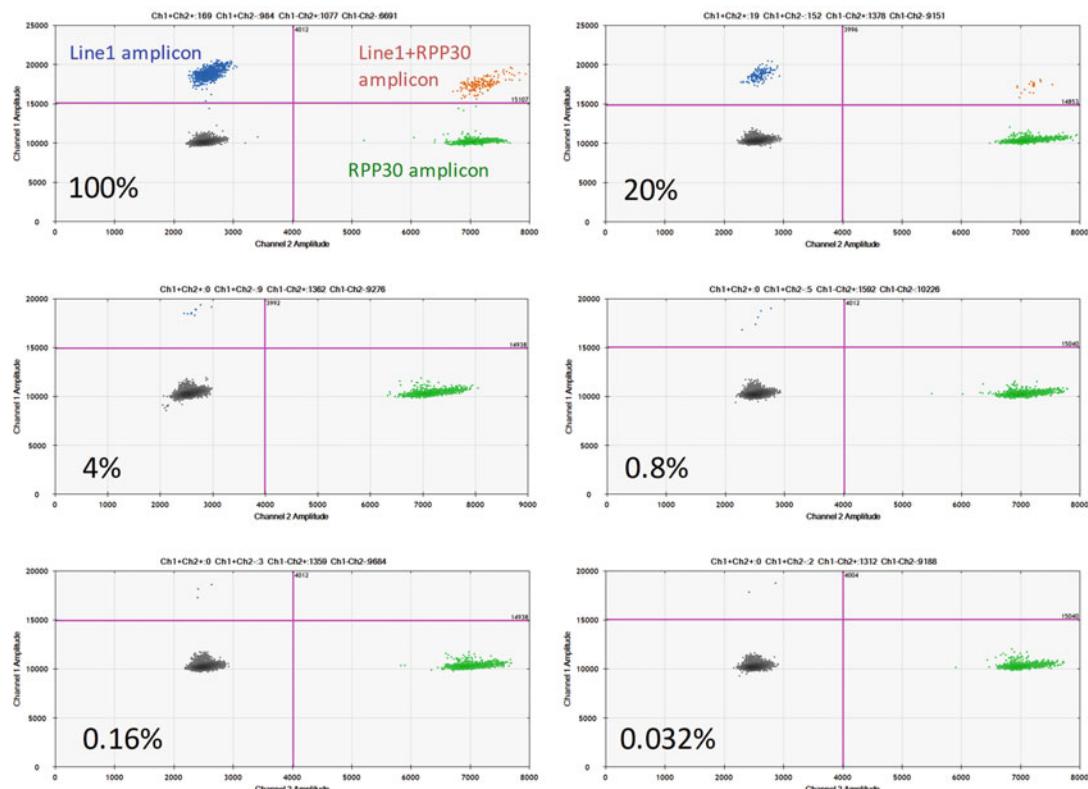


**Fig. 4** Detection of mosaic Line1 insertions with different ddPCR approaches. To test the power of ddPCR to detect highly mosaic Line1 insertions we mixed different levels of GM12878 genomic DNA, containing a rare homozygous insertion of Line1 with precisely known boundaries, and genomic DNA from a cell line that does not contain the insertion (RUID#04C3729). With a total amount of 12 ng genomic DNA, which contains 100%, 20%, 4%, 0.8%, and 0.16% GM12878 DNA, respectively, we tested the performance of three ddPCR protocols: EvaGreen (top panel), TaqMan with FAM/HEX probes (middle panel) and TaqMan with FAM/VIC probes (bottom panel). Overall all three protocols, showed similar very high performance in quantifying MEIs that are as rare as being present in only ~0.1% of the input DNA

Probe-based assays have the advantage of dual channel readout and the additional target specificity introduced by the TaqMan probe. In certain cases, TaqMan probes will be difficult to design; in these cases, only the EvaGreen assay is the only available option. It is also possible that for certain variants, especially small indels and SNVs, the additional sequence specificity conferred by the TaqMan probe is needed.

2. Generally, the more replicate reactions performed in the validation of a variant, the greater the precision of estimation of mosaic level. It is recommended that for low-frequency events additional replicates be performed, instead of overloading a

### Detection of Line-1 insertion with FAM/HEX probe

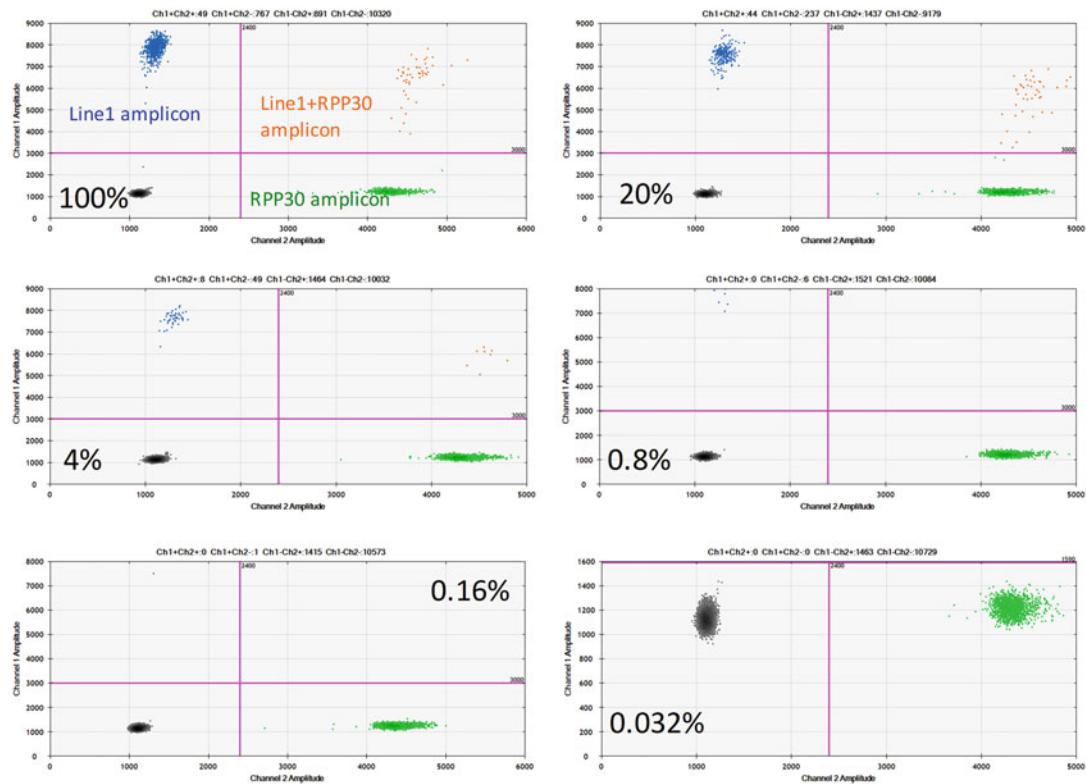


**Fig. 4 (continued)**

reaction with too much template DNA. Always load <5 cpd or <1.6 cpd for maximal precision. Too much template DNA (>5 cpd) could adversely affect droplet formation and also eliminates the number of negative droplets such that Poisson correction of copy numbers cannot be performed leading to inaccurate quantification.

3. It is strongly encouraged to perform the SNV ddPCR assay on positive and negative control templates to first ascertain the optimal ddPCR conditions, primer/probe concentrations, and number of ddPCR cycles before performing the assay on the sample. As a positive control, template DNA that contains the variant of interest may be diluted to desired allele frequency to mimic mosaic events. The optimal ddPCR conditions for SNV assays are generally those in which the droplets that contain the SNV allele has the widest signal (amplitude) separation with those that contain the nonvariant allele only or clusters A and C in the example shown in Fig. 3.

### Detection of Line-1 insertion with FAM/VIC probe



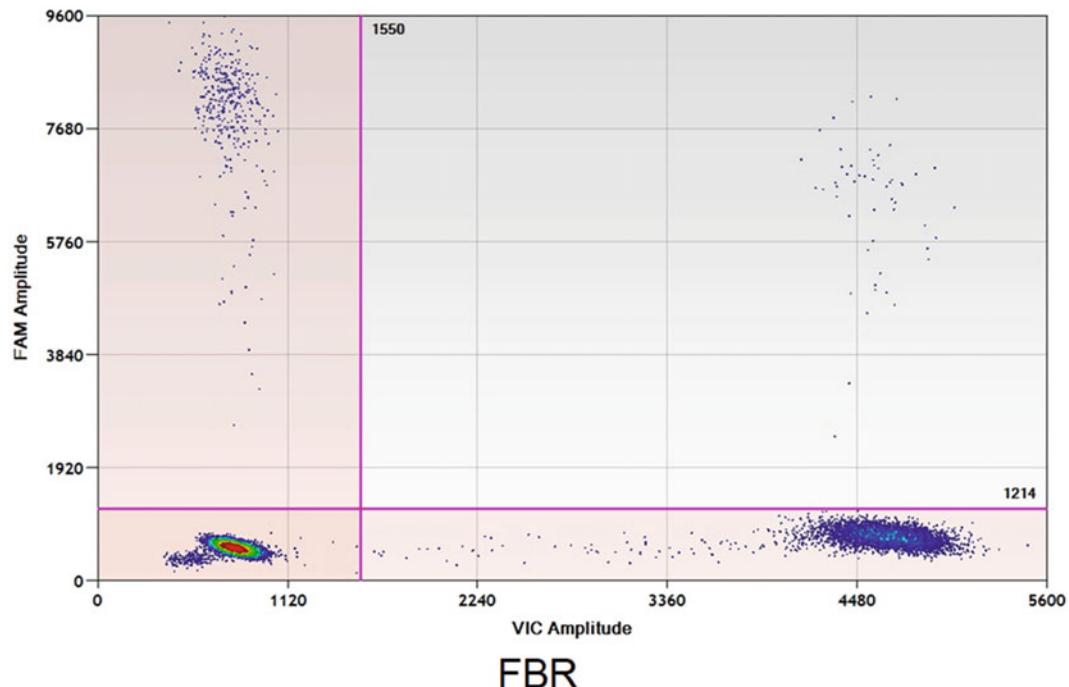
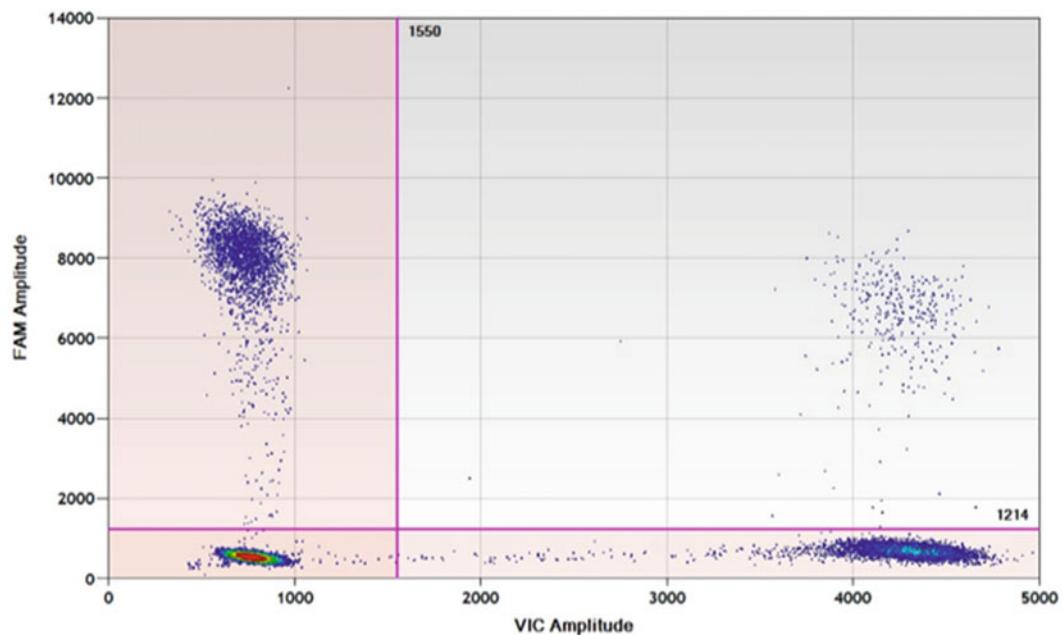
**Fig. 4** (continued)

4. Do not use any intercalating fluorophores other than EvaGreen (e.g., do not use SYBR Green). Intercalating dyes that have not been tested on and approved for the Bio-Rad ddPCR Droplet Reader may potentially leak out of the droplet and damaged the optics or other parts of the machine.

---

## Acknowledgments

Research reported in this chapter was supported by the National Institute of Mental Health of the National Institutes of Health under award numbers R01MH094740 and R01MH100914. We also acknowledge additional funds from Stanford University (Department of Psychiatry and Behavioral Sciences, and Department of Genetics).

**A****FBR****B****hiPSC 7**

**Fig. 5** Detection of mosaic CNV in fibroblast line (A) that is clonally expanded in hiPSC line (B). VIC fluorescent event counts measure copy number of *RPP30* reference gene while FAM fluorescent event counts measure copy number of mosaic CNV by designing the FAM probe such that it spans the CNV breakpoint junction. In this sample, the heterozygous CNV deletion was present in ~10% of fibroblast cells, while the deletion was present in ~100% of the hiPSC 7 cells

## References

1. Abyzov A, Mariani J, Palejev D, Zhang Y, Haney MS, Tomasini L, Ferrandino AF, Rosenberg Belmaker LA, Szekely A, Wilson M, Kocabas A, Calixto NE, Grigorenko EL, Huttner A, Chawarska K, Weissman S, Urban AE, Gerstein M, Vaccarino FM (2012) Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. *Nature* 492:438–442. <https://doi.org/10.1038/nature11629>
2. Campbell IM, Yuan B, Robberecht C, Pfundt R, Szafranski P, McEntagart ME, Nagamani SC, Erez A, Bartnik M, Wisniewiecka-Kowalnik B, Plunkett KS, Pursley AN, Kang SH, Bi W, Lalani SR, Bacino CA, Vast M, Marks K, Patton M, Olofsson P, Patel A, Veltman JA, Cheung SW, Shaw CA, Vissers LE, Vermeesch JR, Lupski JR, Stankiewicz P (2014) Parental somatic mosaicism is under-recognized and influences recurrence risk of genomic disorders. *Am J Hum Genet* 95:173–182. <https://doi.org/10.1016/j.ajhg.2014.07.003>
3. Evrony GD, Lee E, Mehta BK, Benjamini Y, Johnson RM, Cai X, Yang L, Haseley P, Lehmann HS, Park PJ, Walsh CA (2015) Cell lineage analysis in human brain using endogenous retroelements. *Neuron* 85:49–59. <https://doi.org/10.1016/j.neuron.2014.12.028>
4. Edwards JH (1989) Familiarity, recessivity and germline mosaicism. *Ann Hum Genet* 53:33–47
5. Freed D, Stevens EL, Pevsner J (2014) Somatic mosaicism in the human genome. *Genes (Basel)* 5:1064–1094. <https://doi.org/10.3390/genes5041064>
6. Miotke L, Lau BT, Rumma RT, Ji HP (2014) High sensitivity detection and quantitation of DNA copy number and single nucleotide variants with single color droplet digital PCR. *Anal Chem* 86:2618–2624. <https://doi.org/10.1021/ac403843j>
7. Hindson CM, Chevillet JR, Briggs HA, Galli-chotte EN, Ruf IK, Hindson BJ, Vessella RL, Tewari M (2013) Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat Methods* 10:1003–1005. <https://doi.org/10.1038/nmeth.2633>
8. Pinheiro LB, Coleman VA, Hindson CM, Herrmann J, Hindson BJ, Bhat S, Emslie KR (2012) Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Anal Chem* 84:1003–1011. <https://doi.org/10.1021/ac202578x>
9. Dube S, Qin J, Ramakrishnan R (2008) Mathematical analysis of copy number variation in a DNA sample using digital PCR on a nanofluidic device. *PLoS One* 3:e2876. <https://doi.org/10.1371/journal.pone.0002876>
10. Strain MC, Lada SM, Luong T, Rought SE, Gianella S, Terry VH, Spina CA, Woelk CH, Richman DD (2013) Highly precise measurement of HIV DNA by droplet digital PCR. *PLoS One* 8:e55943. <https://doi.org/10.1371/journal.pone.0055943>
11. Abyzov A, Urban AE, Snyder M, Gerstein M (2011) CNVnator: An approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res* 21(6):974–984
12. McDermott GP, Do D, Litterst CM, Maar D, Hindson CM, Steenblock ER, Legler TC, Jouvenot Y, Marrs SH, Bemis A, Shah P, Wong J, Wang S, Sally D, Javier L, Dinio T, Han C, Brackbill TP, Hodges SP, Ling Y, Klitgord N, Carman GJ, Berman JR, Koehler RT, Hiddessen AL, Walse P, Bousse L, Tzonev S, Hefner E, Hindson BJ, Cauly TH 3rd, Hamby K, Patel VP, Regan JF, Wyatt PW, Karlin-Neumann GA, Stumbo DP, Lowe AJ (2013) Multiplexed target detection using DNA-binding dye chemistry in droplet digital PCR. *Anal Chem* 85:11619–11627. <https://doi.org/10.1021/ac403061n>

# **Part IV**

## **Rare Mutation and Rare Allele Detection**



# Chapter 12

## Monitoring of Response and Resistance in Plasma of *EGFR*-Mutant Lung Cancer Using Droplet Digital PCR

**Yanan Kuang, Allison O'Connell, Adrian G. Sacher, Nora Feeney, Ryan S. Alden, Geoffrey R. Oxnard, and Cloud P. Paweletz**

### Abstract

The identification of oncogenic driver mutations has led to the rapid rise of genotype-directed treatments. However, genetic analysis of tumors remains cumbersome and a morbid experience for patients. Noninvasive assessment of tumor genotype, so-called “liquid biopsy,” such as plasma genotyping represents a potentially transformative tool. Here we describe a genotyping protocol of cell-free plasma DNA (cfDNA) using Droplet Digital™ PCR (ddPCR™). ddPCR emulsifies DNA into ~20,000 droplets in which PCR is performed to endpoint in each droplet for both mutant and wild-type DNA. Droplets are run through a modified flow cytometer where mutant and wild-type DNA emit different colored signals. The count of these signals upon Poisson distribution analysis allows sensitive quantification of allelic prevalence.

**Key words** Droplet Digital™ PCR, Cell-free DNA, Liquid biopsy, Genotyping, Non-small-cell lung cancer, Resistance mutations, *EGFR*

---

### 1 Introduction

Genotype-directed therapy is now the standard of care for advanced non-small-cell lung cancer (NSCLC) patients. Somatic mutations in *EGFR*, *ALK*, *ROS1*, and *BRAF* had been found to be predictors of the efficacy of kinase inhibitors and are commonly used in either clinical practice or in clinical trials [1–4]. However, despite the therapeutic success of this approach, not all patients have sufficient tumor tissue available for genotyping. In addition, acquired drug resistance ultimately develops in all patients successfully treated with kinase inhibitors [5]. The identification of treatment-related changes in such cancers is of increasing importance given the development of new targeted therapies specifically for patients with treatment resistance [6, 7].

Repeated tumor biopsy is the gold standard for evaluating genetic changes in lung cancers following treatment with targeted therapies. However, this is not always feasible and can seldom be

performed more than once. Noninvasive techniques for blood-based tumor genotyping (“liquid biopsies”) may realize the potential of genotype-directed lung cancer therapy. Recent studies have suggested that highly sensitive genotyping assays can detect mutations in circulating cell-free plasma DNA (cfDNA) from cancer patients, potentially representing the biology of a patient’s cancer [8–10].

We have now developed [11] a new assay for noninvasive quantitative genotyping of oncogenic mutations in cfDNA of NSCLC patients using Droplet Digital PCR™ (ddPCR™). This technology emulsifies input DNA into ~20,000 droplets so that the proportion carrying mutant versus wild-type alleles can be counted using flow cytometry. We have now used this method to *serially* analyze cfDNA from EGFR-mutant NSCLC patients treated with the EGFR kinase inhibitor, erlotinib. We can identify the presence of the EGFR activating mutation prior to therapy, its reduction or disappearance during therapy and reemergence along with the drug resistance EGFR T790M mutation during treatment [11]. The present chapter guides the reader through several types of ddPCR™ assay designs for plasma derived cfDNA and its clinical application.

---

## 2 Materials

### 2.1 Cell-Free DNA Extraction

1. EDTA lavender-capped vacutainer tubes (Becton and Dickinson #366643).
2. QIAamp circulating nucleic acid kit (Qiagen #551140).
3. Vacuum pump (Qiagen #84010).
4. QIAvac connecting system (Qiagen #19419).
5. QIAvac 24 Plus (Qiagen #19413).
6. 15 mL polypropylene conical tube (Thermo Scientific #339650).
7. 50 mL polypropylene conical tube (Thermo Scientific #339652).
8. 2 mL cryo vial (Thermo Scientific #4000200).
9. Phosphate buffered saline (Thermo Scientific #10010023).
10. Benchtop Vortex mixer.

### 2.2 Droplet Digital PCR

See Note 1.

1. DG8™ Cartridges for QX200™/QX100™ Droplet Generator (Bio-Rad #1864008).
2. DG8™ Gaskets for QX200™/QX100™ Droplet Generator (Bio-Rad #1863009).

**Table 1**  
**Cell lines used as positive controls**

Mutation	Cell line
EGFR del 19	PC9
EGFR L858R	H1975
EGFR T790M	PC9 GR
EGFR wild type	A549

3. DG8™ Cartridge Holder (Bio-Rad #1863051).
4. Droplet Generation Oil for Probes (Bio-Rad #1863005).
5. ddPCR™ Supermix for probes (No dUTP) (Bio-Rad #186-3025).
6. ddPCR™ Droplet Reader Oil (Bio-Rad #1863004).
7. Semiskirted 96-well PCR plate (Thermo Scientific #E951020346).
8. Pierceable Foil Heat Seal (Bio-Rad #1814040).
9. Rainin single and multichannel pipettors (P2, P10, P20, P1000; multi-P50 and multi-P100).
10. Sharp Precision DNase- and RNase-free filter pipette tips (Denville Scientific: #P1096-FR, #P1121, #P1122, and #P1126).
11. Positive DNA controls: High, medium, or low levels of genomic DNA from *EGFR* mutant cell lines (*see* Table 1).
  - (a) High QC: 1000 mutant copies in background of 5000 wt.
  - (b) Medium QC: 100 mutant copies in background of 5000wt.
  - (c) Low QC: 10 mutant copies in background of 5000 wt.
12. Primer and probes are custom ordered from Life Technologies. For good assay design principles we refer the reader to Chapter 2 of Bio-Rad's ddPCR™ application guide [13].
  - (a) *EGFR* L858R primer sequences:  
 Forward: 5'-GCAGCATGTCAAGATCACAGATT-3'  
 Reverse: 5'-CCTCCTTCTGCATGGTATTCTTCT-3'  
*EGFR* L858R probe sequences:  
 5'-VIC-AGTTTGGCCAGCCAA-MGB-NFQ-3'  
 5'-FAM-AGTTTGGCCCGCCAA-MGB-NFQ-3'
  - (b) *EGFR* exon19 deletion primer sequences:  
 Forward: 5'-GTGAGAAAGTTAAAATTCCCGTC-3'  
 Reverse: 5'-CACACAGCAAAGCAGAAC-3'  
*EGFR* exon19 deletion probe sequences are:

5'-VIC-ATCGAGGATTCCTTGTG-MGB-NFQ-3'  
 5'-FAM-AGGAATTAAGAGAACATC-MGB-  
 NFQ-3'

(c) *EGFR* T790M primer sequences:

Forward: 5'-GCCTGCTGGGCATCTG-3'

Reverse: 5'-TCTTGTTGCCGGACATAGTC-3'

*EGFR* T790M probe sequences:

5'-VIC-ATGAGCTGCGTGATGAG-MGB-NFQ-3'

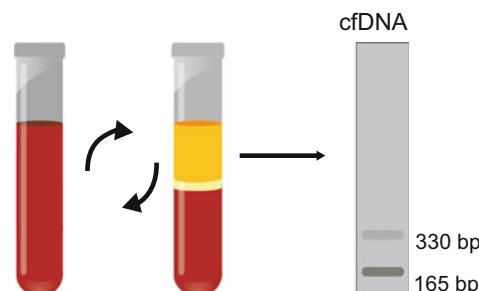
5'-FAM-ATGAGCTGCATGATGAG-MGB-NFQ-3'

### 3 Methods

#### 3.1 Plasma Isolation

See Notes 2–5.

1. Draw venous blood into one [1] 10 mL EDTA lavender-capped vacutainer tube and immediately invert the tube eight to ten times.
2. Immediately centrifuge for 10 min at 1500 ( $\pm 150$ )  $\times g$  in a swinging bucket centrifuge.
3. Pipette the ~2–3 mL plasma layer (upper layer) into a sterile 15 mL polypropylene conical tube (Fig. 1).
4. Centrifuge the 15 mL tube containing the plasma for an additional 10 min at 3000 ( $\pm 150$ )  $\times g$  in a swinging bucket centrifuge. This step removes any remaining white blood cells.
5. Transfer, using a fresh pipette, the supernatant into a second 15 mL polypropylene tube. Leave about 0.3 mL of supernatant in the tube. This leftover contains cellular debris.
6. Using a fresh pipette aliquot plasma into 2 mL cryovials. One can expect ~2–3 mL of pale yellow plasma from a full EDTA vacutainer.
7. Freeze immediately upright at  $-70^{\circ}\text{C}$  or colder until use.



**Fig. 1** Plasma is isolated by whole blood centrifugation. cfDNA fragments range between 135 and 480 base pairs

### 3.2 cfDNA Extraction

#### See Notes 6–9.

This procedure is adapted from the QIAamp circulating nucleic acid handbook. All buffers and enzymes are provided in the QIAamp Circulating Nucleic Acid Kit, but require the addition of 100% ethanol and isopropanol as specified by the manufacturer.

Before you start:

- (a) Prepare all buffers.
  - (b) Wipe down the lab bench, hood and pipets with 10% bleach followed by 70% ethanol.
  - (c) Heat a water bath or heating block for 15 mL conical tubes to 60 °C and a heating block for 1.5 mL microcentrifuge tubes to 56 °C.
  - (d) Set up the QIAvac24 Plus manifold.
  - (e) Thaw 2 mL of plasma to room temperature.
1. Pipet 200 µL QIAGEN Proteinase K into a 15 mL conical tube.
  2. Add 2 mL of plasma into the tube.
  3. Add 1.6 mL Buffer ACL (1.0 µg carrier RNA per plasma sample). Close the cap and mix by pulse-vortexing for 30 s.
  4. Incubate in a 60 °C water bath or heating block for 30 min.
  5. Place the tube back on the lab bench and unscrew the cap.
  6. Add 3.6 mL Buffer ACB to the lysate in the tube. Close the cap and mix thoroughly by pulse-vortexing for 15–30 s.
  7. Incubate the lysate–Buffer ACB mixture in the tube for 5 min on ice.
  8. Insert the QIAamp Mini column into the VacConnector on the QIAvac 24 Plus. Insert a 20 mL tube extender onto the open QIAamp Mini column.
  9. Carefully apply the lysate–Buffer ACB mixture from **step 7** into the tube extender of the QIAamp Mini column. Repeat with all samples. Switch on the vacuum pump. When all the contents have been drawn through the columns completely, switch off the vacuum pump. Carefully remove and discard the tube extender when the pressure has reached 0 mbar.
  10. Apply 600 µL Buffer ACW1 to the QIAamp Mini column. Leave the lid of the column open, and switch on the vacuum pump until all of ACW1 has been drawn through the QIAamp Mini column.
  11. Switch off the vacuum pump and release the pressure to 0 mbar.
  12. Apply 750 µL Buffer ACW2 to the QIAamp Mini column. Leave the lid of the column open, and switch on the vacuum pump. After all of Buffer ACW2 has been drawn through the

QIAamp Mini column, switch off the vacuum pump and release the pressure to 0 mbar.

13. Apply 750 µL of ethanol (>95%) to the QIAamp Mini column. Leave the lid of the column open, and switch on the vacuum pump. After all of ethanol has been drawn through the spin column, switch off the vacuum pump and release the pressure to 0 mbar.
14. Close the lid of the QIAamp Mini column. Remove it from the vacuum manifold and place the QIAamp mini column into a clean 2 mL collection tube. Centrifuge at full speed ( $20,000 \times g$ ) for 3 min.
15. Place the QIAamp Mini column into a new 2 mL collection tube. Open the lid, and incubate the assembly at 56 °C for 10 min to dry the membrane completely.
16. Place the QIAamp Mini column (from **step 15**) into a clean 1.5 mL elution tube and discard the 2 mL collection tube from **step 15**. Carefully apply 100 µL of Buffer AVE to the center of the QIAamp Mini membrane. Close the lid and incubate at room temperature for 3 min.
17. Centrifuge in a microcentrifuge at full speed ( $20,000 \times g$ ) for 1 min. The eluate contains cfDNA.
18. Store upright at –80°C.

### **3.3 Droplet Digital**

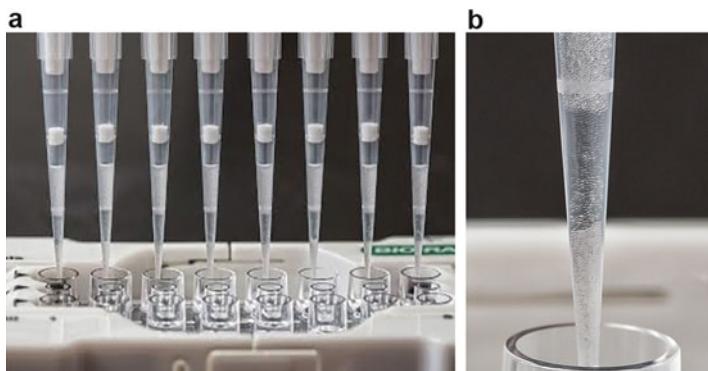
#### **PCR Sample Preparation, Thermocycling, and Reading**

*See Notes 10–12.*

Follow generally accepted PCR rules to avoid contamination:

- (a) Wipe down work surfaces and equipment using 70% ethanol: hood, bench, racks, pipettes, cartridge holders, waste beaker, droplet generator and heat sealer before you start and after you finish. Clean your working area on a weekly basis using DNA Zap.
  - (b) Change gloves frequently: always use CLEAN gloves when preparing a master mix, especially when opening a Taqman probe tube. Change gloves between handling positive controls and patient samples.
  - (c) Use aerosol resistant pipette tips and calibrated pipettes. Check liquid level in the tip before/after pipetting.
1. In a micro centrifuge tube, prepare the reaction mixture in the exact order indicated below, mix five to ten times by gently pipetting up and down.
    - (a) 6.875 µL deionized Water
    - (b) 12.5 µL 2× ddPCR Supermix
    - (c) 0.625 µL 40× Taqman primer–probe mix
    - (d) When preparing master mix for multiple samples, allow 10% extra volume for pipetting.

2. Transfer master mix into a DNase- and RNase-free pipetting reservoir.
3. Use an eight-channel pipettor, transfer 20 µL reaction mixture from **step 1** into wells of a PCR plate.
4. Add 5 µL of isolated cfDNA to each well according to predetermined sample layout. Each assay from one specimen is run in triplicates of 5 µL of isolated cfDNA for a total of 15 µL representing approximately 300 µL of isolated plasma (*see sample plate layout in step 14*). Pipette up and down slowly three times to mix.
5. Secure the DG8 cartridges in cartridge holders. Load 20 µL of reaction mix to the middle row.
6. Fill oil wells, in the lower row of the cartridge, with 70 µL of Droplet Generation Oil.
7. Cover the cartridge with the DG8 gasket and load the cartridge holder with the DG8 cartridge into Droplet Generator.
8. When the light on Droplet Generator turns green, take out the cartridge and holder.
9. Use a manual 50 µL eight-channel pipettor, gently and slowly pipette up 40µL droplets from the top row of the cartridge over a 5 s period. Gently release the droplets into a 96-well PCR plate over a 5 s period with each pipette tip touching the side of a well near—but not at—the well bottom to avoid shearing the droplets (Fig. 2).
10. Repeat droplet generation until all the cartridges are processed and the 96-well plate is filled to the extent desired.
11. Cover the PCR plate with a sheet of Easy Pierce Foil PCR Plate Seal. Mark A1 at the right corner. Seal the PCR plate using a preheated Eppendorf PCR Plate Sealer by pressing to the second tier and counting to six. Rotate the plate 180° and repeat.



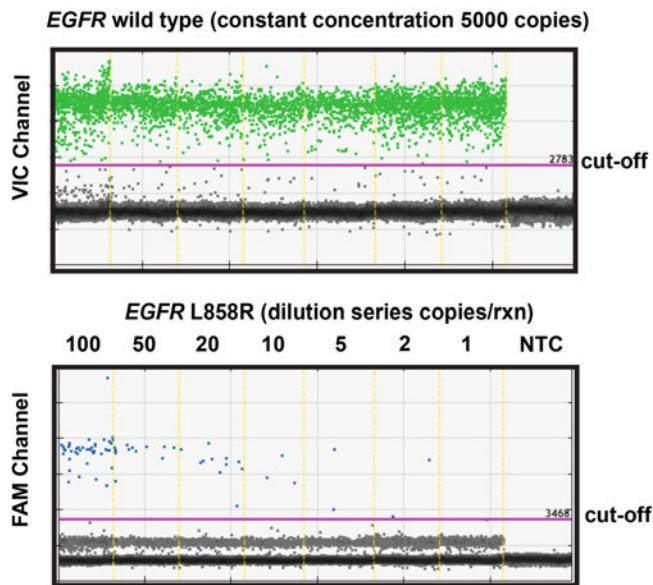
**Fig. 2** Example of good emulsification of droplet generation. A successful emulsification should yield a clear separation of two phases (a) that consist of individual droplets (b)

12. Place the sealed plate in a thermal cycler and run under the following conditions:
  - Lid temperature at 105 °C.
  - (a) 95 °C for 10 min.
  - (b) 94 °C 30 s. Ramp to annealing temperature at 2.5 °C/s.
  - (c) 1 min at annealing temperature.
    - EGFR ex19 = 55 °C.
    - EGFR L858R = 58 °C.
    - EGFR T790M = 58 °C.
  - (d) Repeat b and c for 40 cycles.
  - (e) Hold at 10 °C.
13. Set up a plate layout in QuantaSoft.
14. Prime the QX100 reader with droplet reader oil.
15. Transfer the finished PCR plate to the QX100/200 and read the plate. A general plate layout for 24 samples is shown below:

	1	2	3	4	5	6	7	8	9	10	11	12
A	Hi Ctrl	#1	#1	#1	#9	#9	#9	#17	#17	#17	Hi Ctrl	
B	Med Ctrl	#2	#2	#2	#10	#10	#10	#18	#18	#18	Med ctrl	
C	Lo Ctrl	#3	#3	#3	#11	#11	#11	#19	#19	#19	Lo Ctrl	
D	NTC	#4	#4	#4	#12	#12	#12	#20	#20	#20	NTC	
E	NTC	#5	#5	#5	#13	#13	#13	#21	#21	#21	NTC	
F	NTC	#6	#6	#6	#14	#14	#14	#22	#22	#22	NTC	
G	NTC	#7	#7	#7	#15	#15	#15	#23	#23	#23	NTC	
H	NTC	#8	#8	#8	#16	#16	#16	#24	#24	#24	NTC	

### 3.4 Analysis of EGFR L858R and T790M Droplet Data

1. These assays are analyzed using the QuantaSoft 1D droplet plots. Set the FAM and VIC thresholds based on the amplitude of the positive controls (Fig. 3). General observed fluorescence amplitude ranges for L858R and T790M are listed in Table 2.
2. Sum up the “concentrations/20 μL” of PCR reaction (provided as results output by QuantaSoft) for each triplicate belonging to the assay.
3. Results are normalized to copies/mL of plasma by multiplying by 3.3. This dilution factor accounts for the volume of sample used for ddPCR, the number of replicates, the volume of plasma that was extracted, and the volume the extraction was eluted in.



**Fig. 3** Analysis of EGFR L858R by 1D gating. The top (WT) and bottom (MUT) images are VIC and FAM amplitudes of varying EGFR L858R concentrations in a constant background of 5000 copies of wild type EGFR. Cutoff intervals are generally established by running positive control samples (e.g., H1975 genomic DNA for EGFR L858R) along unknowns

**Table 2**  
Typical amplitudes of individual EGFR mutations using ddPCR

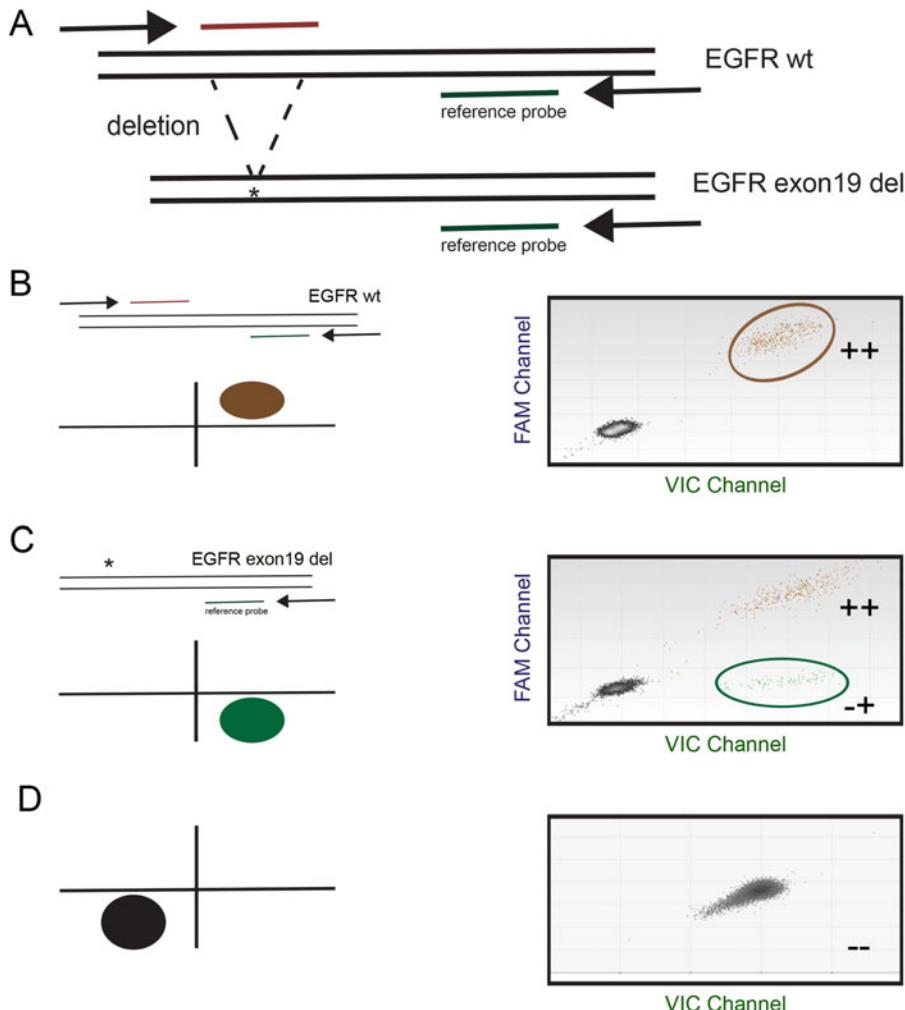
Assay	Amplitude range	
	FAM	VIC
EGFR del19	>3500	>3000
EGFR L858R	>3500	>3000
EGFR T790M	>4000	>2000

### 3.5 Analysis of EGFR Exon19 Deletion Droplet Data Using QuantaSoft 2D Droplet Plots

See Note 13.

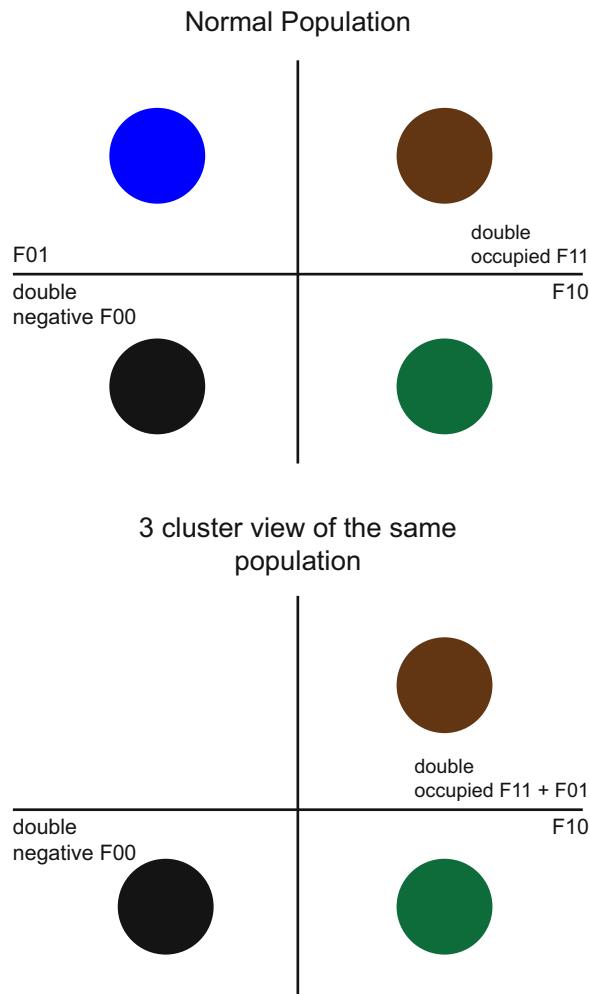
Our EGFR exon19 deletion assay uses a VIC-labeled “reference probe” sequence that is shared by both wild-type and mutant alleles, and thus gives a signal for both, while the FAM-labeled probe sequence spans the exon19 deletion region and thus is only detected on wild-type targets [10]. As a consequence, this assay type only gives three discernible clusters. Figure 4 shows the assay design and 2D droplet plot examples for the three observed clusters.

1. Select 2D Amplitude in QuantaSoft.
2. Click on lasso function under “Analysis” to the left.



**Fig. 4** Analysis of *EGFR* exon 19 deletion by 2D gating. (a) The assay is designed so that a VIC-labeled “reference probe” is shared by wild-type and the deletion 19 exon mutants, while a FAM-labeled probe sequence spans the deletion hotspot. Thus, an *EGFR* wild-type sample will therefore show double positive (brown) droplets (b), while an *EGFR* exon19 mutant will only have VIC-labeled droplets (bottom right) (c). Negative droplets with no template are shown in d and are in the bottom left corner

3. Lasso the (FAM+VIC+) cluster and the VIC+ only positive cluster based on the amplitudes of the positive controls. The former contains the total wild-type *EGFR* population and the latter *EGFR* exon19 deletion population only. An example of a wild-type only population is shown in Fig. 4b and a sample that additionally contains *EGFR* exon19 deletions in Fig. 4c.
4. Sum the copies/20  $\mu$ L well of PCR of each triplicate for the FAM+ population, and the copies/20  $\mu$ L well of PCR of each triplicate for the VIC+ cluster.
5. Subtract the value of FAM+ from the value of VIC+. This value estimates the concentration of *EGFR* exon19 deletion. See



**Fig. 5** In a nonstandard, three-cluster, drop-off assay, WT targets appear as doubly positive droplets rather than single positive droplets. As a result, the QuantaSoft algorithms used to compute the channel concentrations (where normally, one dye channel represents the presence of one species and double positive droplets represent the presence of both species) must be interpreted differently. The clusters are designated as C1 (F00), C2(F10) and C3 (F11 + F01) and refer to the number of droplets in that particular cluster. Based on Poisson Statistics the concentration (in copies/uL) of WT and mutant alleles in the ddPCR reaction are given by: (1)  $\text{Conc}_{\text{Ch1}} = \text{Conc}_{\text{WT}} = \{-\ln [(C1 + C2)/(C1 + C2 + C3)]\} \times [(1000\text{nL}/\text{uL})/(0.85\text{nL}/\text{droplet partition})]$  (2)  $\text{Conc}_{\text{Ch2}} = \text{Conc}_{\text{WT}} + \text{Conc}_{\text{MUT}} = \{-\ln [C1/(C1 + C2 + C3)]\} \times [(1000\text{nL}/\text{uL})/(0.85\text{nL}/\text{droplet partition})]$  (3)  $\text{Conc}_{\text{Ch2}} - \text{Conc}_{\text{Ch1}} = \text{Conc}_{\text{MUT}} = \{-\ln [C1/(C1 + C2)]\} \times [(1000\text{nL}/\text{uL})/(0.85\text{nL}/\text{droplet partition})]$ . These formulas assume that the VIC (or HEX) only channel is monitoring the mutant allele

Fig. 5 for further explanations how concentrations are calculated for three-cluster, drop-off assays.

6. Results are normalized to copies/mL of plasma by multiplying by 3.3.

### **3.6 Analytical Considerations for ddPCR Assays of cfDNA**

Sensitivity and specificity of ddPCR assays:

In our experience the lowest % of mutant allele that can be reliably detected in a wild type background is 0.05%. We recommend that prior to testing plasma cfDNA samples one should test their technique on low/high copy positive and negative constructed samples to help assess the sensitivity and specificity of *EGFR* T790M, *EGFR* exon19 deletion, and L858R in their own laboratory. Positive control samples can be purchased from Horizon Discovery. Typical acceptance criteria for constructed specimens are:

1. Sensitivity: Seven molecules for all *EGFR* mutations (T790M, L858R, exon19 deletion) in a “low” wild-type background (2000 genome equivalent wt background), or 20 molecules for all *EGFR* mutations in a “high” wild-type background (50,000 genome equivalent wt background)
2. Specificity: 100% for all *EGFR* mutations in both low and high wild-type backgrounds
3. No single positive droplets in any no-template controls. Run at least 5 NTC samples per plate.

### **3.7 Interpretation of Plasma Genotyping Results**

*See Note 14.*

Interpretation of ddPCR-based plasma genotyping results is highly dependent on clinical context, disease characteristics and the manner in which the assay is reported. The dynamic range of plasma ddPCR in advanced NSCLC is wide and may be influenced by disease burden and tumor biology [11, 13]. As such, reporting ddPCR assay results as copies per mL of plasma provides information on both tumor genotype as well as disease status and prognosis [14]. The same factors that affect dynamic range are also key determinants of assay sensitivity with false positive results being more likely in patients with more limited disease, presumably secondary to lower rates of cfDNA shed from tumor [13]. The timing of sample collection with respect to treatment also has a significant impact on assay sensitivity with maximal assay sensitivity occurring at initial diagnosis of metastatic disease as well as at subsequent points of clear disease progression. Changes in quantitative plasma cfDNA can be expected in response to therapy with decreasing levels associated with response to effective therapy.

Plasma genotyping using a ddPCR-based assay exhibits exquisite specificity for the detection of targetable genomic alterations in newly diagnosed advanced NSCLC [11, 12]. The detection of *EGFR* sensitizing mutations using this method can thus reliably identify patients that will benefit from targeted therapy. Conversely, the detection of nontargetable driver mutations such as KRAS codon 12 mutations can be used as a reliable indicator that targetable alterations are unlikely to be identified by other methods with

the possible exception of *BRAF*V600E alterations [15]. Apparent false positives in newly diagnosed patients should prompt careful scrutiny of tissue genotyping results and consideration of repeat biopsy, particularly if high levels of mutant cfDNA are detected. The specificity of plasma ddPCR for the detection of acquired resistance mutations such as *EGFR* T790M is more complex to interpret. Heterogeneity of acquired resistance mechanisms across metastatic sites within the same patient is not uncommon. Tissue genotyping can thus only detect apparent resistance mutations at a single metastatic site, whereas plasma ddPCR represents an average of resistance across all metastatic sites. The thresholds of plasma *EGFR* T790M copy number, which predict response to *EGFR* T790M specific inhibitors represents an active area of investigation. However, the gold standard comparisons for evaluating plasma genotyping methods in the acquired resistance setting are challenging given the limitations of tissue genotyping in this clinical context.

---

#### 4 Notes

1. Cheaper pipette tips often contain rough tip edges that can shed microscopic plastic particulates into sample preps and ddPCR reactions resulting in shredding of droplets.
2. Time from draw to freezing of plasma should be less than 4 h.
3. During centrifugation, brake switch must be off so the cell–plasma interface is not disturbed.
4. When transferring plasma, do not dip the tip of the pipette into the plasma–cell interface. Leave a thin plasma layer intact over the interface.
5. Sometimes the color varies from pink to red due to hemolysis, which will not affect downstream ddPCR genotyping. If plasma is too viscous for pipetting, it may indicate lysis of white blood cells, which causes overwhelming copies of wild-type allele being detected and may affect the ability to detect rare events.
6. During lysis, make sure that a visible vortex forms in the tube. To ensure efficient lysis, it is essential that the sample and Buffer ACL are mixed thoroughly. Do not interrupt the procedure at this time. Proceed immediately to **step 5**.
7. When setting up the QIAamp Mini columns over the QIAvac 24 Plus, make sure that the tube extender is firmly inserted into the column in order to avoid leakage of sample. Keep the collection tube for the dry spin in **step 14**.
8. To avoid cross-contamination, be careful not to move the tube extenders over neighboring QIAamp Mini Columns.

9. Do not use plasma that has been frozen and thawed more than once to prevent cryoprecipitates from clogging the QIAamp Mini column. In case cryoprecipitates are clearly visible the sample may be centrifuged for 5 min at 16,000 × g.
10. Set up ddPCR reactions in an amplicon-free PCR work station that is in an amplicon-free room.
11. The following protocol below is for the manual preparation of droplet generation. More recently Bio-Rad's automated droplet generator (Bio-Rad #1864101) can also be used.
12. Always carry no-template controls (NTCs) to monitor for contaminations. If positive droplets are observed in NTC wells, results from the plate cannot be used.
13. During droplet generation process, both sample wells and oil wells must contain liquid or the droplet generator will not generate droplets. Supermix (1×) can be used in the sample wells that are not being used.
14. Using 2 mL of plasma one expects to isolate between 10 and 150 ng DNA (~3000–45,000 wild-type events).

## References

1. Shaw AT, Kim DW, Nakagawa K et al (2013) Crizotinib versus chemotherapy in advanced ALK-positive lung cancer. *N Engl J Med* 368:2385–2394
2. Mok TS, Wu YL, Thongprasert S et al (2009) Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 361:947–957
3. Bergethon K, Shaw AT, Ignatius Ou SH et al (2012) ROS1 rearrangements define a unique molecular class of lung cancers. *J Clin Oncol* 30:863–870
4. David Planchard JM, Riely GJ, Rudin CM et al (2013) Interim results of phase II study BRF113928 of dabrafenib in BRAF V600E mutation-positive non-small cell lung cancer (NSCLC) patients. *J Clin Oncol* 31 (Supplement):8009
5. Garraway LA, Janne PA (2012) Circumventing cancer drug resistance in the era of personalized medicine. *Cancer Discov* 2:214–226
6. Zhou W, Ercan D, Chen L et al (2009) Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. *Nature* 462:1070–1074
7. Shaw AT, Kim DW, Mehra R et al (2014) Ceritinib in ALK-rearranged non-small-cell lung cancer. *N Engl J Med* 370:1189–1197
8. Diaz LA Jr, Williams RT, Wu J et al (2012) The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature* 486:537–540
9. Kuang Y, Rogers A, Yeap BY et al (2009) Non-invasive detection of EGFR T790M in gefitinib or erlotinib resistant non-small cell lung cancer. *Clin Cancer Res* 15:2630–2636
10. Yung TKF, Chan KCA, Mok TSK et al (2009) Single-molecule detection of epidermal growth factor receptor mutations in plasma by microfluidics digital PCR in non-small cell lung cancer patients. *Clin Cancer Res* 15:2076–2084
11. Oxnard GR, Pawletz CP, Kuang Y et al (2014) Noninvasive detection of response and resistance in EGFR-mutant lung cancer using quantitative next-generation genotyping of cell-free plasma DNA. *Clin Cancer Res* 20:1698–1705
12. [http://www.bio-rad.com/webroot/web/pdfs/lsr/literature/Bulletin\\_6407.pdf](http://www.bio-rad.com/webroot/web/pdfs/lsr/literature/Bulletin_6407.pdf)

13. Sacher AG, Paweletz CP, Alden R, et al. (2015) A prospective study of rapid plasma genotyping utilizing sequential ddPCR and NGS in newly diagnosed advanced NSCLC patients. IASLC World Conference on Lung Cancer, Denver, CO
14. Mok TS, Wu YL, Soo Lee J et al (2015) Detection and dynamic changes of EGFR mutations from circulating tumor DNA as a predictor of survival outcomes in NSCLC patients treated with first-line intercalated erlotinib and chemotherapy. *Clin Cancer Res* 21:3196–3203
15. Sholl LM, Aisner DL, Varella-Garcia M et al (2015) Multi-institutional oncogenic driver mutation analysis in lung adenocarcinoma: the lung cancer mutation consortium experience. *J Thorac Oncol* 10:768–777



# Chapter 13

## Detection of Cancer DNA in Early Stage and Metastatic Breast Cancer Patients

Arielle J. Medford, Riaz N. Gillani, and Ben Ho Park

### Abstract

Breast cancer is the leading cause of cancer in women and the second leading cause of cancer-related death. There are many subtypes of breast cancer, which can be identified through the process of molecular and genetic profiling. While the current standard of care utilizes tumor tissue biopsy to subclassify breast cancer, plasma tumor DNA (ptDNA) can be detected through droplet digital PCR (ddPCR) of plasma obtained from a simple blood draw. Tissue biopsy is not only more invasive but because tumors exhibit heterogeneity it can be less accurate. Blood collects DNA shed from normal and cancerous cells alike, thus ddPCR of plasma offers a broader picture of a cancer's genetic makeup. This chapter summarizes how patients with breast cancer can be screened for specific cancerous mutations in both tissue and plasma through the use of ddPCR.

**Key words** Breast cancer, Droplet digital, PCR, Liquid biopsy, Biopsy, PIK3CA, Plasma, Blood

---

### 1 Introduction

Every year, over 230,000 women are diagnosed with breast cancer, and nearly 40,000 die from the disease. It is the most commonly diagnosed cancer in women and the second leading cause of cancer-related death [1, 2].

Typically, breast cancer is detected via mammography and confirmed through tissue sampling. From here, patients undergo specific therapies targeted toward their class of breast cancer, which is defined principally by the presence or absence of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER-2). Researchers have shown that within these subgroups, there are specific, idiosyncratic mutations, which

---

Arielle J. Medford and Riaz N. Gillani contributed equally to this work.

can be useful for assessing burden of disease, predicting response to therapy, and monitoring the latter in a rapid and minimally invasive approach. These include mutations in PIK3CA, ESR1, TP53, MAP3K1, GATA3, RPTOR, ERBB2, and ERBB3 [3–6].

Current studies of receptor mutations, however, face a significant barrier: they rely on single site biopsy of tumor tissue. This method is problematic because tumors have been shown to exhibit heterogeneity, so sampling a single piece of tissue may miss sites of the specific mutations listed above. This same heterogeneity can exist within metastatic tumors, and furthermore, there has been documented heterogeneity between metastases at different sites of the body, even if the metastases originated from the same primary tumor [7, 8]. In other words, tissue biopsy limits clinicians from seeing the whole genetic picture.

In order to circumvent this challenge, researchers can now use droplet digital PCR (ddPCR) on samples of plasma to screen for these same mutations, through a procedure colloquially termed “liquid biopsy.” The body’s blood supply feeds primary tumors as well as metastases, and preliminary research suggests there is no selectivity in terms of which tumor cells’ DNA is deposited into the blood. In fact, normal and cancerous cells deposit into circulation, which makes the bloodstream an effective repository of cellular DNA, cancer and otherwise [9–11]. Thus, circulating blood contains traces of all variations of tumor DNA, which makes plasma tumor DNA (ptDNA) most representative of the varied genetic makeup of cancer cells. Sampling peripheral blood is also much easier than tissue biopsy, and the procedure comes with far fewer risks.

There are many burgeoning applications of ddPCR within oncology, and breast cancer is poised to greatly benefit from the approach. The efficacy of ptDNA/ddPCR for detecting PIK3CA mutations in breast cancer has already been demonstrated with 93.3% sensitivity and 100% specificity in a study of 29 patients with early stage breast cancer [12]. It is important to clarify that this approach is specific for plasma, not serum. Research has shown the latter to not be reliable, whereas plasma has reproducible data supporting its efficacy [13, 14]. From identifying specific mutations sensitive to targeted therapy, to monitoring response to treatment, to detecting residual disease, this highly sensitive, cost effective approach offers improvement of breast cancer treatment and forward progress in studying the science behind it. This protocol describes the methodology for using ddPCR to detect tumor DNA in both plasma and primary tumor tissue.

---

## 2 Materials

### 2.1 Design and Testing of ddPCR Primers and Probes

See Subheading 2.5.

### 2.2 Blood Handling and Preparation

1. EDTA or Cell-free DNA BCT tubes (Streck).
2. 23-gauge needle (*see Note 1*).

### 2.3 Blood Sample Processing and Plasma Separation

1. Isolated hood space to prevent contamination.
2. Bleach wipes (1%).
3. QIAamp Circulating Nucleic Acid Kit (Qiagen cat#: 55114).
4. DNase/RNase-Free distilled water.
5. Centrifuge.
6. DNase/RNase-free filter pipette tips (*see Note 2*).

### 2.4 Processing and DNA Extraction from Primary Tumor Samples

1. Five to ten- $\mu$ m thick unstained FFPE histology slides (10–15 slides depending on tumor size and the percentage of tumor cells in the slices).
2. Container for soaking of the histology slides in xylene, ethanol, and water.
3. Adjacent tissue slide with H&E stain, with tumor tissue demarcated from normal by a pathologist.
4. >99% xylene.
5. DNase/RNase-free distilled water.
6. 100, 70, and 50% ethanol.
7. Pinpoint Slide DNA Isolation System (Zymo Cat#: D3001).
8. QIAamp DNA FFPE tissue kit (Qiagen Cat#: 56404).

### 2.5 ddPCR of ptDNA or FFPE Samples

#### 2.5.1 Preamplification

1. Isolated hood or sterilized space to prevent contamination (separate from blood and plasma processing and preferably in separate room).
2. Sample to be used as template (prepared from plasma or FFPE tissue).
3. DNase/RNase-Free distilled water.
4. dNTPs (10 mM concentration).
5. Dimethyl sulfoxide (DMSO).
6. 5 $\times$  Phusion Buffer: MgCl<sub>2</sub>, DMSO (NEB Cat#: B0518S).
7. Forward and reverse primers for amplification of locus of interest (50  $\mu$ M).

8. Phusion High-Fidelity DNA polymerase (NEB Cat#: M0530S).
9. Thermal cycler with 96-deep well reaction module.
10. 96-well PCR plates.
11. Foil heat seals.
12. DNase, RNase-free filter pipette tips.
13. 1.5 mL Eppendorf tubes.
14. QIAquick PCR purification kit (Qiagen Cat#: 28104).

#### 2.5.2 ddPCR

1. ddPCR sterile hood (separate from pre-amp and blood/plasma prep hoods and preferably in a separate room).
2. Sample prepared in preamplification step.
3. DNase/RNase-free distilled water.
4. Bleach wipes (1%).
5. Sterile troughs.
6. DG8 cartridges for QX200 droplet generator (Bio-Rad Cat#: 1864008).
7. DG8 cartridge holder (Bio-Rad Cat#: 1863051).
8. DG8 gaskets for QX200 droplet generator (Bio-Rad Cat#: 1863009).
9. Pipettes with tips: p10, p20, p100, p50 multichannel, p100 multichannel.
10. 1.5 mL Eppendorf tubes.
11. ddPCR supermix for probes (No dUTP) (Bio-Rad Cat# 186-3023).
12. Droplet generation oil for probes (Bio-Rad Cat# 186-3005).
13. Forward and reverse primers (50 µM)—these may be the same primers used for locus-specific preamplification (Ordered as “custom oligos” through [IDTDNA.com](http://IDTDNA.com)).
14. Wild-type and mutant probes (100 µM).
15. Promega Female (cat#: Promega G1521) (*optional*).
16. Four-to-six cutter restriction enzyme (*optional*).
17. 96-well semiskirted plates.
18. Pierceable foil heat seal (Bio-Rad Cat#: 1814040).
19. QX200 droplet generator (Bio-Rad Cat# 1864002).
20. PX1 PCR plate sealer (Bio-Rad Cat#: 1814000).
21. C1000 Touch Thermal Cycler with 96-Deep Well Reaction Module (Bio-Rad Cat#: 1851197).
22. QX200 droplet reader (Bio-Rad Cat#: 1864003).
23. QX200 droplet reader oil (Bio-Rad Cat#: 1863004).

24. Laptop computer with Windows operating system.
25. Quantasoft Software.

## **2.6 Data Interpretation and Statistical Analysis**

1. Laptop computer with windows operating system.
2. Quantasoft Software.
3. Microsoft Excel.

## **3 Methods**

*The following is the complete protocol for use with blood samples. For use with FFPE, use the following protocol with the modifications specified in the addendum at the end of the Methods section.*

### **3.1 Design and Testing of ddPCR Primers and Probes**

1. Primers and probes can be ordered at [idtdna.com](http://idtdna.com).
2. The ideal amplicon size ranges from 80 to 150 base pairs (*see Note 3*).
3. When designing primers, incorporate the following parameters:
  - (a) Maintain a length of 20–28 base pairs.
  - (b) Melting temperature ( $T_m$ ) should range from 45 to 55 °C.
  - (c) Do not include more than three repeated base pairs in a row.
  - (d) Confirm there is no primer heteroduplexing.
4. Probes are fluorescently labeled oligonucleotides. When designing these, incorporate the following parameters:
  - (a) Maintain a length of 20–30 base pairs.
  - (b) The melting temperature ( $T_m$ ) should be 5–10 °C higher than that of the primers.
  - (c) The 5'-end of the probe should NOT include guanine.
  - (d) The 5'-end should include the fluorophore.
  - (e) The 3'-end should include the fluorophore quencher.
  - (f) The probe should start at least 2 base pairs away from primer binding site.
  - (g) Wild type probes are typically ordered with the fluorophore Hex.
  - (h) Mutant probes are typically ordered with the fluorophore 6-Fam (*see Note 4*).
  - (i) Both probes should be quenched with Zen 3'Iowa Black FQ.
5. When testing probes for sensitivity and specificity, first create a positive control. This can be from a cell line that carries the

mutation, synthetic double stranded DNA or cloned plasmid DNA. The positive control will confirm the probe's capacity to bind the target mutant allele.

6. Test the sensitivity and specificity of the positive control by first ordering a “mini-order” of the mutant and wild type probes and follow the “Subheading 3.5” detailed below, starting with step 6 of “reaction preparation.”
1. Draw the desired quantity of blood from patients (~10 mL per tube) into either an EDTA tube or a Cell-free DNA BCT Tube (*see Notes 5–8*).
  - (a) If using EDTA, extract plasma from the blood within 2 h. Plasma can be stored at –80 °C.
  - (b) If using Cell-free DNA BCT tubes, blood can be stored up to 7 days at room temperature. Tubes MUST be stored at room temperature, and cold and heat should be avoided (acceptable temperature range: 6–37 °C).
2. If transporting blood, assure proper transport temperature (*see step 1a and 1b*).

### **3.3 Blood Sample Processing and Plasma Separation**

1. Put on two layers of gloves and open the biohazard bag containing the patient sample in the PCR hood. Remain in the hood for the remainder of the blood sample processing steps, except for centrifugation. This precaution minimizes risk of contamination of this very sensitive assay.
2. Open the EDTA tube, and transfer its contents to a 15 mL conical tube. After the blood is removed, dispose of the tube and the outer layer of gloves in the original biohazard bag to minimize contamination.
3. Centrifuge the 15 mL conical tubes at  $1500 \pm 150$  rcf for 10 min. This step can be done at room temperature or at 4 °C.
4. Use a 10 mL serological pipette to transfer the supernatant to a new 15 mL conical tube. Do not combine plasma from the same patient at this step. In other words, each tube should translate to its own conical tube at this step. Do not touch the buffy coat. It is good practice to leave a bit of plasma volume behind to ensure minimum contamination of cells (*see Note 9*).
  - (a) Optional: Buffy coats can be saved or processed at this stage for baseline germline DNA (*see Note 10*).
5. Centrifuge the conical tubes at  $3000 \pm 150$  rcf for 10 min.
6. Use a 5 mL serological pipette to transfer the supernatant to a new 15 mL conical tube. At this point, plasma from the same patient can be combined if there is no visible hemolysis. Leave 0.3 mL (~7 mm) on the bottom of the tube to avoid cellular contamination.

7. Gently mix and record the plasma volume. Expect ~4 mL plasma per 10 mL blood.
8. *Optional:* Aliquot 1–1.5 mL plasma to a 1.5 mL vial, and store at –80 °C until further use.

### **3.4 Plasma DNA Extraction**

1. Perform all extraction steps in a sterile hood. Prior to beginning extraction, wipe down all surfaces, equipment, and gloves with bleach.
2. If plasma was stored at –80 °C, thaw it on ice.
  - (a) Use a pipette to transfer thawed plasma to a 15 mL conical tube.
  - (b) Centrifuge at 3600 rpm (3000 rcf) for 15 min at 4 °C or room temperature (*see Note 11*).
3. Upon returning to the hood, wipe everything again with bleach. Use a pipette to transfer the supernatant to a new 15 mL conical tube. Leave 0.3 mL (~7 mm), since the pellet will likely not be visible. The plasma may now be combined in up to 3 mL per tube (two 1.5 mL vials). This is the maximum volume specified for the QIAamp® Circulating Nucleic Acid Kit.
4. Prepare up to 3 mL plasma according to the QIAamp® Circulating Nucleic Acid Kit manufacturer instructions. Elute DNA in 50 µL of water. They can be used immediately or else stored at 2–8 °C if they are to be used in the next 24 h, or –15 to 30 °C until ready for use (*see Notes 12–15*).

### **3.5 ddPCR of ptDNA or FFPE Samples**

#### **3.5.1 Preamplification to Increase the Absolute Count of Alleles or Genome Equivalents**

1. Collect prepared sample to be used as DNA template for ddPCR reaction (from plasma or from FFPE tissue).
2. Genomic DNA (gDNA) from normal cells will serve as an appropriate negative control in most cases, as gDNA should contain the wild-type locus of interest without mutant alleles. Alternatively, Promega sells a female genomic DNA that should also have purely wild-type loci.
  - (a) Prior to use, gDNA or Promega will need to be digested by a restriction enzyme with ubiquitous restriction sites throughout the genome, but no restriction site present in the locus of interest. For most assays, fragments <5 kb are appropriately sized, though depending on the assay, fragments may need to be as small as <500 bp. For this protocol, a four- to six-cutter enzyme (e.g., CviQI) generated appropriately sized fragments (*see Note 16*).
3. In a bleach-cleaned dry hood, create a 50-µL PCR using with separate reactions for the sample, positive control, negative control, and NTC. This is a separate hood from the one to be used for the droplet generation. Each 50 µL reaction will

comprise the following: 2  $\mu$ L of template, 10  $\mu$ L of 5 $\times$  Phusion Buffer, 1  $\mu$ L of dNTPs (10 mM), 1.5  $\mu$ L of DMSO, 1  $\mu$ L of forward primer (50  $\mu$ M), 1  $\mu$ L of reverse primer (50  $\mu$ M), 33  $\mu$ L of DNase/RNAse-Free distilled water, and 0.5  $\mu$ L of Phusion High-Fidelity DNA polymerase. Phusion polymerase should be added last to the reaction.

4. Aliquot each 50  $\mu$ L reaction into a single well of the 96-well plate. Once all samples have been aliquoted, seal plate with foil and centrifuge plate for <15 s.
5. Set PCR cycling protocol as follows on thermal cycler: (1) 30 s at 98 °C; (2) 15 s at 98 °C; (3) 30 s at an annealing temperature that is below the melting point of the primers used for amplification (*see Note 1*); (4) 30 s at 72 °C. Cycle through step 2–4 for 10 cycles. Then resume with steps: (5) 5 min at 72 °C; (6) infinite hold at 4 °C. Ramp setting should be ~2 °C/s.
6. After cycling, run each PCR reaction through PCR purification per “QIAquick PCR purification kit” manufacturer instructions. Elute end-product into 50  $\mu$ L of DNase/RNAse-free distilled water.

### 3.5.2 Reaction Preparation (Perform in Clean Pre-PCR Hood)

1. Prepare for entry into the ddPCR hood by using bleach wipes to clean the exterior, bench top, and any other surfaces that may be encountered during set up of the assay.
2. Obtain ddPCR Supermix from –20 °C freezer and thaw at room temperature.
3. If product from preamplification has been stored at –20 °C, thaw the product at room temperature.
4. After the product has been thawed, mix well and make serial tenfold dilutions of the product using distilled water. These dilutions should span approximately four orders of magnitude and will be used as the templates for the ddPCR assay to ensure that at least some dilutions fall within the dynamic range of the ddPCR system (~5 logs) and give adequate precision. Aliquot each dilution into a 1.5 mL Eppendorf tube.
5. The same serial tenfold dilutions should be done for positive and negative DNA controls and NTC. Aliquot distilled water to be used as a no-template control (NTC) into a 1.5 mL Eppendorf tube.
6. Make the 20 $\times$  primer–probe mixture to be used for the ddPCR reaction according to the proportions below (*see Note 18*):

Reagent	Vol (μL)
Primer F (50 μM)	36
Primer R (50 μM)	36
Probe WT (100 μM)	5
Probe Mut (100 μM)	5
Water	18
Total	100

7. Bring a sterile trough into the ddPCR hood and remove from the protective plastic cover. Place supermix, primer–probe mix, and droplet generation oil in the ddPCR hood.
8. Fill the sterile trough with droplet generation oil using a p1000 pipette with at least 100 μL of oil available for each reaction.
9. Label an Eppendorf tube for every dilution of the product being assayed. If the template has been diluted by tenfold over four orders of magnitude, there should be four Eppendorf tubes (*see Note 19*).
10. Within each Eppendorf tube, the reaction contents prior to droplet generation will contain the supermix, primer–probe mixture, and appropriate dilution of the template. A single reaction volume for ddPCR will be 20 μL, divided according to the table below (Table 1). If each dilution is being run in triplicate (as is recommended), four volumes of reaction should be prepared to accommodate for any pipetting error and to avoid introducing bubbles into the sample well of the DG8. Therefore, 40 μL of Supermix, 4 μL of primer–probe mixture, and 36 μL of template should be combined to form an 80 μL reaction mixture in each Eppendorf tube.

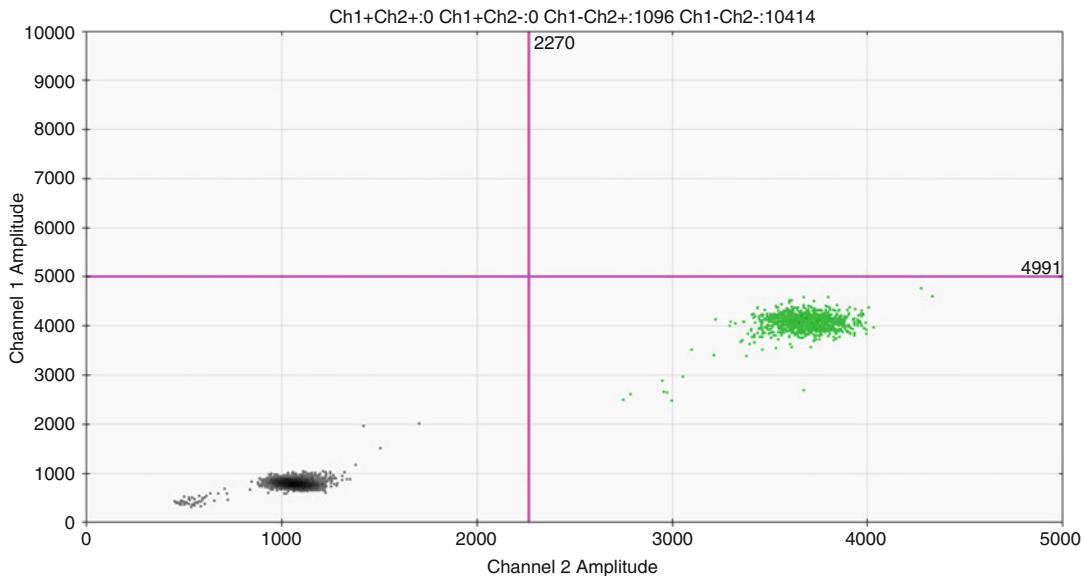
*3.5.3 Cartridge Preparation, Droplet Generation, and PCR Plate Preparation (Fig. 1)*

Suggested Plate Layout

1. Place a single DG8 cartridge into the cartridge holder, sliding the two ends of the cartridge holder inward to lock into place; the cartridge holder can be locked only if the cartridge is inserted with the proper orientation.
2. The cartridge is comprised of 3 rows of 8 wells, with the wells in each row being of a different diameter than those in the other rows. If the cartridge holder is oriented so that text may be read upright by the user, the bottom row of wells will contain the droplet generation oil (as indicated by the label on the cartridge holder). Vortex the contents of each PCR reaction in each Eppendorf tube, and then spin down. In the middle row of wells, labeled as “sample,” aliquot 20 μL of reaction into each well. If running each reaction in triplicate,

**Table 1**  
**Mastermix/primer probe mix**

Reagent	1 × (μL)
Template	9
20× primer–probe mix	1
ddPCR Supermix	10



Template: ESR1 D538G WT plasmid

Probes: ESR1 D538G MUT and WT probe

**Fig. 1** Two-dimensional ddPCR amplitude plot showing that the assay detects the D538G gene (X-axis) and does not erroneously detect the ESR1 wild type (Y-axis) in a D538G mutant plasmid

aliquot 20 μL of each reaction into three adjacent wells. This is ideally done with an eight-channel pipettor, but if this is not possible, it is okay to load 1 well at a time.

3. All 8 wells in the middle row of the cartridge should be filled. Wells remaining in the middle row of the cartridge may be used for a positive and negative control, or else they may be filled with water (*see Notes 20–23*). For the example where each reaction is run in triplicate, a single cartridge may have the wells 1–3 filled with the reaction contents for one dilution of the template, wells 4–6 filled with the reaction contents for the another dilution of the template, well 7 filled with the reaction contents for the positive control, and well 8 filled with the reaction contents for the negative control.

4. Use the p100 multichannel pipette to draw up 70 µL of droplet generation oil per channel; then, aliquot 70 µL of oil into each well in the bottom row of the cartridge (*see Notes 24–26*).
5. After the bottom and middle rows of the cartridge are filled, cover the cartridge with the DG8 gasket, stretching it from one end of the cartridge over to the other and placing it over the prongs to secure it in place.
6. Place the cartridge holder containing the filled cartridge with secured gasket into the QX200 droplet generator. The machine will detect the cartridge and indicate this by lighting up a green light. The generator may then be closed with the click of the top button. Once closed, droplet generation will commence unless there is some error with cartridge preparation. The droplet light on the right of the machine will flash during generation.
7. After several minutes, droplet generation will be complete, which will be indicated by the droplet light on the far right of the machine remaining lit. Open the droplet generator using the button on the sliding green cover, remove the cartridge, and place it back into the ddPCR hood. Remove the gasket, and observe that the upper row of the cartridge is now filled with thousands of droplets individually encapsulating PCR reactions.
8. Place a 96-well semiskirted plate into the ddPCR hood. Use a p50 multichannel pipette set at 43 µL to slowly withdraw the contents from the upper row of the cartridge, and deposit these into a single column of the 96-well semiskirted plate (*see Note 27*).
9. Repeat steps 1–8 as many times as necessary, depending on the number of dilutions that are being analyzed.
10. Once all ddPCR reaction contents have been aliquoted into the 96-well semiskirted plate, place a pierceable foil heat seal over the plate. Turn on the PX1 PCR plate sealer using the power switch on the back of the machine. The plate sealer will heat up to 180 °C over the course of approximately 2 min.
11. Once heating is complete, press the eject button on the touch screen to open the plate sealer; insert the plate with foil on top into the sealer, with the red line on the foil oriented closer to the machine. Select the “seal” option on the touch screen. Once plate sealing is complete, the plate may be removed from the plate sealer, and the plate sealer turned off.

### 3.5.4 PCR and Plate Reading

1. Set PCR cycling protocol as follows on thermal cycler:

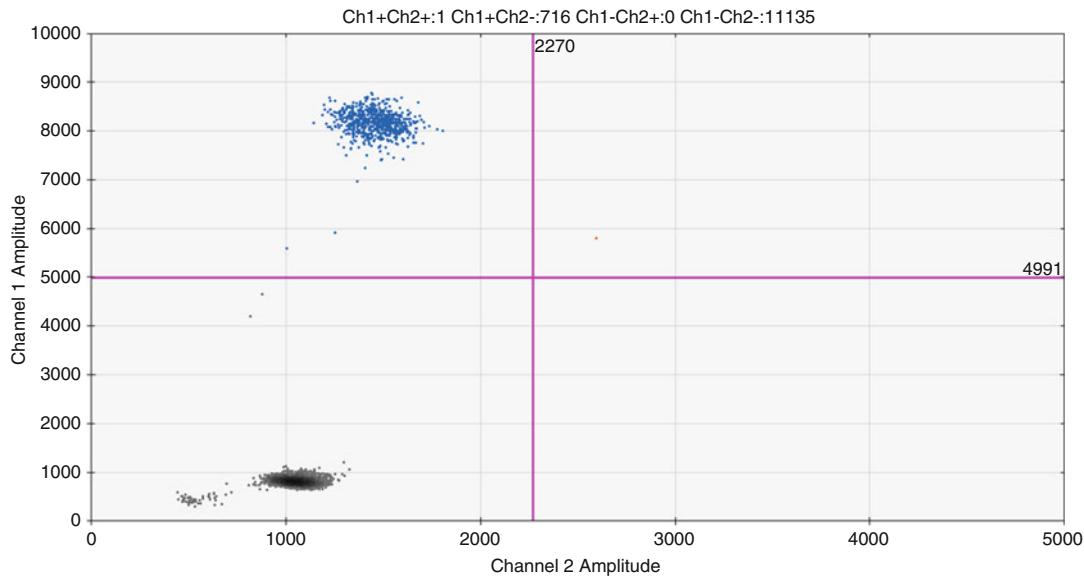
Step	Temp	Time	Ramp	Cycle
1	95	10 min	2°C/s	1
2	94	30 s	2°C/s	40
3	Variable	1 min	2°C/s	
4	98	10 min	2°C/s	1
5	4	Hold		1

2. After the PCR cycling protocol is complete, remove plate from thermal cycler. The QX200 droplet reader should be attached via a USB cable to a computer with a Windows operating system and the Quantasoft software installed. Open the droplet reader using the button on the sliding green cover. Using the latch on either side, remove the metal cover that will be used to secure the 96-well plate in place, slot the plate into place, and then replace and relatch the cover. Close the sliding green cover.
3. On the attached computer, open an instance of Quantasoft. Under the “Setup” option on the left side of the screen, open a prior template or plate, or create a new plate. Label each well on the plate diagram with the sample name and dilution. Once the plate has been saved, follow the prompts to assign the FAM, VIC, and/or HEX probes to the appropriate channels.
4. Select “Run” from the left side of the screen. Indicate the correct dyes (HEX or VIC) (*see Note 28*). The droplet reader will take approximately 2 min to process each well.
5. After the entire plate has been read, the “Analyze” option may be selected on the left-hand side of the screen to review the data from the run.

### 3.6 Data Interpretation and Statistical Analysis (Figs. 2, 3, and 4)

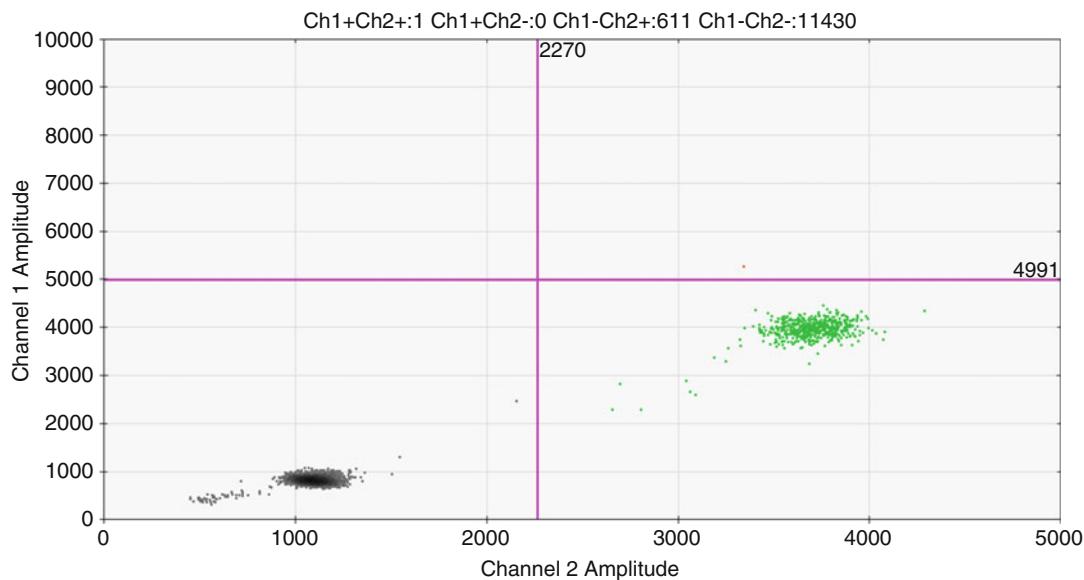
*Caption: Example 2D droplet plots.*

1. Subsequent data analysis is contingent upon fluorescence intensity gating for “positive” droplets containing the mutation of interest. The amplitude threshold for a positive droplet may be set by selecting the “Analyze” option and viewing the “2D amplitude” plots for the various samples. This can be done with the controls first to maximize sensitivity, and then the experimental samples may be superimposed. Place the threshold as low as possible in order to maximize specificity. The positive and negative controls should be evaluated with similar DNA input amounts. An appropriate amplitude cutoff will vary by experiment and should be set based on the amplitude of droplets for controls.



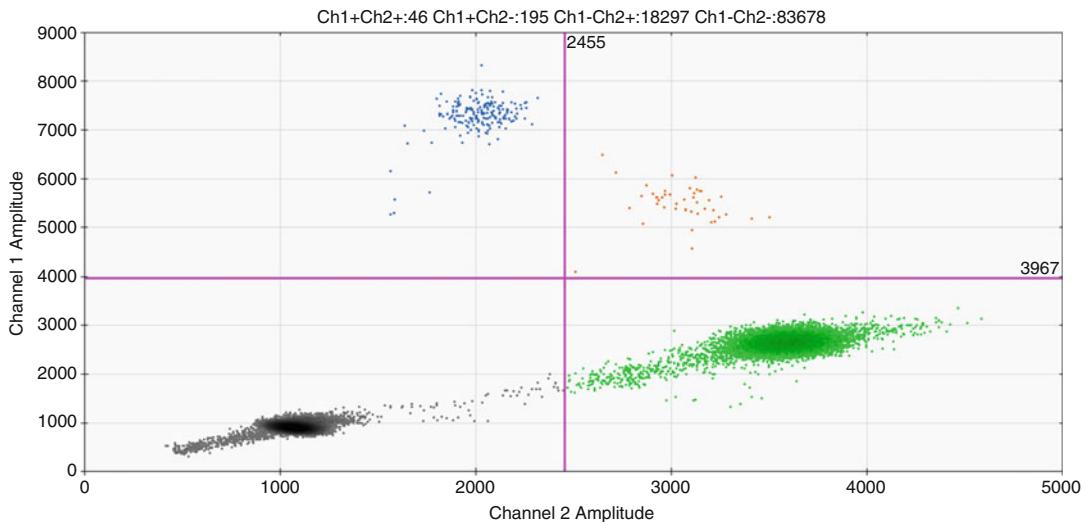
Template: ESR1 D538G MUT plasmid  
Probes: ESR1 D538G MUT and WT probe

**Fig. 2** Two-dimensional ddPCR amplitude plot showing that the assay detects the D538G gene (Y-axis) and does not erroneously detect the ESR1 wild type (X-axis) in a D538G mutant plasmid



Template: digested MCF7 gDNA  
Probes: ESR1 D538G MUT and WT probe

**Fig. 3** Two-dimensional ddPCR amplitude plot showing that the assay detects the wild type ESR1 gene (X-axis) and does not erroneously detect the D538G mutation (Y-axis) in digested MCF7 gDNA



Template: patient cfDNA with known ESR1 D538G mutation  
Probes: ESR1 D538G MUT and WT probe

**Fig. 4** Two-dimensional ddPCR amplitude plot showing that the assay detects the wild type ESR1 gene (X-axis) and the D538G mutation (Y-axis) in patient ctDNA

2. Once this threshold has been set, subsequent analyses may be done. One such analysis is the determination of concentration, which takes a Poisson distribution of occupied and unoccupied droplets into account. Under “Analyze”, the “Concentration” option allows for the visualization of concentration of wild-type and mutant amplicons as copies/ $\mu$ L. Under this same option, “Fractional Abundance” may be selected, allowing for the visualization of concentration of mutant as percentage of total.
3. Another analysis under the “Events” option allows for a visualization of total number of droplets, as well as droplets that are occupied by mutant and/or wild-type amplicons (*see Notes 29–30*).

\*For analysis of FFPE, perform the below DNA extraction and then begin above protocol at Subheading 3.5.

### 3.7 Processing and DNA Extraction from Primary Tumor Samples (FFPE)

1. Heat a histology slide on a heating plate having a 5 to 10- $\mu$ m thick, unstained FFPE section at ~60 °C for ~5–10 min until paraffin wax melts (*see Note 30*).
2. Submerge slide into >99% xylene at room temperature for ~30 min.
3. After 30 min, discard this xylene and resubmerge in new xylene; incubate at room temperature for another 30 min.

4. Discard this second batch of xylene, and then submerge the slide in 100% ethanol for ~2 min. The slide should then be submerged in 70% ethanol, 50% ethanol, and distilled water for ~2 min each, with the goal of gradual hydration.
5. Next, air-dry the slide for ~30–60 min; the slide will be dry when it becomes visibly white.
6. Apply the pinpoint solution from the “Pinpoint Slide DNA Isolation System” to the area of tumor tissue per the manufacturer’s protocol, using another prepared H&E slide of an adjacent cross section of tissue as a guide (*see Note 31*).
7. Remove embedded tissue from the slide with a clean and sharp scalpel, and transfer tissue to a clean Eppendorf tube; centrifuge briefly to bring the blue tissue pieces to the bottom.
8. Follow **step 10** onward from the QIAamp DNA FFPE tissue kit protocol to isolate gDNA from this tissue sample, ultimately eluting into 20–100 µL of distilled water.

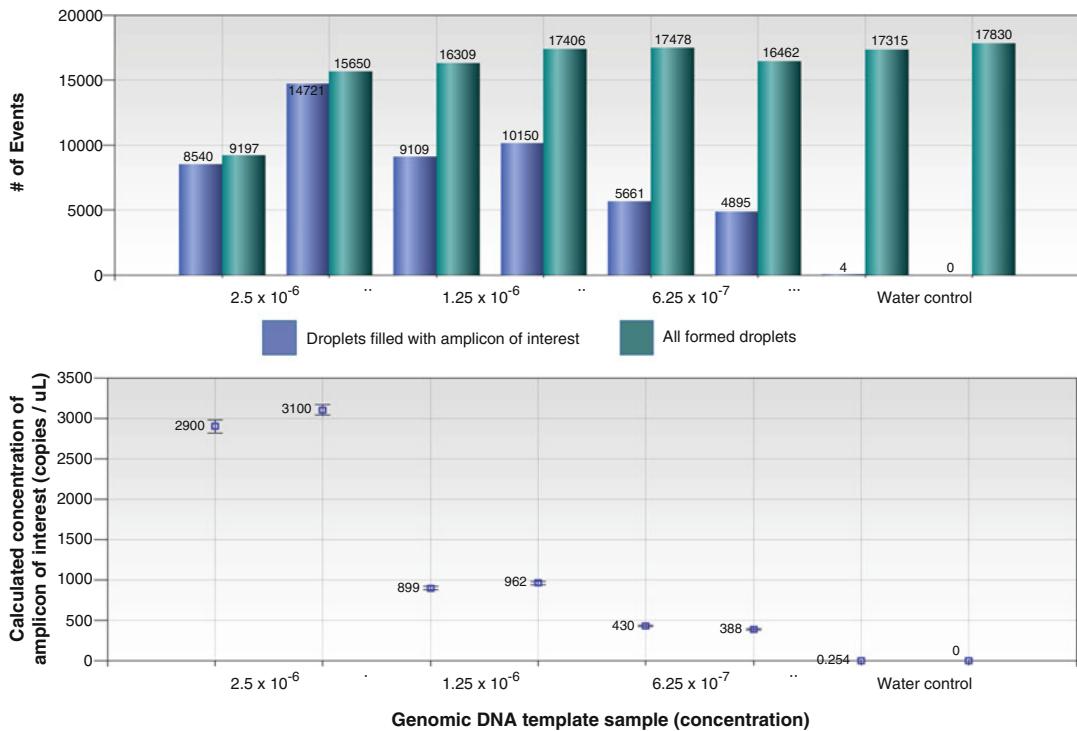
---

#### 4 Notes

1. When drawing blood for use with ptDNA, it is best to use a 23 gauge needle. Smaller sizes risk shearing cells, and larger sizes increases patient discomfort.
2. Bio-Rad protocol specifies Rainin® tips.
3. In general, the smaller the amplicon, the more sensitive the ddPCR detection. Larger amplicons may be used to enable multiple mutation sites to be detected by a single amplicon, but large sizes (>500 bp) begin to lose fluorescence with increasing length.
4. This is the convention. Most importantly, mutant and wild type should have different fluorophores.
5. For this original protocol, researchers drew three tubes of blood (~30 mL).
6. For every 10 mL of blood, expect 4 mL plasma to be isolated.
7. Store EDTA or Cell-free DNA BCT tubes horizontalFly to minimize lysis. When lysed samples are processed, they will have a pink-to-red tinge to the supernatant. Nonlysed samples should have clear-to-yellow supernatants.
8. Tumor DNA is detected in plasma.
9. When transferring the supernatant during blood sample preparation, avoid using bulbs for suction, which may inadvertently lead to suctioning too quickly and disruption of layers.
10. The buffy coat consists of white blood cells, which can be used to determine the underlying germline DNA.

11. Centrifugation is necessary for thawed plasma due to cryoprecipitate, which could clog the DNA extraction columns.
12. During the cell lysis reaction of the QIAamp® Circulating Nucleic Acid Kit protocol, make sure the contents are vortexed to become homogenous
13. Do not pause the experiment at the lysis step of the nucleic acid isolation.
14. At the last elution step of the QIAamp® Circulating Nucleic Acid Kit protocol, centrifuge for 2 min (the protocol calls for 1 min).
15. During centrifugation, if the cap in the Eppendorf tube is left open, it risks picking up contaminants from other materials spun in the same centrifuge. Thus, remove the cap before spinning, and place a new cap on the tubes after the spin.
16. Sanger sequencing may be completed on FFPE and plasma samples prior to ddPCR to confirm that the amplicon of interest is present. This is a test to confirm that the amplicon is there but will likely detect only wild type. ddPCR will detect specific mutants, which will be present in much lower abundance.
17. Exact annealing temperature varies with primers. First, test the primers with a full-cycle temperature gradient PCR on genomic DNA to determine optimal annealing temperature.
18. Primer probe mix can be stored at –20 °C for later use.
19. Serial tenfold dilutions of the template allow for the creation of at least one concentration that appropriately saturates droplets to a concentration in the “digital” range, which may vary greatly depending on the target, but in this assay ranges from ~1000 to 5000 template-filled droplets out of ~15,000 total droplets. In prior experiences of working with patient samples, a dilution of approximately  $10^{-3}$  has led to concentrations in the digital range. The ideal dilution will vary by sample as well as each instance of the protocol, and is contingent on final elution volumes.
20. When aliquoting samples for ddPCR, start with water, then wild type, then mutant. This will minimize contamination.
21. Thoroughly vortex the ddPCR supermix before pipetting. Do the same before pipetting the mix into the ddPCR cartridge. The contents of the mix tend to settle out of solution, so this step is important for adequate distribution of contents.
22. All wells in the middle and bottom row must be filled in order for droplets to be generated.
23. Avoid air bubbles when aliquoting into ddPCR cartridge. This will disrupt adequate droplet formation.

24. Oil should be added only after sample has been added to all 8 wells—adding oil first will compromise droplet quality.
25. All oil must be at the same level. If air gets into the system, the droplet generation terminates.
26. After adding oil, proceed to droplet generation within 2 min.
27. When suctioning droplets from the uppermost row of the cartridge into the 96-well semiskirted plate, take the utmost care to suction slowly and evenly. This will help to prevent the shearing of droplets, maximizing the number of droplets per reaction well, and therefore improving the utility of the ddPCR assay. Once thermocycled, droplets are no longer fragile.
28. If the incorrect dyes are selected before the run, it is possible to correct the information after the run.
29. Poisson statistics are used to calculate the copies/ $\mu\text{L}$  concentration metric based on the formula:  
$$\# \text{of copies} = -\ln \left( 1 - \frac{\text{Positives}}{\text{Total counted}} \right) \times \text{Total counted}$$
This formula yields the expected value of copies in the portion of a sample that was counted based on the number of droplets filled with the amplicon of interest (*positives*) and the number of total droplets counted (*total counted*). In the QuantaSoft analysis software, this result is expressed as copies/ $\mu\text{L}$  which can be related back to the amount of sample placed into the 20  $\mu\text{L}$  ddPCR reaction to determine the number of copies of mutant or wildtype targets quantified in the reaction.
30. The figures below demonstrate the effect of varying levels of sample abundance on the calculated concentration of the amplicon of interest. The top graph shows total formed droplets for ddPCR in green, with droplets filled with the amplicon of interest in blue, at three concentrations of template DNA (serial twofold dilutions) and with water control. The bottom graph illustrates the corresponding calculated concentrations (copies/ $\mu\text{L}$ ) of the amplicon of interest, in this case the ligand-binding domain of ESR1 ex10. As can be seen, there is concordance between an approximate 50% reduction in the concentration of template DNA at dilutions of  $1.25 \times 10^{-6}$  to  $6.25 \times 10^{-7}$ , and the resulting approximate 50% reduction in calculated concentration of the amplicon of interest. However, at a sample abundance where there is near full saturation of generated droplets, such as a dilution of  $2.5 \times 10^{-6}$  in this experiment, the calculated concentration of the amplicon of interest is less reliable. This may be explained by the fact that the former dilutions are closer to the digital range, defined as saturation of a few thousand droplets of the approximately 15,000–20,000 that are generated in total, where the concentration of the amplicon of interest may be more reliably measured (Fig. 5).



**Fig. 5** Top: Bar graph of ddPCR analysis of varying concentrations of samples. The number of droplets positive is indicated at the top of each bar (blue—droplets containing mutated DNA, green—total number of droplets analyzed). Water is used as a negative control. Bottom: Increasingly diluted concentrations of genomic DNA entered into assay

31. Tissues are prone to DNA contamination due to the normal processing of FFPE blocks. Fresh microtome blades should be used for each specimen, and if possible, it is preferable to discard the first 2–3 slices.
32. When using the “Pinpoint Slide DNA Isolation System,” use a pipette tip to swab and spread the pinpoint solution over the area of interest.

## References

1. American Cancer Society (2014) Breast cancer facts & figures 2013–2014. American Cancer Society Inc., Atlanta
2. Eckhardt BL, Francis PA, Parker BS, Anderson RL (2012) Strategies for the discovery and development of therapies for metastatic breast cancer. *Nat Rev Drug Discov* 11:479–497
3. Toy W, Shen Y, Won H et al (2013) ESRI ligand-binding domain mutations in hormone-resistant breast cancer. *Nat Genet* 45:1439–1445
4. Cancer Genome Atlas Network (2012) Comprehensive molecular portraits of human breast tumours. *Nature* 490:61–70
5. Robinson DR, Wu YM, Vats P et al (2013) Activating ESRI mutations in hormone-resistant metastatic breast cancer. *Nat Genet* 45:1446–1451
6. Bose R, Kavuri SM, Searleman AC et al (2013) Activating HER2 mutations in HER2 gene amplification negative breast cancer. *Cancer Discov* 3:224–237

7. Gerlinger M, Rowan AJ, Horswell S et al (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 366:883–892
8. Higgins MJ, Jelovac D, Barnathan E et al (2012) Detection of tumor PIK3CA status in metastatic breast cancer using peripheral blood. *Clin Cancer Res* 18:3462–3469
9. Swarup V, Rajeswari MR (2007) Circulating (cell-free) nucleic acids – a promising, non-invasive tool for early detection of several human diseases. *FEBS Lett* 581:795–799
10. Hodgson DR, Wellings R, Orr MC et al (2010) Circulating tumour-derived predictive biomarkers in oncology. *Drug Discov Today* 15:98–101
11. Casciano I, Vinci AD, Banelli B et al (2010) Circulating tumor nucleic acids: perspective in breast cancer. *Breast Care (Basel, Switzerland)* 5:75–80
12. Beaver JA, Jelovac D, Balukrishna S et al (2014) Detection of cancer DNA in plasma of patients with early-stage breast cancer. *Clin Cancer Res* 20:2643–2650
13. El Messaoudi S, Rolet F, Mouliere F, Thierry AR (2013) Circulating cell free DNA: preanalytical considerations. *Clin Chim Acta* 424:222–230
14. Oshiro C, Kagara N, Naoi Y et al (2015) PIK3CA mutations in serum DNA are predictive of recurrence in primary breast cancer patients. *Breast Cancer Res Treat* 150:299–307



# Chapter 14

## Droplet Digital PCR for Minimal Residual Disease Detection in Mature Lymphoproliferative Disorders

Daniela Drandi, Simone Ferrero, and Marco Ladetto

### Abstract

Minimal residual disease (MRD) detection has a powerful prognostic relevance for response evaluation and prediction of relapse in hematological malignancies. Real-time quantitative PCR (qPCR) has become the settled and standardized method for MRD assessment in lymphoid disorders. However, qPCR is a relative quantification approach, since it requires a reference standard curve. Droplet digital<sup>TM</sup> PCR (ddPCR<sup>TM</sup>) allows a reliable absolute tumor burden quantification withdrawing the need for preparing, for each experiment, a tumor-specific standard curve. We have recently shown that ddPCR has a good concordance with qPCR and could be a feasible and reliable tool for MRD monitoring in mature lymphoproliferative disorders. In this chapter we describe the experimental workflow, from the detection of the clonal molecular marker to the MRD monitoring by ddPCR, in patients affected by multiple myeloma, mantle cell lymphoma and follicular lymphoma. However, standardization programs among different laboratories are needed in order to ensure the reliability and reproducibility of ddPCR-based MRD results.

**Key words** Minimal residual disease (MRD), Follicular lymphoma (FL), Mantle cell lymphoma (MCL), Multiple myeloma (MM), Immunoglobulin heavy-chain gene (IGH), t(14;18) translocation, t(11;14) translocation, BCL2/IGH, BCL1/IGH, Droplet digital PCR

---

### 1 Introduction

Minimal residual disease (MRD) monitoring in hematology is defined as any approach aimed at detecting and possibly quantifying residual tumor cells beyond the sensitivity level of routine imaging and laboratory techniques. Whenever a patient achieves complete clinical remission a number of different scenarios might actually take place including full eradication of the neoplastic clone, long-term persistence of quiescent or nonclonogenic or immunologically regulated tumor cells or persistence of clonogenic cells capable of giving rise to a full clinical relapse within months or years. MRD analysis is currently a valuable early predictor of outcome in several lymphoid malignancies, having made a considerable impact on clinical research. Actually, the persistence of detectable MRD after

a successful treatment heralds the disease recurrence and thus MRD monitoring might drive preemptive treatments, aimed at avoiding or delaying the onset of clinical relapse [1–8].

Different technologies have been used for MRD study. Molecular methods—including polymerase chain reaction (PCR), real-time quantitative PCR (qPCR), and the most recently introduced next-generation sequencing (NGS)—are widely employed for MRD monitoring in acute lymphoblastic leukemia (ALL) and lymphomas [9–19], while multicolor flow cytometry (MFC) has a major role for MRD detection in chronic lymphocytic leukemia (CLL) and multiple myeloma (MM) [20–22]. Although nowadays qPCR is a well-established tool for MRD monitoring in lymphoproliferative disorders, it has the main disadvantage of relying upon a serial dilution standard curve for target quantification. One of the latest evolutions of PCR, droplet digital PCR (ddPCR), has introduced several practical advantages, compared to qPCR, most notably its absolute quantification nature does not require a reference standard curve [23–33]. Moreover, we have recently shown that ddPCR has a good concordance with qPCR with respect to sensitivity, accuracy, and reproducibility, in the context of MRD evaluation in mature lymphoproliferative disorders [34]. The greater applicability and reduced labor intensiveness, compared to qPCR, recommend ddPCR as an attractive alternative method for MRD assessment.

In this chapter, we firstly provide a technical roadmap for a suitable clonal-specific tumor marker identification from circulating tumor cells. Then we describe our ddPCR-based MRD detection approach based on different target types: (1) the immunoglobulin heavy chain gene rearrangement (IGH) for MM and mantle cell lymphoma (MCL), (2) the BCL1/IGH fusion gene, derived from the chromosomal translocation t(11;14), for MCL, and (3) the BCL2-MBR/IGH gene, arising from the chromosomal translocation t(14;18), for follicular lymphoma (FL).

Finally, we point out that ddPCR prognostic value still needs to be validated in the context of prospective clinical trials. Most importantly, before entering into the clinical practice, the promise of ddPCR, compared to qPCR and MFC, needs to be proved in the context of standardization programs. It is part of the current activities of the Euro-MRD group, a division of the European Scientific foundation for Laboratory Hemato Oncology (ESLHO), to test and validate these ddPCR assay at the inter-laboratory level, as has already been done several years ago for qPCR. Since this is currently a work in progress, the ddPCR procedures described in this manuscript are likely to be updated in the not too distant future.

## 2 Materials

### 2.1 Lab Workflow Requirements

PCR protocols need appropriate caution to safeguard from contaminations. This aspect is particularly important in ddPCR, in which even a single contaminant molecule could be detected. To avoid contamination it is crucial to use clean supplies (i.e., gloves, racks, tubes, plates, etc.) and dedicated pipettes (*see Note 1*). Therefore, before PCR experiment setup, verify the following workspace availability to prevent the risk of DNA contamination:

- **Pre-PCR zone:** dedicated to mix preparation. Keep any DNA or PCR products away from this area.
- **DNA zone:** dedicated to adding template and droplet generation. Placed in a room separate from the pre-PCR zone.
- **Post-PCR zone:** dedicated to PCR amplification, PCR product analysis and droplet reading.

### 2.2 Instruments and Software

1. Fume hood for agarose gel setup.
2. Vortexer and minicentrifuge.
3. Nanodrop UV-Vis spectrophotometer (Fisher Thermo Scientific, Waltham, MA, USA).
4. Benchtop centrifuge with microplate carriers and microcentrifuge.
5. Thermal Cycler with heated lid, i.e., T100 (Bio-Rad, Laboratories, Hercules, CA, USA) or thermal cycler with similar characteristics able to reach a ramp rate of 2.5 °C/s.
6. Horizontal electrophoresis apparatus and power supply (2–300 V/4–500 mA output voltage/current range).
7. Digital camera or any gel documentation and imaging systems.
8. DNA Sequencing facility for Sanger sequencing analysis.
9. Thermomixer (i.e., Eppendorf ThermoMixer C).
10. QX100<sup>TM</sup> or QX200<sup>TM</sup> ddPCR<sup>TM</sup> System (Bio-Rad, Laboratories, Hercules, CA, USA).
11. PXI<sup>TM</sup> PCR plate sealer (Bio-Rad, Laboratories, Hercules, CA, USA).
12. Multichannel pipettes and filtered pipette tips.
13. Chromas (free licenced software) for electropherogram visualization.
14. IMGT/V-QUEST tool ([www.imgt.org](http://www.imgt.org)) for IGH genes sequence analysis.
15. Blastn ([www.blast.ncbi.nlm.nih.gov](http://www.blast.ncbi.nlm.nih.gov)) for BCL2-MBR/IGH and BCL1/IGH genes sequence analysis.

16. Primerquest and Oligo Analyzer 3.1 tools ([www.eu.idtdna.com](http://www.eu.idtdna.com)) or Primer3Plus ([www.primer3plus.com](http://www.primer3plus.com)) for forward and reverse allele-specific oligonucleotide (ASO) primers and probes design.
17. QuantaSoft™ ddPCR analysis software (Bio-Rad, Laboratories, Hercules, CA, USA).

### **2.3 Reagents and Kits**

#### *2.3.1 Bone Marrow (BM) and Peripheral Blood (PB) Sample Processing*

1. Erythrocyte lysis buffer (0.155 M NH<sub>4</sub>Cl, 10 mM KHCO<sub>3</sub>, 0.1 mM ethylenediamine tetraacetic acid (EDTA), pH 7.2–7.4).
2. Genomic DNA (gDNA) extraction from circulating tumor cells can be performed by different methods such as DNAzol (Thermo Fisher Scientific, Waltham, MA, USA), NucleoSpin Tissue (Macherey-Nagel, Bethlehem, PA, USA), or automated systems, such as the Maxwell RSC (Promega, Madison, WI, USA).

#### *2.3.2 Tumor-Specific Molecular Marker Assessment*

1. Forward and reverse consensus primers for IGH [[16](#), [35](#), [36](#)], BCL1/IGH [[11](#)] and BCL2-MBR/IGH [[9](#)] PCR screening and sequencing (Table 1).
2. PCR reagents: GoTaq Flexi (5 U/μL), 5× Green GoTaq Flexi buffer, 25 mM MgCl<sub>2</sub> (Promega, Madison, WI, USA), dNTPs mix (2 mM) (nonspecific company), and nuclease-free water.
3. 1× TAE buffer (40 mM Tris-acetate, 1 mM EDTA, in distilled water, pH 8.0).
4. Agarose gel (2%).
5. Ethidium bromide 1% or RealSafe staining solution (Durviz, Spain).
6. 100 bp DNA mass ladder (i.e., φX174-HaeIII digest).
7. QIAquick gel extraction kit (Qiagen GmbH, Germany).
8. QIAquick PCR purification kit (Qiagen GmbH, Germany).
9. TOPO-TA cloning Kit (Thermo Fisher Scientific, Waltham, MA, USA).
10. ImMedia AmpBlue and ImMedia AMPLiquid agar for IPTG/X-gal plates and low-salt LB liquid medium for bacteria growing, respectively (Thermo Fisher Scientific, Waltham, MA, USA).
11. Wizard Plus SV minipreps DNA purification System (Promega, Madison, WI, USA).
12. gDNA (100 ng/μL) from a BCL2-MBR/IGH-positive cell line (such as DOHH-2 or RL).
13. gDNA (100 ng/μL) from a BCL1/IGH-positive cell line (such as JVM2).

**Table 1**

**Primer sequences for IGH, BCL1/IGH and BCL2-MBR/IGH screening by PCR or nested-PCR. Reverse primers are italicized**

Target	Primer	Sequence (5'-3')
IGH	VH1Fs	CAGGTGCAGCTGGTCARYCTG
	VH2Fs	CAGRTCACCTTGAAGGAGTCTG
	VH3Fs	GAGGTGCAGCTGGTGSAGTCYG
	VH4aFs	CAGSTGCAGCTGCAGGAGTCAG
	VH4bFs	CAGGTGCAGCTACARCACTGGG
	VH5Fs	GAGGTGCAGCTGKTCAGTCTG
	VH6Fs	CAGGTACAGCTGCAGCAGTCAG
	VH1D	CCTCAGTGAAGGTCTCCTGCAAGG
	VH2D	TCCTGCGCTGGTGAAAGCCACACA
	VH3D	GGTCCCTGAGACTCTCCTGTGCA
	VH4aD	TCGGAGACCCCTGTCACCTGCA
	VH4bD	CGCTGTCTGGTTACTCCATCAG
	VH5D	GAAAAAGCCCAGGGAGTCTGAA
	VH6D	CCTGTGCCATCTCCGGGGACAGTG
	<i>JHD</i>	<i>ACCTGAGGAGACGGTGACCAGGGT</i>
	VH1-FR2	CTGGGTGCGACAGGCCCCCTGGACAA
	VH2-FR2	TGGATCCGTCAGCCCCAGGGAGG
	VH3-FR2	GGTCCGCCAGGCTCCAGGGAA
	VH4-FR2	TGGATCCGCCAGCCCCAGGGAGG
	VH5-FR2	GGGTGCGCCAGATGCCGGGAAAGG
	VH6-FR2	TGGATCAGGCAGTCCCCATCGAGAG
	VH7-FR2	TTGGGTGCGACAGGCCCCCTGGACAA
	<i>JHC</i>	<i>CTTACCTGAGGAGACGGTGACC</i>
BCL1/IGH	BCL1-P2	GAAGGACTTGTGGTTGC
	BCL1-P4	GCTGCTGTACACATCGGT
	<i>JH3</i>	<i>ACCTGAGGAGACGGTGACC</i>
BCL2-MBR/IGH	MBR2	CAGCCTGAAACATTGATGG
	<i>JH3</i>	<i>ACCTGAGGAGACGGTGACC</i>
	MBR3	TATGGTGGTTGACCTTTAG
	<i>JH4</i>	<i>ACCAGGGTCCCTTGGCCCCA</i>

### 2.3.3 MRD Detection by ddPCR

1. Forward and reverse allele-specific oligonucleotide (ASO) primers and probes for individual IGH regions and for BCL1/IGH translocations. Consensus forward and reverse primers and probes for BCL2-MBR/IGH. See Table 2 and Fig. 6.
2. 2× ddPCR Supermix for Probes (No dUTP) (cat. 186-3024), DG8 cartridges (cat. 186-4008), DG8 gaskets (cat. 186-3009), ddPCR droplet generation oil for probes (cat. 186-3005), pierceable foil heat seals (cat. 181-4040), ddPCR droplet reader oil (cat. 186-3004) (Bio-Rad Laboratories, Hercules, CA, USA).

**Table 2**  
**Primers and probes sequences used in MRD based ddPCR**

Target	Oligos type	NAME	5'-SEQUENCE-3'
IGH	PRIMER forward	Pt code-F	Pt specific sequence derived from CDR2 region
	PRIMER reverse	Pt code-R	Pt specific sequence derived from CDR3 region
	PROBES	LVH1	GCACAGCCTACATGGAGCTGAGCAG
		MVH2	ACCACCTGGTTTTGGAGGTGTCCTT
		LVH3	TCCTCGGCTCTCAGGC
		MVH3TER	CTCTGGAGATGGTGAATCGGC
		MVH4	TGTCTGCAGCGGTACAGA
		MVH4TER	TGTCCGCAGCGGTACAGA
		LVH5	CTTCAGGCTGCTCCACTGCAGGTAG
		LVH1bis	TACATGGAGCTGAGCAGCCTGA
		LVH4	CAACCCCTCCCTCAAAGAGTC
		VH3DD3	GATTACCATCTCCAGAGACA
		DVH4	TGTTTGCAGCGGTACAGA
		VH3MG	CGGCCCTCACGGAGTCT
	PRIMER forward	Pt code-F	Pt specific sequence derived from “N” region
BCL1/IGH	PRIMER consensus	JH3	ACCTGAGGAGACGGTGACC
	PROBES	JH1-4-5	ACCTGGTCACCGTCTCCTCAGGTG
		JHDD1	ACGTCTGGGGCAAAGGGACACGG
		JHG2	CGATCTCTGGGGCGTGGCAC
BCL2-MBR/ IGH	PRIMER forward	Mbr2/Q	CTATGGTGGTTGACCTTAGAG
	PRIMER reverse	JH32-short	CCTGAGGAGACGGTGACC
	PROBE	BCL2	CTGTTAACACAGACCCACCCAGAG

3. 96-well PCR plates and optical adhesive films or 0.2 mL strip tubes with cups (used for mix preparation and collection before droplets generation).
4. Scotch tape.
5. Hard-shell high-profile 96-well semiskirted PCR plates (cat. 0030 128.575) (Eppendorf, Hamburg, Germany).
6. gDNA (100 ng/µL) from a BCL2-MBR/IGH-positive cell line (such as DOHH-2 or RL).
7. gDNA (100 ng/µL) from buffy coat (BC), pooled from 5 to 10 healthy donors, to use as negative control in IGH and BCL1/IGH ddPCR reactions.
8. gDNA (100 ng/µL) from a BCL2-MBR/IGH-negative sample as MCF-7 human breast cancer cell line, a chemotreated patients or any subject previously assessed by PCR to be BCL2-MBR/IGH negative, since this translocation can be also present in healthy donors without lymphoma [37].

### 3 Methods

Methods described below provide a technical support for clonal-specific tumor marker identification and ddPCR-based MRD monitoring in MM, MCL and FL. Figure 1 shows the schematic experimental workflow, for each disease, summarizing the steps from sample acquisition through patient's unique tumor marker determination to MRD quantification by ddPCR.

#### **3.1 Bone Marrow and Peripheral Blood Samples Processing**

1. For sample collection, the use of sodium citrate or EDTA is preferred (*see Note 2*) over Heparin.
2. Resuspend blood samples in erythrocyte lysis buffer (NH<sub>4</sub>Cl) (dilute BM 1:4 and PB 1:2). Leave blood tubes for 15 min at room temperature (lying flat in the dark) then centrifuge for 10 min at 450 ×  $\text{g}$  at room temperature. Discard the supernatant wash in NH<sub>4</sub>Cl and centrifuge for 10 min at 450 ×  $\text{g}$  at room temperature. Remove supernatant, resuspend cell pellets in PBS or 0.9% NaCl count and dispense 5–10 × 10<sup>6</sup> cells in 1.5 mL tubes; centrifuge for 1 min at 13000 ×  $\text{g}$  and discard supernatant. Cells can now be stored indefinitely, as dried pellets, at –80 °C for further gDNA extraction.
3. Perform gDNA extraction with a commonly used method or kit.
4. Estimate the gDNA quality and concentration before experimental use (*see Note 3*).
5. Prepare 200–300 μL of 100 ng/μL stock solution and store at –20 °C (*see Note 4*).

#### **3.2 Tumor-Specific Molecular Marker Identification**

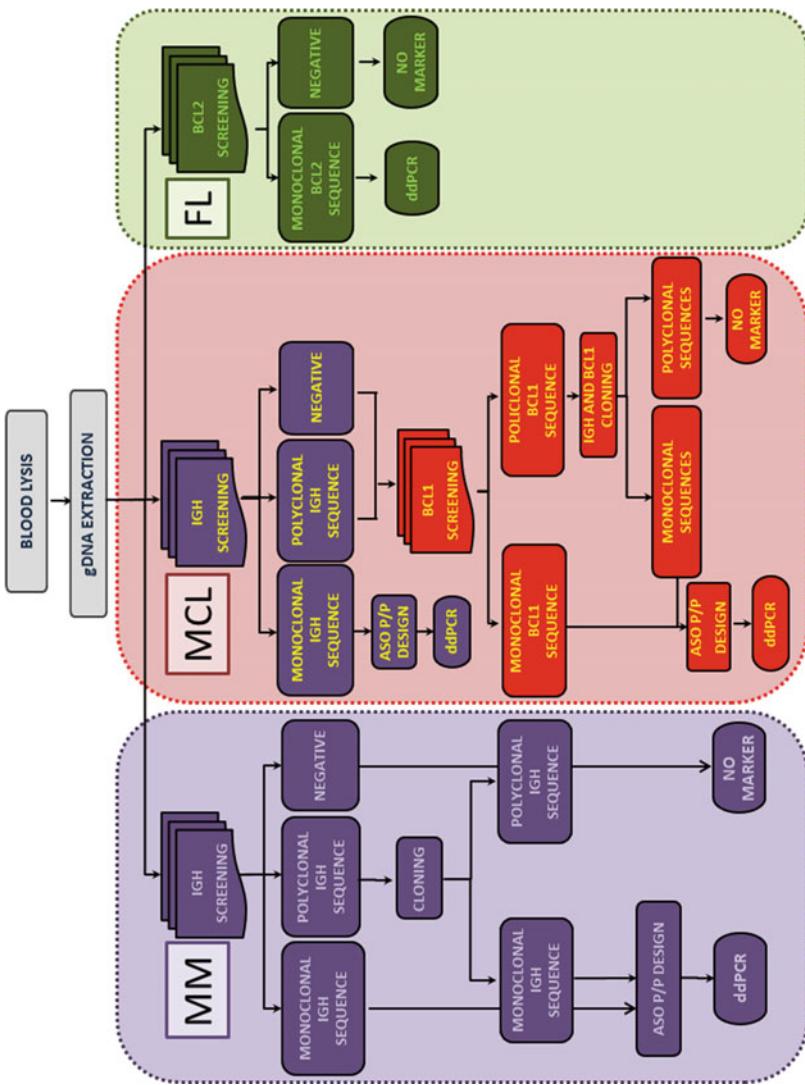
Be aware that samples containing <5% of tumor cells present a lower rate of success in marker identification (especially for IGH sequencing) by Sanger sequencing. Therefore, diagnostic samples containing a high percentage of tumor cells (>5%) are required.

Multiple strategies to obtain the tumor-specific marker are currently available [15–17].

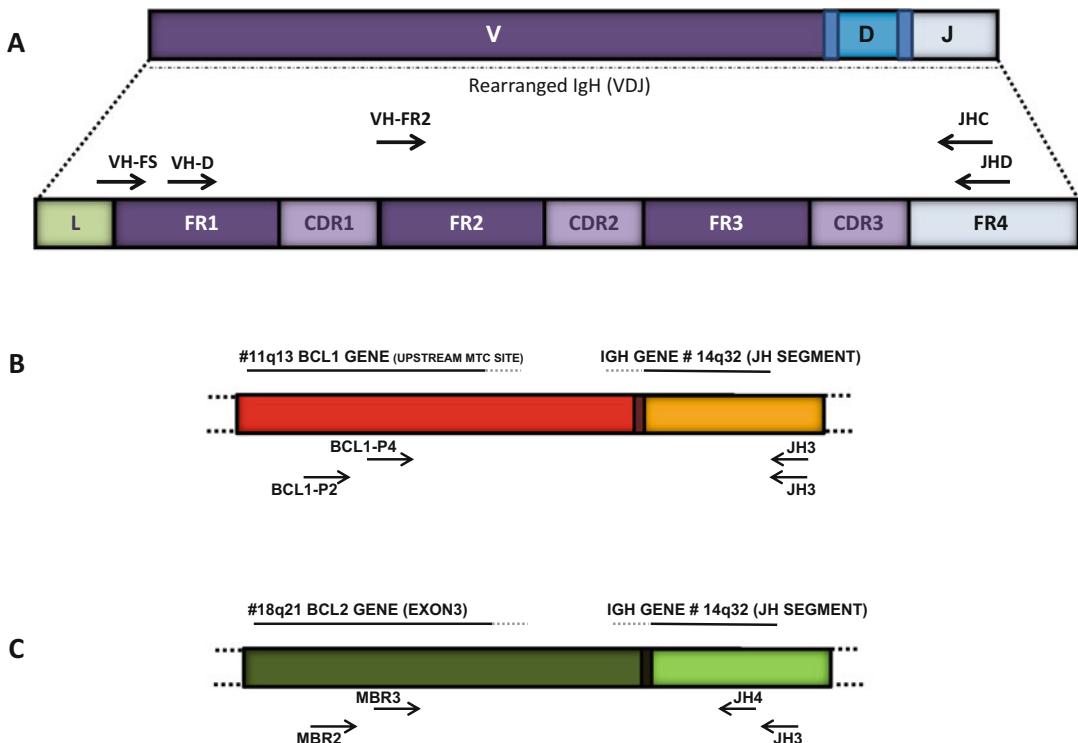
Here we present a PCR based approach (Fig. 2a–c), established as previously published [9, 11, 36], to identify and to direct sequencing:

##### **3.2.1 IGH Clonal Rearrangement Screening at Baseline**

1. IGH rearrangements in MM and MCL.
2. BCL1/IGH gene translocation, in MCL (if IGH rearrangement is not successfully identified).
3. BCL2-MBR/IGH gene translocation, in FL.
1. The strategy used for IGH screening adopts three sets of consensus forward primers (VH-FS, VH-D and VH-FR2) derived from framework regions (FR1, FR2), each set is



**Fig. 1** Experimental workflow for MRD monitoring by ddPCR. For each type of disease, steps from sample acquisition, tumor marker screening and MRD quantification by ddPCR are described. Albeit MCL could harbor both IGH and BCL1/IGH tumor markers, the acceptance of only one target for MRD monitoring depends on the protocol and clinical trial purpose (ref. 11, 16). Here we describe the approach for MRD monitoring on patient's unique tumor marker where BCL1/IGH is performed only if IGH rearrangement is not successfully identified. Samples in which all steps fail to identify a clonal target sequence ("no marker" patients), can not be monitored for MRD studies. gDNA (genomic DNA), BCL1 (BCL1/IGH), BCL2 (BCL2/IGH), ASO P/P (Allele Specific Oligonucleotide Primers/ Probe), multiple myeloma (MM), mantle cell lymphoma (MCL), and follicular lymphoma (FL)



**Fig. 2** Schematic representation of PCR strategies used, for marker identification, for: (a) IGH rearrangements screening in MM and MCL, (b) BCL1/IGH gene translocation, in MCL and (c) BCL2-MBR/IGH gene translocation in FL, performed on diagnostic sample. IGH rearrangement derives from the combination of V-(D)-J segments that code for the variable domains of the IGH molecules. In both BCL1/IGH and BCL2-MBR/IGH translocations the IGH locus at 14q32.3 is involved and the BCL1 gene from 11q13 and BCL2 region from 18q21 are translocated, respectively, to the IGH locus. IGH gene regions: L (leader), F (framework), CD (complementary determining)

combined separately with the appropriate JH antisense consensus primer (JHD or JHC; Fig. 2a). We suggest performing a first PCR screening by individually combining each of the VH-FS family primers with the reverse JHD primer. In case of no amplification, switch to the VH-D and then to VH-FR2 set of primers combined one-at-a-time with the JHD or JHC primers, respectively (Table 1) [10, 16, 35, 36].

2. In the Pre-PCR zone, prepare a 45 µL PCR mix solution in 0.2 mL tubes. Figure 3 shows an example of how to prepare the PCR reaction mix for the IGH screening with VH-FS primers.
3. In the DNA zone, add 5 µL of gDNA (100 ng/µL) in each of the patient and control tubes. Before adding the gDNA to the mix, ensure that the DNA is thoroughly mixed. Vortex, spin down, and then pipet the gDNA sample up and down a few times.

IGH REARRANGEMENT	
Reagents	1 Reaction ( $\mu$ l)
dNTPs mix (2mM)	5
5x GoTaq Flexi Green Buffer	10
MgCl <sub>2</sub> (20mM)	5
GoTaq Flexi 5U/ $\mu$ l (Promega)	0,25
VH-FS (10pmol/ $\mu$ l)	1
JHD (10pmol/ $\mu$ l)	1
H <sub>2</sub> O	22,75
TOT mix	45
gDNA	5
TOT	50

	Sample	gDNA code	VH-F primer
1	CTR+	gDNA or PL	VH-specific
2	pt1		VH-FS1
3	pt1		VH-FS2
4	pt1		VH-FS3
5	pt1		VH-FS4a
6	pt1		VH-FS4b
7	pt1		VH-F5
8	pt1		VH-F6
9	CTR-	H <sub>2</sub> O	all VH-F primers

Initial denaturation: 94°C for 1 minute

Denaturation: 94°C for 30 seconds

Annealing: 62°C for 30 seconds

Extension: 72°C for 30 seconds

Final extension: 72°C for 10 minutes

}

repeated for 33 cycles

**Fig. 3** PCR reaction mix and setup for the IGH screening with VH-FS primers. The IGH strategy adopts three sets of consensus forward primers (VH-FS, VH-D, and VH-FR2) each set combined separately with the appropriate JH antisense consensus primer (JHD or JHC). The PCR reaction mix here described combines each of the VH-FS family primers with the reverse JHD primer. In case of no amplification or no monoclonal IGH gene identification, the VH-D and then the VH-FR2 set of primers are combined one-at-a-time with the JHD or JHC primers, respectively, *see Table 1* for primers sequence. (VH-FS (VH-FS forward primer), gDNA (genomic DNA), CTR+ (positive control), PL (plasmid), CTR- (negative control), pt (patient)

4. Perform PCR with thermal cycling conditions and control DNA's as shown in Fig. 3.
5. Analyze the amplification product by running 10  $\mu$ L of PCR product on a 2% agarose gel, together with a DNA molecular weight marker (*see Note 5*).
6. Visualize with UV light and capture the image by a digital photo camera, in order to have documentation of the experiment result (*see Note 6*). The IGH product size must be between 300 and 500 bp.
7. In case of lack of amplification, repeat the PCR using a new set of primers (i.e., VH-D and VH-FR2; *see Note 7*). If no amplification is observed with all three sets of primers, no MRD analysis based on this molecular marker can be performed. In case of MCL sample, move to Subheading 3.2.3 for BCL1/IGH screening. In case of MM, the patient is marked as "IGH negative: no marker available for MRD analysis".
8. If the PCR product appears on the gel as a unique band, the whole volume of amplified gDNA can be directly purified by "QIAquick PCR purification kit," following the manufacturer's instructions. Otherwise, use the "QIAquick gel extraction kit." In this case, DNA containing the proper amplification band must be sliced from the gel, weighed on a scale, put in a

- 1.5 mL microcentrifuge tube and stored at 4 °C until the extraction. (It is recommended to store no longer than for 1 day).
9. Direct Sanger sequencing from both strands, with the same primers used in the successful PCR reaction described above can now be performed. To check the validity of the sequence, visualize the electropherogram by “Chromas” or other known software (*see Note 8*).
  10. In case of identification of a monoclonal sequence, proceed with sequence analysis using the IMGT/V-QUEST database ([www.imgt.org](http://www.imgt.org)). IMGT allows you to verify that your sequence is an IGH and to display the gene regions where primers and probe must be depicted (Subheading 3.3.1).
  11. If the sequence is polyclonal, proceed with IGH cloning (Subheading 3.2.2; *see Fig. 1*). This circumstance is more frequent in MM samples, mostly due to their high rate of somatic hypermutation, which may significantly reduce primer efficiency.

### 3.2.2 Cloning of Patient-Specific Rearrangement

To be used when direct Sanger sequencing does not allow identification of a clonal sequence. For MCL patients, cloning should be performed only in cases where both IGH and BCL1/IGH have been screened and a polyclonal sequence has been identified.

1. Set up the ligation reaction on ice using the TA-TOPO cloning kit: Combine 1 μL of TA-TOPO Cloning vector, 1 μL of Salt solution, 4 μL of fresh patient-specific PCR product. Mix together well. Incubate at room temperature for 30 min and, if not immediately used for bacterial transformation, store at –20 °C.
2. Thaw competent E. coli (stored –80 °C) on ice, add 3 μL of ligation product to competent cells, shake gently and incubate for 30 min, on ice.
3. Heat-shock at 42 °C for 30 s on a Thermomixer (without shaking) and immediately put on ice for 2 min.
4. Add 250 μL of S.O.C Medium (supplied with the TA-TOPO cloning kit), incubate for 60 min at 37 °C in a thermomixer under vigorous shaking (800–1000 rpm).
5. Spread 50–70 μL of transformed cells on agar petri dish, previously prepared with “ImMedia Amp blue” (prepared following the vendor’s instructions). Air-dry for few minutes. Incubate at 37 °C overnight (*see Note 9*).
6. Grow up the transformed cells (10–15 white colonies) into 1.5 mL of “ImMedia Amp Liquid” incubating overnight at 37 °C under vigorous shaking (800–1000 rpm) on a thermomixer. The next day, perform a PCR amplification using 2 μL of

the overnight grown transformed cells, adjusting the volume of PCR mix with 25.75 µL of water, instead of 22.75 µL. In order to verify the presence of the cloned rearrangement the same primers, and the same conditions previously used at the screening, must be used. Run on a 2% gel and verify which plasmids contain product with the desired molecular weight band.

7. Perform plasmid DNA isolation and purification (Wizard Plus SV Miniprep DNA purification system, Promega, Madison, WI) from those plasmids showing similar amplification size.
8. Check the sequence of at least ten different plasmids. Sequencing only one strand from the forward side is usually enough for good sequence comparison. If an identical sequence shows up in at least three different plasmids, this sequence can be considered to represent the tumor clone [12]. Then, proceed with sequence analyses by IMGT (for IGH) or Blastn (for BCL1/IGH) and the design of clone-specific primers and probe (*see* Subheading 3.3.1).

### 3.2.3 BCL1/IGH

#### *Translocation Screening at Diagnosis*

MCL patients not successfully screened for an IGH rearrangement are then screened for the BCL1/IGH translocation, by a semi-nested PCR approach (*see* Fig. 2b). This consists of a first PCR amplification using a sense 5' primer (BCL1-P2) located in the major translocation cluster (MTC) upstream of the BCL1 gene, followed by a second PCR reaction performed with a sense 5' primer (BCL1-P4) located 43 bp downstream from the BCL1-P2 primer. Both amplifications employ an antisense 3' joining region (JH3) consensus primer [11].

1. In the Pre-PCR zone, prepare a 45 µL mix solution in 0.2 mL tubes, as shown in Fig. 4, using the 1st amplification primer pair specified (BCL1-P2/JH3).
2. In the DNA zone, add 5 µL of gDNA (100 ng/µL). Before adding the gDNA to the mix, ensure that the DNA is thoroughly mixed. Vortex, spin down, and then pipet the gDNA sample a few times.
3. Run 1st round PCR under the conditions shown in Fig. 4.
4. Set up the 2nd round PCR: in the Pre-PCR zone, prepare a 48 µL mix solution as shown in Fig. 4, (adjusting the mix with 25.75 µL of water, instead of 22.75 µL), using the 2nd amplification primer pair (BCL1-P4/JH3).
5. In the post PCR-zone, add 2 µL of 1st amplification PCR product.
6. Run the 2nd round PCR at the thermal cycling conditions shown in Fig. 4.

BCL1/IGH		
Reagents	1st round PCR ( $\mu$ l)	2nd round PCR ( $\mu$ l)
dNTPs mix (2mM)	5	5
5x GoTaq Flexi Green Buffer	10	10
MgCl <sub>2</sub> (15mM)	5	5
GoTaq Flexi 5U/ $\mu$ l (Promega)	0.25	0.25
BCL1-MTC (10pmol/ $\mu$ l)	1	1
JH (10pmol/ $\mu$ l)	1	1
H <sub>2</sub> O	22.75	25.75
TOT mix	45	48
gDNA	5	2
TOT	50	50

	Sample	gDNA code
1	CTR+	JVM-2 ( $10^2$ )
2	Pt1	
...	...	
...	Ptn	
n	CTR-	H <sub>2</sub> O

Primers	BCL1-P2/JH3	BCL1-P4 /JH3

Denaturation: 94°C for 60 seconds

Annealing: 58°C for 30 seconds

Extension: 72°C for 30 seconds

repeated for 33 cycles (1st round PCR)

repeated for 30 cycles (2nd round PCR)

Final extension: 72°C for 10 minutes

**Fig. 4** BCL1/IGH rearrangement reaction and setup. BCL1-MTC (forward primer, BCL1-P2 or BCL1-P4), JH: reverse primer (JH3 or JH4), gDNA (genomic DNA), CTR+ (positive control), CTR- (negative control), pt. (patient), JVM-2 (BCL1/IGH positive cell line)

7. Follow the steps 5, 6, 8, and 9 described in Subheading 3.2.1. BCL1/IGH PCR product size should be in the size range between 100 and 300 bp (see Note 7).
8. If the sequence electropherogram is monoclonal, proceed with sequence analysis using Blastn ([www.blast.ncbi.nlm.nih.gov](http://www.blast.ncbi.nlm.nih.gov)) to verify the BCL1/IGH translocation sequence and proceed with primers and probe design as described in Subheading 3.3.1, otherwise proceed with cloning as above (see Subheading 3.2.2).

### 3.2.4 BCL2-MBR/IGH Translocation Screening at Diagnosis

A number of variant translocations involving the BCL2 gene have been described; however only MBR translocations (Major Breakpoint Region), and to a lesser extent mcr (minor cluster region) translocations, have been extensively exploited for MRD analysis [6, 9, 38]. Thus, only a subset of FL patients, ranging from 50% to 65%, can currently be assessed using the BCL2/IGH nested PCR approach [6, 39, 40]. The procedure here described is focused only on the most common BCL2-MBR/IGH translocation (Fig. 2c). However, if no clonal signal is observed for the BCL2-MBR/IGH translocation, we suggest trying with other primers sets for minor breakpoint region detection [38].

1. In the Pre-PCR zone, prepare a 45  $\mu$ L mix solution, in 0.2 mL tubes, as shown in Fig. 5, using the BCL2-MBR2/JH3 primer pair.
2. In the DNA zone, add 5  $\mu$ L of gDNA (100 ng/ $\mu$ L). Before adding the gDNA to the mix, ensure that the DNA is thoroughly mixed. Vortex, spin down, and then pipet the gDNA sample few times.
3. Follow 1st round PCR thermal cycler conditions as shown in Fig. 5 (27 cycles).
4. In the Pre-PCR zone, set up the 2nd round PCR: prepare a 45  $\mu$ L mix solution, as shown in Fig. 5, using the BCL2-MBR3/JH4 primer pair.
5. In the post PCR-zone, add 5  $\mu$ L of 1st amplification product and run the 2nd round PCR (Fig. 5).
6. Follow the steps 5, 6, 8, and 9 as described in Subheading 3.2.1. BCL2-MBR/IGH PCR product size should range between 150 and 400 bp (*see Note 7*).
7. Proceed with Blastn ([wwwblast.ncbi.nlm.nih.gov](http://wwwblast.ncbi.nlm.nih.gov)) to verify the BCL2-MBR/IGH translocation sequence.

### 3.3 MRD Monitoring by ddPCR

#### 3.3.1 Clone-Specific Primers and Probe Design

ddPCR monitoring is performed based on the nucleotide sequence of the clonal tumor-specific marker. Here we propose one strategy for marker identification albeit other approaches, as gene scanning or next generation sequencing, are now available [16, 17].

The junctional “N” regions, located in IGH rearrangements (Fig. 6a), as well as in the BCL1/IGH (Fig. 6b) and BCL2-MBR/IGH rearrangements (Fig. 6c), represent a “fingerprint” of the lymphoma cell, and therefore their correct characterization is crucial for tumor-specific primers and probe design and for patient-specific MRD monitoring by ddPCR.

Particularly for the IGH and BCL1/IGH gene rearrangements, based on the sequences of interest, primers and probe should be designed following the widely known common recommendations [12, 15] (i.e., optimal melting temperature (Tm) between 57 and 62 °C, primers length between 13 and 26 mer, % CG content between 30 and 70%, probes with a Tm 6–8 °C higher than the primers etc.). Moreover, we suggest PrimerQuest and Oligo Analyzer 3.1 ([www.eu.idtdna.com](http://www.eu.idtdna.com)) or Primer3Plus ([www.primer3plus.com](http://www.primer3plus.com)) as helpful tools for assay design and optimization.

Different primer and probe design strategies need to be performed for each target:

1. The strategy for the IGH rearrangement-based MRD monitoring, here described, is named “ASO primer” and consists in as follows: based on the IGH sequence, tumor-specific primers

BCL2-MBR/IGH	
Reagents	1 Reaction ( $\mu$ l)
dNTPs mix (2mM)	5
5x GoTaq Flexi Green Buffer	10
MgCl <sub>2</sub> (20mM)	5
GoTaq Flexi 5U/ $\mu$ l (Promega)	0.25
BCL2-MBR (20pmol/ $\mu$ l)	1
JH (20pmol/ $\mu$ l)	1
H <sub>2</sub> O	22.75
	TOT mix 45
gDNA (1st PCR) or 1st PCR product (2ndPCR)	5
	TOT 50

	Sample	gDNA code
1	CTR+	DOHH2 (10 <sup>2</sup> )
2	Pt1	
...	...	
...	Ptn	
n	CTR-	H <sub>2</sub> O

1st amplification	BCL2-MBR2/JH3
2nd amplification	BCL2-MBR3/JH4

### **1st PCR reaction**

Denaturation: 94°C for 60 seconds  
 Annealing: 55°C for 60 seconds }  
 Extension: 72°C for 60 seconds } repeated for 27 cycles  
 Final extension: 72°C for 10 minutes

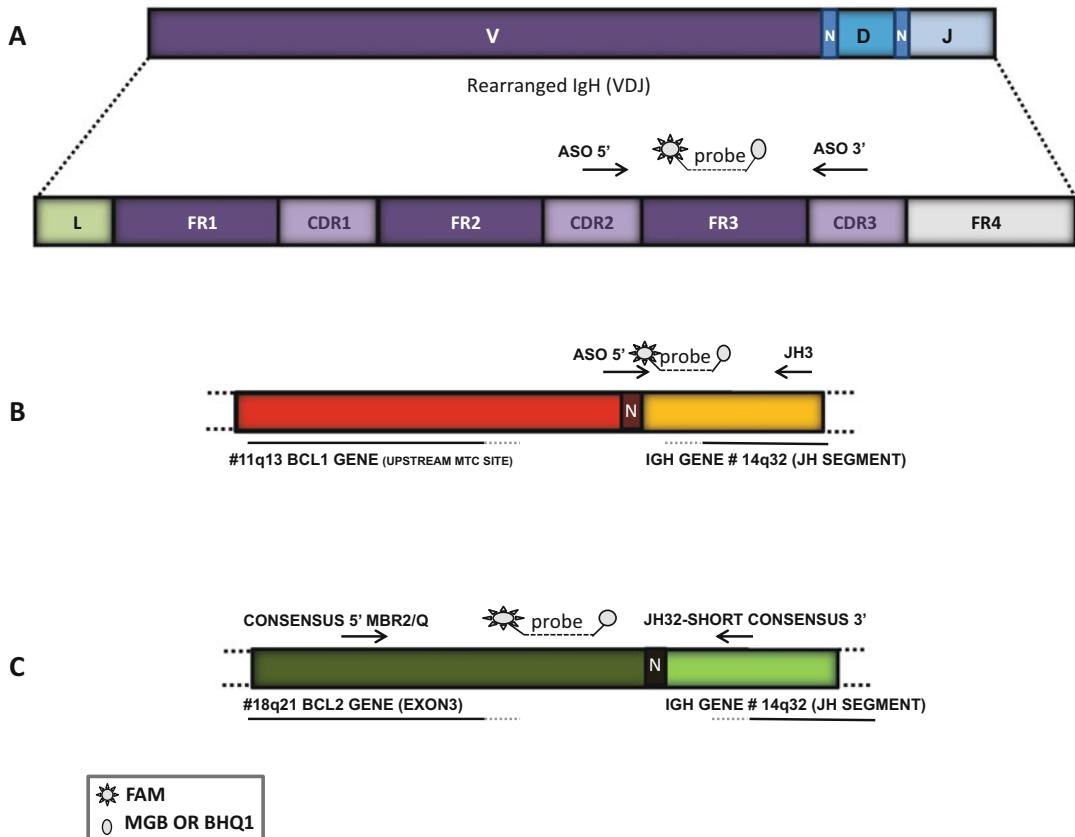
### **2nd PCR reaction**

Denaturation: 94°C for 60 seconds  
 Annealing: 58°C for 60 seconds }  
 Extension: 72°C for 60 seconds } repeated for 30 cycles  
 Final extension: 72°C for 10 minutes

**Fig. 5** BCL2-MBR/IGH rearrangement reaction and setup. BCL2-MBR (forward primer, BCL2-MBR3 or BCL2-MBR4), JH (reverse primer JH3 or JH4), gDNA (genomic DNA), CTR+ (positive control), CTR- (negative control), pt (patient), DOHH2 (BCL2-MBR/IGH positive cell line)

are positioned on the most hypervariable areas located in the IGH-V second and third complementary-determining region (CDR2-CDR3), while the consensus probe is designed on the most constant FR3 region (Fig. 6a) and can be employed for a large proportion of patients [2, 7, 8, 12, 41] (see Note 10). In our ongoing MRD studies, we routinely and successfully use a panel of consensus probes listed in Table 2 that are suitable for a high proportion of patients. However, in cases with frequent somatic mutations in the IGH gene rearrangements, a completely sequence-based design of primers and probe is recommended.

2. For BCL1/IGH primer and probe design (Fig. 6b): a forward patient-specific ASO primer, depicted within the junctional “N” insertion or overlapping the breakpoint in the BCL1 sequence and should be used in combination with the same JH primer used in the seminested PCR for IGH screening



**Fig. 6** Schematic representation of strategies used for designing: (a) IgH, (b) BCL1/IgH, and (c) BCL2-MBR/IgH primers and probes for MRD analysis by ddPCR. N (N region with target specific nucleotide insertion), IgH gene regions: L (leader), F (framework), CD (complementary determining), all strategies use FAM or BHQ1 probes

(JH3). In order to reach high specificity, the probe is positioned in the JH region adjacent to the forward primer. Since the translocation involves different JH genes (JH1 to JH6), probes are not consensus for all patients but are represented by a panel of probes (Table 2), depending on the JH sequence rearranged. In this manner, the same probe can be suitable for different patients sharing the same JH gene (*see Note 11*).

- For BCL2-MBR/IgH translocation (Fig. 6c) the strategy uses consensus primer-probe, previously described [6, 40], universally suitable for different FL patients, making MRD analysis in this type of lymphoma much cheaper and faster compared to ASO primers approach (*see Note 12*).

### 3.3.2 ddPCR Reaction Setup

Major concepts to be aware of, before planning a ddPCR experiment:

- No standard curve generation is needed for ddPCR experiment setup, and therefore no determination of percentage of tumor

cell, by flow cytometry, is required. Tumor burden quantification is calculated as number of copies of target in  $1 \times 10^5$  cells (corresponding approximately to 500 ng of gDNA).

2. As for any kind of PCR experiments, a positive control is mandatory (*see Note 13*). The majority of BM or PB diagnostic samples for MCL contain less than 50% of tumor cells infiltration. However, in order not to waste precious diagnostic gDNA and to avoid system overloading, in case of highly infiltrated samples we suggest loading a  $10^{-1}$  or even a  $10^{-2}$  diagnostic gDNA dilution point as positive control.
3. Since nonspecific amplification can be observed, negative controls are mandatory. Nonspecific amplification (buffy coat) should be run in six replicates and no template control (NTC) in triplicates, for each specific marker quantification.
4. As the load of less template will result in loss of sensitivity, particular attention must be used in gDNA quantification of follow-up samples. So far, no consensus has been reached on reference gene usage. Thus, in the context of the Italian MRD lymphoma network, supported by the Italian Lymphoma Foundation (FIL), an ongoing study is currently evaluating several reference genes to establish which is the more stable and reliable for gDNA quality/quantity assay. In the meantime, we suggest to test, especially for follow-up samples, a reference gene (i.e., albumin or RNaseP) on 100 ng of gDNA, in only one replicate, preferably in the same ddPCR reaction as the target gene.
  - (a) In the Pre-PCR zone, prepare the  $20\times$  primer and probe mix and the ddPCR reaction mix, as shown in Fig. 7 (*see Note 14*).
  - (b) Dispense the reaction mix in one well, based on the number of technical replicates needed, as described in Fig. 7 (i.e., for 1 replicate 16.5  $\mu\text{L}$  and for 3 replicates 49.5  $\mu\text{L}$  of mix).
  - (c) In the DNA zone, load the amount of gDNA (100 ng/ $\mu\text{L}$ ) depending on the number of planned replicates per well, as indicated in Fig. 7 (i.e., for 1 replicate = 16.5  $\mu\text{L}$  of mix, add 5.5  $\mu\text{L}$  of gDNA; for 3 replicates = 49.5  $\mu\text{L}$  ( $16.5 \times 3$ ) of mix, add 16.5  $\mu\text{L}$  ( $5.5 \times 3$ ) of gDNA). From this well, 20  $\mu\text{L}$  of mix will be taken for each replicate droplet generation. To ensure that the gDNA is thoroughly mixed: vortex, spin down, and then pipet the sample a few times before adding the gDNA to the mix (*see Note 15*).
  - (d) Seal carefully the plate or the strips with optical adhesive film or caps, mix and spin down briefly.

ddPCR reaction Mix	
Reagents	1 well Reaction ( $\mu$ l)
2X ddPCR Supermix for Probes (No dUTP)	11
20X Target Primers/Probe mix	1.1
HINFI (2U/ $\mu$ l)	1.1
H <sub>2</sub> O	3.3
TOT mix	16.5
Total amount of input gDNA (100 ng)	5.5
TOT	22
<b>20X Target Primers/Probe mix (50<math>\mu</math>l): 38 <math>\mu</math>l H<sub>2</sub>O+2 <math>\mu</math>l PROBE+5 <math>\mu</math>l EACH PRIMER</b>	
MIX for 1 replicate	16,5 $\mu$ l
gDNA (100 ng)	5,5 $\mu$ l
MIX for 3 replicates	49,5 $\mu$ l
gDNA (100 ng)	16,5 $\mu$ l
MIX for n replicates/1well	16,5 $\mu$ l x n
gDNA (100 ng)	5,5 $\mu$ l x n
Total Volume loaded for droplet generation	20 $\mu$ l
Total amount of droplets mix	40 $\mu$ l

**Fig. 7** ddPCR reaction and setup

- (e) Proceed with droplet generation, loading 20  $\mu$ L of reaction mix and 70  $\mu$ L of droplet generation oil into the proper DG8 cartridge wells, following manufacturer's instruction (see Note 16).
- (f) Transfer 40  $\mu$ L of generated droplets to the hard-shell high-profile, 96-well semiskirted PCR plates and seal immediately with a strip of scotch tape.
- (g) After all samples are loaded, remove the tape and seal with a pierceable foil on PX1 PCR plate sealer (Bio-Rad, Laboratories, Hercules, CA, USA). Run on thermocycler using the default Bio-Rad thermal cycling protocol (95 °C, 10 min; 94 °C, x 30 s, and (optimized or gradient) Tm°C x 1 min, for 40 cycles; 98 °C for 10 min) adjusting the Tm according to the ASO primers for MM and MCL and to 59 °C for BCL2-MBR/IGH (Fig. 7).
- (h) Load the post-PCR 96-wells plate in the QX100-QX200 droplet reader. Before run on the droplet reader, PCR plate can also be stored at 4 °C, no longer than 24 h.

For data interpretation, MRD tumor burden quantification is expressed as total copies of target gene generated by performing

the merge of replicates by QuantaSoft (Bio-Rad Inc.) (*see Notes 17 and 18*). Examples of data quality and analysis, for each target, are shown in Figs. 8, 9, and 10.

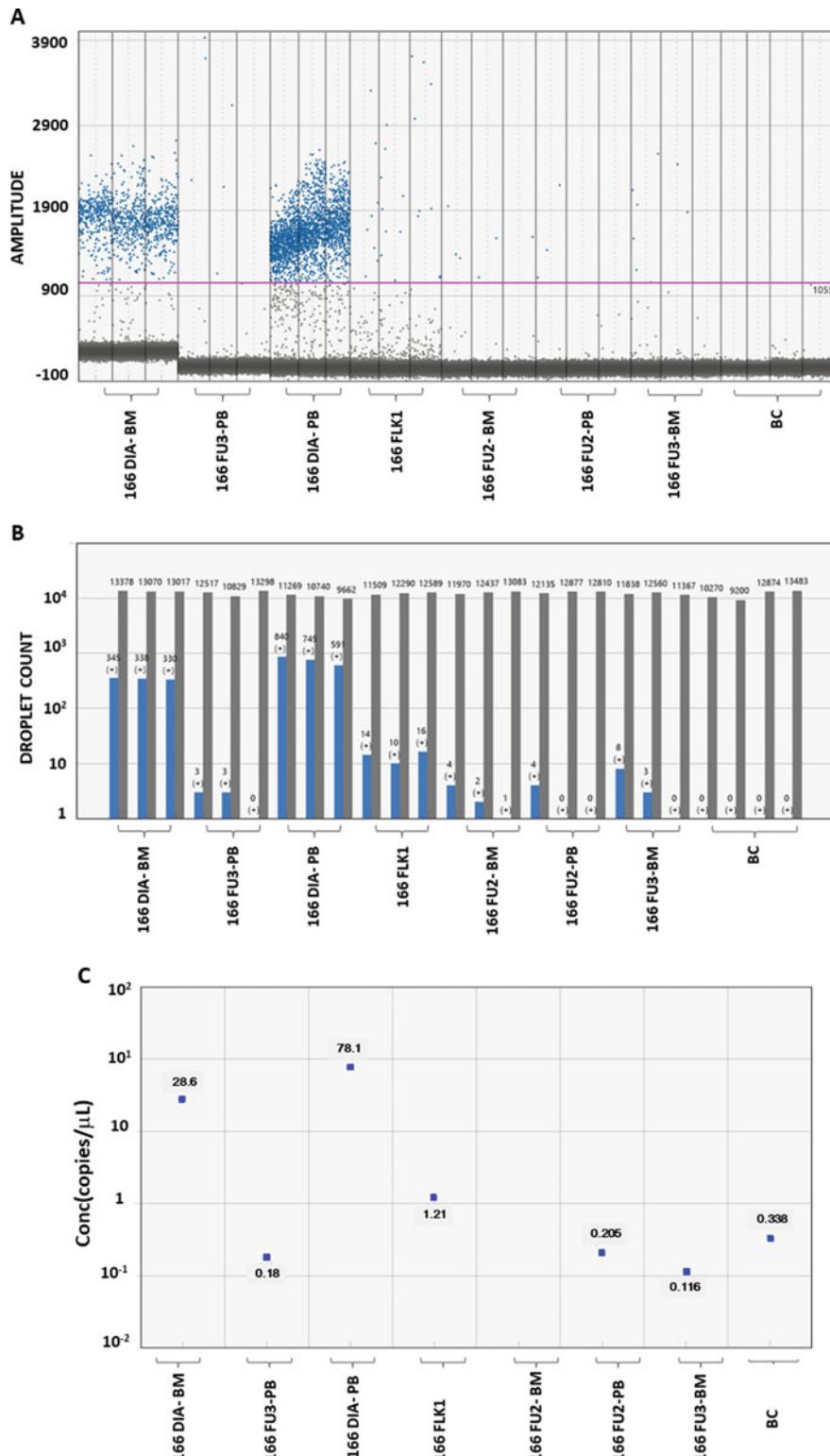
### 3.4 Final Remarks

Recent advances in gene quantification strategies have led to increased sensitivity and reduced variability and risk of contamination. ddPCR appears to be a feasible and attractive alternative method for MRD assessment in lymphoproliferative disorders and might complement or even substitute qPCR in routine clinical laboratories. One milestone of MRD monitoring by qPCR has been the multilaboratory standardization established by large cooperative efforts such as the Euro-MRD group (*see Note 19*). Currently ddPCR still has to undergo such a critical validation step. However, considering the reduced labor intensiveness of the experimental setup and the current robust structure of the interlaboratory quality control groups, we believe that ddPCR standardization would be definitely easier and faster than qPCR. However, the real advantages and predictive value of ddPCR still need to be further investigated in the context of prospective clinical trials.

---

## 4 Notes

1. Keep PCR reagents in a room where template is not isolated or stored. Use sterile filter tips for pipetting to minimize contamination from aerosols. We suggest to verify the quality of tips used for droplet loading or to use what recommended by Bio-Rad in the instruction manual. We recommend having separate pipettes and pipette tips for PCR setup, DNA loading, droplet generation, and post-PCR. Moreover, we suggest to regularly decontaminate PCR work areas (with 10% bleach) as well as pipettes (under UV).
2. Based on personal experimental observations (unpublished data), we suggest collecting samples in EDTA or sodium citrate instead of Heparin.
3. For assessment of gDNA concentration, the NanoDrop method is quite accurate. However, always check the absorbance spectrum and not only the concentration value. 260/230 and 260/280 ratios and the shape of the curve can give you important information about your sample quality. When measuring your gDNA on the NanoDrop, firstly ensure that the DNA is thoroughly mixed, preferably using vortex and pipetting several times. If amount of gDNA is below 100 ng/ $\mu$ L or 260/280 ratio is not 1.8–1.9, sample must be repurified.
4. We suggest to estimate gDNA concentration after 100 ng/ $\mu$ L stock preparation and before any PCR reaction, especially if sample is stored at 4 °C.

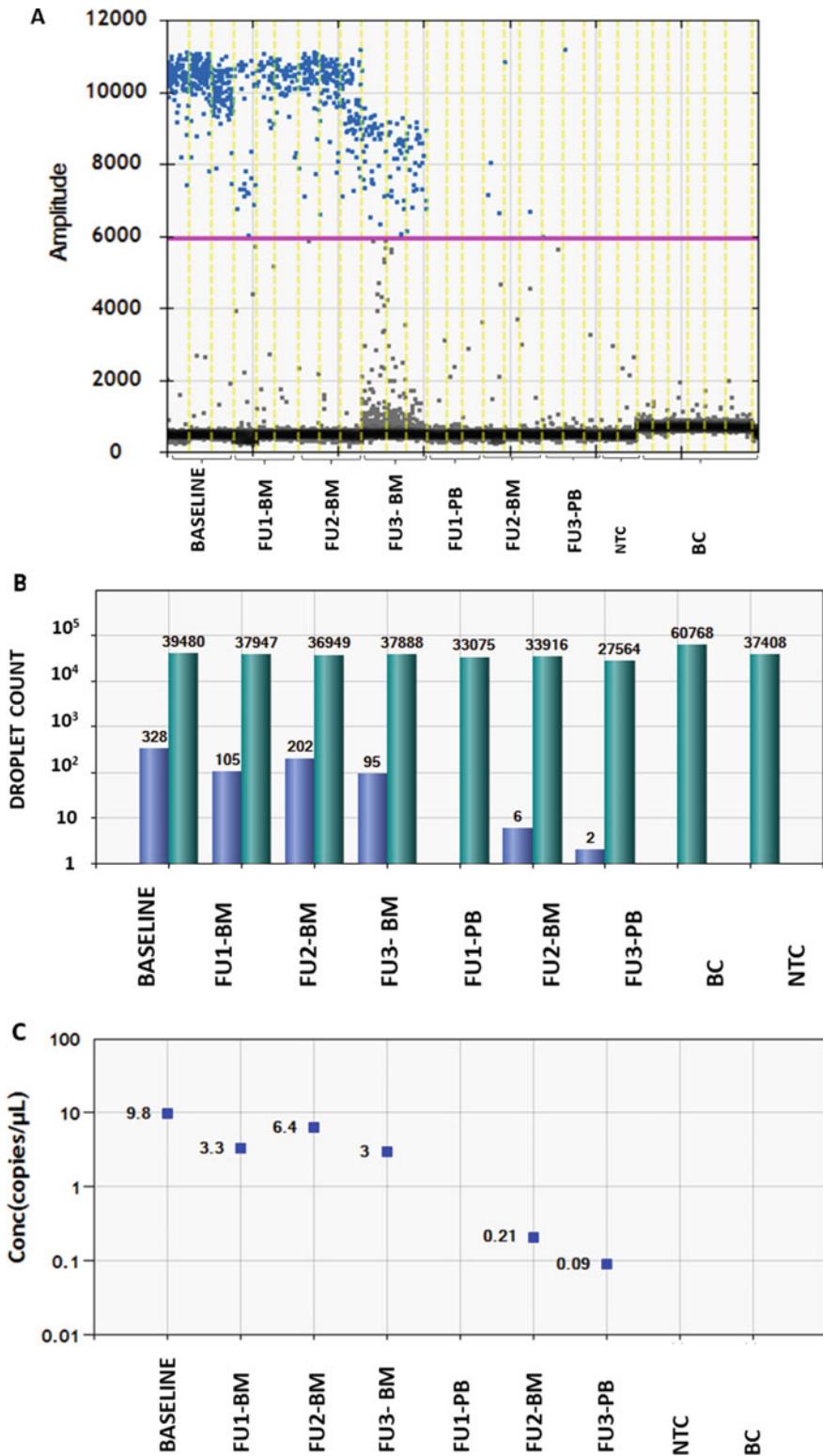


**Fig. 8** MRD monitoring of IGH rearrangement in an MCL patient: ddPCR was performed with patient-specific ASO primers and LVH5 probe. In each panel samples are reported from diagnostic to last follow-up. (a) 1D

5. Always include a DNA molecular weight marker such as φX174 DNA, cleaved with Hae III, ranging from 72 to 1353 bps, to check the correct PCR product size. We suggest to use the 5× GoTaqFlexi Green Buffer which allows loading the amplification reaction directly onto the gel, without need of loading buffer and which is compatible with “QIAquick PCR purification kit.” Run the gel at 120 V for 15–20 min, or until the two color dye markers are well separated on the gel. One percent ethidium bromide can be substituted by RealSafe Nucleic acid staining solution at the concentration suggested by the manufacturer (Durviz, S.L., Spain).
6. Each qualitative PCR reaction must include a positive control (to verify the quality of amplification) and a “no template control” (NTC) (to check for contamination). If the positive control fails or in case of PCR contamination, if the NTC is positive, review all the procedures and change all reagents.
7. In case of lack of amplification, but good performance of PCR confirmed by positive control sample, always verify that you have good sample gDNA quantity/quality by performing a PCR reaction for a housekeeping gene (e.g., Albumin, RNase P, and p53).
8. For electropherogram analysis “Chromas” is ideal for basic analysis of sequences and has no multiple sequences alignment capability. To create a sequence database for routinely checking for sequence contaminations, we suggest to gather all the sequences by a multiple sequence alignment software such as free CLC sequences viewer software (<http://www.clcbio.com>).
9. The efficiency of the ligation is determined by the comparison of the growth of the experimental versus negative control plates. If problems with cloning are encountered, verify that: (a) you used fresh PCR product for ligation, since TOPO-TA cloning requires fresh PCR; (b) competent cells are stored properly at –80 °C, since they are very sensitive to

---

**Fig. 8** (continued) plot for patient coded “166” shows: bone marrow (BM), peripheral blood (PB) at diagnosis (DIA), and post therapy follow-up (FU) at three time points (FLK1, FU2, and FU3). A single threshold has been set in between the highest amplitude value of BC and the positive control (DIA). (b) Plot for positive events (shorter bars) and total droplets (taller bars) for each sample analysed. Of note one event occurs in one of the BC replicates and it is omitted from the analysis since it shows to be a nonspecific signal (overly high amplitude value 11.000, data not shown in plot A since outside the amplitude range for graphical). (c) Representation of concentration in terms of copies/ $\mu$ L



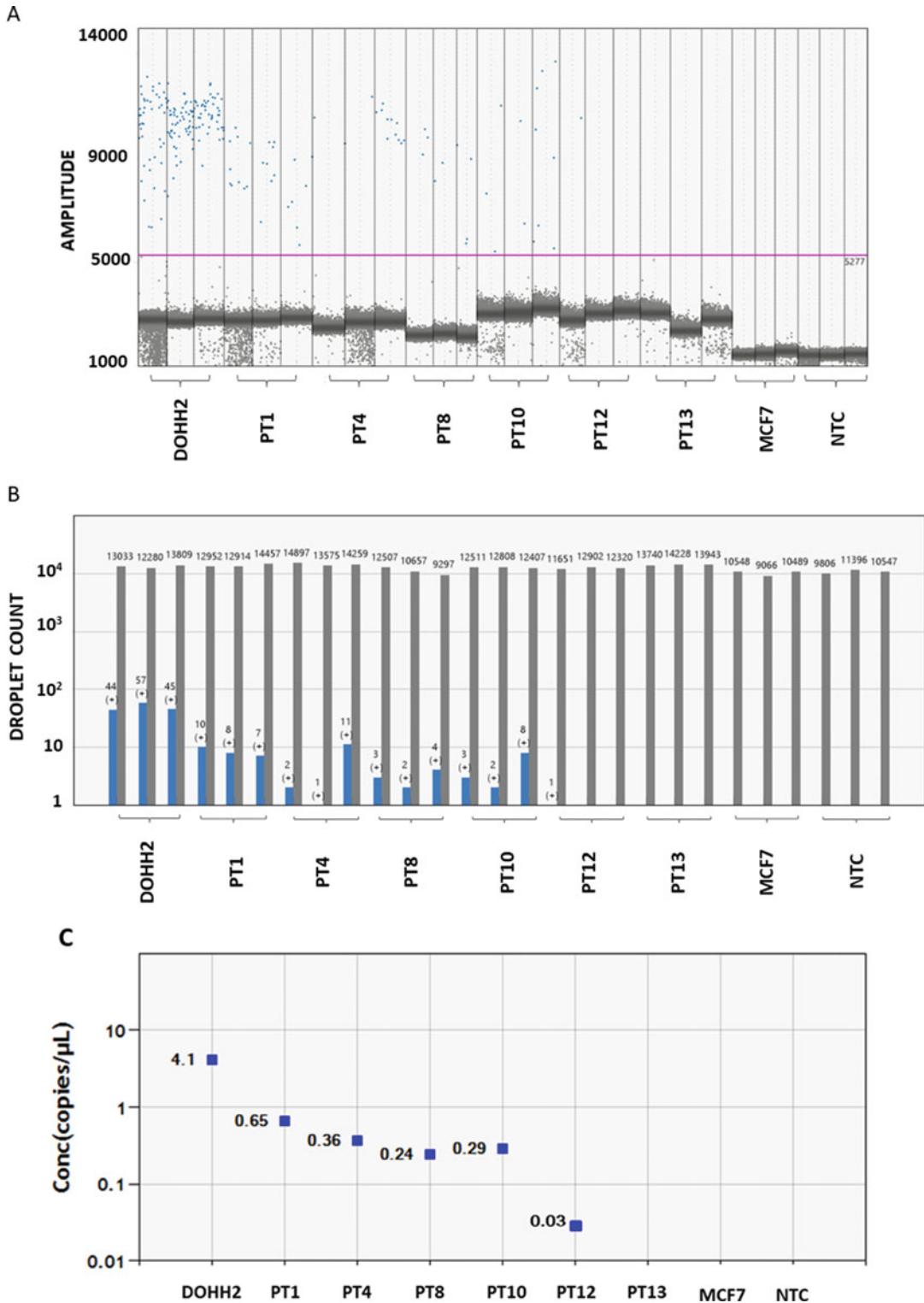
**Fig. 9** MRD monitoring of BCL1/IGH translocation in an MCL patient: ddPCR performed with a patient-specific ASO forward primer, JH1-4-5 probe and JH3 reverse primer. **(a)** 1D plot for patient samples at diagnosis

temperature; (c) thermomixer temperature for heat-shock procedure is correct; (d) SOC medium has not been stored for too long.

10. Different groups use an “ASO probe” approach, derived from the Euro-MRD guidelines for ALL, a neoplasia that is not affected by somatic hypermutation of IGH sequences. The “ASO primer” approach has shown to be a more powerful strategy in the context of B-cell lymphoproliferative disorders, characterized by a variable mutational load of IGH rearrangement [40]. Moreover, “ASO primer” strategy has the advantage to be much cheaper since no patient-specific probes are needed.
11. Within the Euro-MRD network, the Lymphoma MRD group recently planned to compare the specificity of different strategies for BCL1/IGH primers and probe design, for MRD monitoring (24th Euro-MRD meeting, Zurich 6–7 November 2015).
12. Clonal BCL2-MBR/IGH translocation is quantified by ddPCR with a consensus primer-probe system (Table 2) that does not require information about the sequence. However, we recommend to always check the BCL2-MBR/IGH sequence at diagnosis, in order to confirm the patient’s specific clone and for its further monitoring in follow-up samples, as well as to monitor for possible contaminations.
13. With respect to qPCR, ddPCR has the advantage of being an absolute quantification method. However, although a standard curve is not needed, in each experiment a positive control is required. For IGH rearrangements and BCL1/IGH based ddPCR, each patient requires his own baseline (or positive) sample as positive control, while for BCL2/IGH based ddPCR reactions gDNA from DOHH2 cell line is used as positive control. The annealing temperature for each ASO primer set can be different. For ASO primers test, use a thermal gradient ( $\pm 2$  °C from the calculated primers Tm) to optimize ddPCR results. If the positive control shows positive droplets at greater than one amplitude level (graphically represented by two clouds in 2D droplet plots or 2 bands in 1D droplet plots), the annealing temperature can be increased in order to improve the specificity; alternatively new ASO primers must be designed until a single positive cluster is achieved (and the concentration

---

**Fig. 9** (continued) (BASELINE) and at three follow-up time points. A single threshold has been set in between the highest amplitude value of BC and the positive control (b) Positive events (shorter bars) and total droplets (taller bars) for each sample analysed (c) Representation of concentration in terms of copies/ $\mu$ L



**Fig. 10** MRD monitoring of BCL2-MBR/IGH translocation in follow-up samples from six FL patients: ddPCR performed with consensus primers and probe (see Table 2). (a) 1D plot for diluted DOHH2 positive control cell

is consistent over a range of temperatures) even while the separation between negative and positive clusters may vary.

14. We observed that for all assays, final primer and probe concentrations of 500 nM and 200 nM, respectively, work properly in the ddPCR reaction (for 25  $\mu$ L of 20 $\times$  probe primer mix: 1  $\mu$ L probe (100  $\mu$ M), 2.5  $\mu$ L each primers (100  $\mu$ M), 19  $\mu$ L H<sub>2</sub>O). For ddPCR analysis, MGB or BHQ1 probes are required. TAMRA probes must be avoided since they lead to high background and noisy signals. As shown in Fig. 7, we normally prepare the ddPCR mix for a volume increased by 10% to be sure to have enough reaction mix volume for each replicate, and to not risk introducing air bubbles into the DG8 cartridges when loading the samples.
15. A total of 500 ng of gDNA is loaded in each replicate. Be aware that, depending on the DNA extraction protocol, gDNA can be more viscous and this characteristic can affect droplet formation. For this reason, in case of sticky DNA, it is essential to use restriction enzymes to decrease the size of the input DNA. We observed that 2 U/ $\mu$ L of restriction enzyme (HINFI) directly added into the reaction mix are sufficient for a good ddPCR performance and proper droplet generation. However, before using the enzyme, verify that primers and probe sequences will not be damaged. To prepare a ddPCR reaction mix containing the enzyme, add 1.1  $\mu$ L of 2 U/ $\mu$ L of enzyme to the 22  $\mu$ L of ddPCR reaction mix, adjusting properly with water.
16. We suggest: (a) to load the 20  $\mu$ L of ddPCR mix in to the DG8 cartridge using a multichannel pipet with filtered tips; (b) to paying attention when removing the DG8 gasket and always remove it from the NTC, or BC well position, to the positive control sample; (c) to load the ddPCR reaction mix into the cartridge always before the oil; (d) to transfer the 40  $\mu$ L of droplets/well to the hard-shell, high-profile, 96-well semi-skirted PCR plates using a multichannel pipet with no filter tips and immediately seal the wells with a scotch tape; (e) to

---

**Fig. 10** (continued) line (first three replicates), six follow-up samples from FL patients (PT) (three replicates each) and negative controls (MCF7 and NTC). Of note, three different gDNA samples form MCF7 cell line, (two replicates each) had been used as negative control. Variability in negative droplets amplitude depends on sample type and purity and only in one patient (PT13, wells E04-E06) it is between technical replicates. Moreover, this assay uses consensus primers and probe for different patients and different amplitude could represents differences in assay efficiency. A single threshold has been set in between the highest value of amplitude of the background and the positive control. (b) Positive events (shorter bars) and total droplets (taller bars) for each sample analysed (c) Representation of concentration in terms of copies/ $\mu$ L

load the technical replicates in different DG8 cartridges to avoid losing a full sample reaction, due to technical error that can happen during pipetting or the droplet generation procedure; (f) to remove carefully, any bubble created into the DG8 cartridge “sample” well during sample loading.

17. If you find a contamination (positive events in the range of the positive control amplitude) in your NTC sample: (a) replace all reagents and thoroughly clean PCR preparation areas; (b) check whether the probe is degraded, since free dye can give rise to signal and/or high background.
18. Only replicates with more than 9000 droplets must be considered for the analysis. We define as MRD positive by ddPCR those samples that have all positive replicates and negative those that have all negative replicates. Based on the higher variability observed at low target concentrations and applying the Euro-MRD guidelines, cases showing alternatively positive or negative replicates must be judged with caution, especially if less than totally three positive events are detected in the three technical replicates. In such a case, ddPCR of the sample should be performed again on more replicates (at least six) to see if the result appears consistent. One critical limitation of qPCR is its drawback to providing a reliable target quantification for a proportion of samples with tumor burden between the sensitivity and the quantitative range of the method. Samples falling in this window of unreliable quantification, by qPCR, are currently defined as “positive nonquantifiable” (qPNQ). ddPCR showed to be able to partially recover these qPNQ samples [34]. However, specific guidelines for data interpretation still need to be established and the potential advantages and predictive value of ddPCR results, especially for these cases, need to be further validated.
19. The Euro-MRD group, (<http://www.euromrd.org>) is composed by international MRD-PCR laboratories, under the umbrella of ESLHO consortium (European Scientific foundation for Laboratory Hemato Oncology). Euro-MRD group is responsible for developing of new strategies and techniques, as well as for developing guidelines for the interpretation of qPCR based MRD data and in the near future also for developing guidelines for ddPCR based MRD.

---

## Acknowledgments

We gratefully acknowledge Elisa Genuardi, Barbara Mantoan, Martina Ferrante, Luigia Monitillo, Manuela Gambella, Daniela Barbero, Irene Della Starza, Elena Ciabatti, Nadia Dani, and Marta Varotto for their excellent technical support. Moreover, we are grateful to the Italian Lymphoma Foundation (FIL) that is supporting our ongoing research projects on MCL and FL.

## References

1. Andersen NS, Pedersen LB, Laurell A et al (2009) Pre-emptive treatment with rituximab of molecular relapse after autologous stem cell transplantation in mantle cell lymphoma. *J Clin Oncol* 27:4365–4370
2. Ladetto M, Pagliano G, Ferrero S et al (2010) Major tumor shrinking and persistent molecular remissions after consolidation with bortezomib, thalidomide, and dexamethasone in patients with autografted myeloma. *J Clin Oncol* 28:2077–2084
3. Pott C, Hoster E, Delfau-Larue MH et al (2010) Molecular remission is an independent predictor of clinical outcome in patients with mantle cell lymphoma after combined immunochemotherapy: a European MCL intergroup study. *Blood* 115:3215–3223
4. Pott C (2011) Minimal residual disease detection in mantle cell lymphoma: technical aspects and clinical relevance. *Semin Hematol* 48:172–184
5. Korthals M, Sehnke N, Kronenwett R et al (2012) The level of minimal residual disease in the bone marrow of patients with multiple myeloma before high-dose therapy and autologous blood stem cell transplantation is an independent predictive parameter. *Biol Blood Marrow Transplant* 18:423–431
6. Ladetto M, Lobetti-Bodoni C, Mantoan B et al (2013) Persistence of minimal residual disease in bone marrow predicts outcome in follicular lymphomas treated with a rituximab-intensive program. *Blood* 122:3759–3766
7. Ferrero S, Ladetto M, Drandi D et al (2015) Long-term results of the GIMEMA VEL-03-096 trial in MM patients receiving VTD consolidation after ASCT: MRD kinetics' impact on survival. *Leukemia* 29(3):689–695
8. Ladetto M, Ferrero S, Drandi D et al (2016) Prospective molecular monitoring of minimal residual disease after non-myeloablative allografting in newly diagnosed multiple myeloma. *Leukemia* 30(5):1211–1214
9. Gribben JG, Freedman A, Woo SD et al (1991) All advanced stage non-Hodgkin's lymphomas with a polymerase chain reaction amplifiable breakpoint of bcl-2 have residual cells containing the bcl-2 rearrangement at evaluation and after treatment. *Blood* 78(12):3275–3280
10. van Dongen JJ, Langerak AW, Brüggemann M et al (2003) Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 concerted action BMH4-CT98-3936. *Leukemia* 17:2257–2317
11. Rimokh R, Berger F, Delsol G et al (1994) Detection of the chromosomal translocation t(11;14) by polymerase chain reaction in mantle cell lymphomas. *Blood* 83(7):1871–1875
12. Ladetto M, Donovan JW, Harig S et al (2000) Real-time polymerase chain reaction of immunoglobulin rearrangements for quantitative evaluation of minimal residual disease in multiple myeloma. *Biol Blood Marrow Transplant* 6:241–253
13. Brüggemann M, Droese J, Bolz I et al (2000) Improved assessment of minimal residual disease in B cell malignancies using fluorogenic consensus probes for real-time quantitative PCR. *Leukemia* 14:1419–1425
14. van der Velden VH, Cazzaniga G, Schrauder A et al (2007) Analysis of minimal residual disease by Ig/TCR gene rearrangements: guidelines for interpretation of real-time quantitative PCR data. *Leukemia* 21:604–611
15. van der Velden VH, Hochhaus A, Cazzaniga G et al (2003) Detection of minimal residual disease in hematologic malignancies by real-time quantitative PCR: principles, approaches, and laboratory aspects. *Leukemia* 17(6):1013–1034
16. Pott C, Brüggemann M, Ritgen M et al (2013) MRD detection in B-cell non-Hodgkin lymphomas using Ig gene rearrangements and chromosomal translocations as targets for real-time quantitative PCR. *Methods Mol Biol* 971:175–200
17. Ladetto M, Brüggemann M, Monitillo L et al (2014) Next-generation sequencing and real-time quantitative PCR for minimal residual disease detection in B-cell disorders. *Leukemia* 28:1299–1307
18. Martinez-Lopez J, Lahuerta JJ, Pepin F et al (2014) Prognostic value of deep sequencing method for minimal residual disease detection in multiple myeloma. *Blood* 123:3073–3079
19. van Dongen JJ, van der Velden VH, Brüggemann M et al (2015) Minimal residual disease diagnostics in acute lymphoblastic leukemia: need for sensitive, fast, and standardized technologies. *Blood* 125(26):3996–4009
20. Böttcher S, Ritgen M, Fischer K et al (2012) Minimal residual disease quantification is an independent predictor of progression-free and overall survival in chronic lymphocytic leukemia: a multivariate analysis from the randomized G CLLSG CLL8 trial. *J Clin Oncol* 30:980–988

21. Rawstron AC, Child JA, de Tute RM et al (2013) Minimal residual disease assessed by multiparameter flow cytometry in multiple myeloma: impact on outcome in the Medical Research Council Myeloma IX Study. *J Clin Oncol* 31:2540–2547
22. Paiva B, Gutiérrez NC, Rosiñol L et al (2012) High-risk cytogenetics and persistent minimal residual disease by multiparameter flow cytometry predict unsustained complete response after autologous stem cell transplantation in multiple myeloma. *Blood* 119:687–691
23. Markey AL, Mohr S, Day PJ (2010) High-throughput droplet PCR. *Methods* 50:277–281
24. Hindson BJ, Ness KD, Masquelier DA et al (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem* 83:8604–8610
25. Pinheiro LB, Coleman VA, Hindson CM et al (2012) Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Anal Chem* 84:1003–1011
26. Day E, Dear PH, McCaughan F (2013) Digital PCR strategies in the development and analysis of molecular biomarkers for personalized medicine. *Methods* 59(1):101–107
27. Didelot A, Kotsopoulos SK, Lupo A et al (2013) Multiplex picoliter-droplet digital PCR for quantitative assessment of DNA integrity in clinical samples. *Clin Chem* 59:815–823
28. Dingle TC, Sedlak RH, Cook L et al (2013) Tolerance of droplet-digital PCR vs real-time quantitative PCR to inhibitory substances. *Clin Chem* 59:1670–1672
29. Hayden RT, Gu Z, Ingwersen J, Abdul-Ali D (2013) Comparison of droplet digital PCR to real-time PCR for quantitative detection of cytomegalovirus. *J Clin Microbiol* 51:540–546
30. Hindson CM, Chevillet JR, Briggs HA et al (2013) Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat Methods* 10:1003–1005
31. Huggett JF, Whale A (2013) Digital PCR as a novel technology and its potential implications for molecular diagnostics. *Clin Chem* 59:1691–1693
32. Nixon G, Garson JA, Grant P et al (2014) Comparative study of sensitivity, linearity, and resistance to inhibition of digital and nondigital polymerase chain reaction and loop mediated isothermal amplification assays for quantification of human cytomegalovirus. *Anal Chem* 86:4387–4394
33. Huggett JF, Cowen S, Foy CA (2015) Considerations for digital PCR as an accurate molecular diagnostic tool. *Clin Chem* 61 (1):79–88
34. Drandi D, Kubiczkova-Besse L, Ferrero S et al (2015) Minimal residual disease detection by droplet digital PCR in multiple myeloma, mantle cell lymphoma, and follicular lymphoma. A comparison with real-time PCR. *J Mol Diagn* 17(6):652–660
35. Deane M, McCarthy KP, Wiedemann LM et al (1991) An improved method for detection of B-lymphoid clonality by polymerase chain reaction. *Leukemia* 5:726
36. Voena C, Ladetto M, Astolfi M et al (1997) A novel nested-PCR strategy for the detection of rearranged immunoglobulin heavy-chain genes in B cell tumors. *Leukemia* 11:1793
37. Summers KE, Goff LK, Wilson AG et al (2001) Frequency of the Bcl-2/IgH rearrangement in normal individuals: implications for the monitoring of disease in patients with follicular lymphoma. *J Clin Oncol* 19:420–424
38. Albingher-Hegyi A, Hochreutener B, Abdou MT et al (2002) High frequency of t(14;18) translocation breakpoints outside of major breakpoint and minor cluster regions in follicular lymphomas improved polymerase chain reaction protocols for their detection. *Am J Pathol* 160:823–832
39. Ladetto M, De Marco F, Benedetti F et al (2008) Prospective, multicenter randomized GITMO/IIL trial comparing intensive (R-HDS) versus conventional (CHOP-R) chemoimmunotherapy in high-risk follicular lymphoma at diagnosis: the superior disease control of R-HDS does not translate into an overall survival advantage. *Blood* 111 (8):4004–4013
40. Ladetto M, Sametti S, Donovan JW et al (2001) A validated real-time quantitative PCR approach shows a correlation between tumor burden and successful ex vivo purging in follicular lymphoma patients. *Exp Hematol* 29 (2):183–193
41. Della Starza I, Cavalli M, Del Giudice I et al (2014) Comparison of two real-time quantitative polymerase chain reaction strategies for minimal residual disease evaluation in lymphoproliferative disorders: correlation between immunoglobulin gene mutation load and real-time quantitative polymerase chain reaction performance. *Hematol Oncol* 32(3):133–138



# Chapter 15

## Quantitation of JAK2 V617F Allele Burden by Using the QuantStudio™ 3D Digital PCR System

Elena Kinz and Axel Muendlein

### Abstract

The JAK2 V617F mutation is highly prevalent in patients with myeloproliferative neoplasms (MPN). Furthermore, it has been shown that its allelic burden correlates with hematologic characteristics, drug response, and clinical endpoints in MPN patients. Digital PCR is an emerging technology for sensitive mutation detection and quantitation, based on dilution and high-grade partitioning of a sample. Here, we describe the use of the nanofluidic chip-based QuantStudio™ 3D Digital PCR System for quantitation of the JAK2 V617F mutation.

**Key words** Digital PCR, Nanofluidic chip, JAK2 V617F, Allele burden, Myeloproliferative neoplasms

---

### 1 Introduction

The detection of the acquired mutation JAK2 V617F in patients with Philadelphia-negative myeloproliferative neoplasms (MPNs), including polycythemia vera (PV), essential thrombocythemia (ET), and primary myelofibrosis (PMF) in 2005 [1–4], has revolutionized MPN diagnostics. The mutation occurs in more than 95% of PV cases and about 50–60% of those with ET or PMF. Consequently, presence of the JAK2 V617F mutation has been included as a major diagnostic criterion for these diseases in the World Health Organization (WHO) classification [5].

Subsequent studies indicated that the different MPN phenotypes had different mutational loads [6, 7]: Patients with ET have the lowest allele burden (<50%), those with PV and PMF intermediate levels, and those with post-PV myelofibrosis the highest burden [8]. Therefore, at MPN diagnosis, measurement of the allele burden may help to discriminate between different MPN phenotypes: when an allele burden is greater than 50%, the likelihood of a masked PV or myelofibrotic evolution exists [9]. JAK2 allele burden may have prognostic significance as well, since it correlates with clinical endpoints in MPN patients [10]: A high allele burden in

patients with ET or PV has been associated with increased risk of thrombotic events [11] and transformation to myelofibrosis [12, 13]. On the other hand, in PMF patients, a low allele burden ( $\leq 25\%$ ) has been correlated with poorer survival, probably due to increased leukemic transformation [14] or systemic infections [15]. Furthermore, JAK2 V617F quantitation has been used to assess efficacy of treatment with significant reductions in allele burden reported in PV patients following pegylated interferon  $\alpha$ -2a treatment [16, 17]. Therefore, quantitation of JAK2 V617F allele burden may be used in stratification of patients with MPN regarding prognosis as well as drug treatment.

Several techniques have been described to quantify JAK2 V617F allele burden; out of those, until recently, allele specific real-time quantitative polymerase chain reaction (AS-qPCR) has been proposed to be the most reliable and sensitive one [18, 19].

Recently, we showed a high correlation between a AS-qPCR kit for quantitation of JAK2 allele burden, registered for in vitro diagnostic use, and chip-based digital PCR (dPCR) using the QuantStudio™ 3D Digital PCR System (Thermo Fisher Scientific, Waltham, Massachusetts, USA) providing sensitivity to detect less than 0.1% JAK2 V617F allelic burden [20]. Accuracy and reproducibility of results obtained by AS-qPCR strongly depend on used standards needed for quantitation of target molecules. In contrast, chip-based dPCR allows assessment of the total copy number of target molecules without the need for standard curves facilitating standardization of quantitation. Particularly, chip-based digital PCR demonstrates an appropriate technique for the detection of mutations present at low allele frequencies because partitioning of sample reduces wild-type background signals and therefore enhances accuracy of detection. Therefore, chip-based dPCR appears very suitable for measurement of JAK2 V617F allelic burden.

The QuantStudio™ 3D Digital PCR System platform is composed of a GeneAmp 9700 thermal cycler (including a chip adapter kit), an automatic chip loader, and the QuantStudio™ 3D Instrument. The QuantStudio™ 3D Digital PCR 20K Chip consists of individual 20,000 partitions and each partition is a chamber for an individual PCR reaction. Genomic DNA (gDNA) samples are diluted down to a limiting quantity, such that most individual PCR reactions contain either zero or one DNA molecule. Digital PCR is an endpoint analysis based on a 5'-exonuclease assay using fluorogenic TaqMan® probes targeting the mutation. The absolute quantification of target is calculated by Poisson statistics, based on the number of negative partitions, which contain no DNA [21]. After mixing diluted gDNA sample with the premixed primer and probe assay and a PCR master mix, the PCR reaction mix is loaded onto the chip, which in turn is placed in the PCR machine to run the reaction. Next, the chip is placed in the QuantStudio™ 3D

Instrument to read out fluorogenic signals. Data are analyzed using QuantStudio™ 3D AnalysisSuite™ Software.

According to the specifications of the manufacturer as well as to our observations, QuantStudio™ 3D Digital PCR System platform is able to quantitate JAK2 V617F mutants at a prevalence as low as 0.1%, reaching sensitivity recommended for residual disease monitoring [9].

Importantly, sensitivity can be further increased using additional chips, which data are aggregated by the software and analyzed as one chip. Therefore, chip-based digital PCR represents an appropriate technique for detection and quantitation of the JAK2 V617F allelic burden.

## 2 Materials

### 2.1 Sample Collection and DNA Preparation

1. Peripheral whole blood, purified blood granulocytes and bone marrow aspirates, respectively, is suitable as a source for genomic DNA used in JAK2 V617F quantitation (*see Note 1*).
2. Kit for isolation of genomic DNA providing high DNA quality e.g.: QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany), peqGOLD Blood DNA Mini kit (VWR, Vienna, Austria), or others (*see Note 2*).
3. Fluorometer (e.g. Qubit™, Thermo Fisher) or spectrophotometer (e.g. Nano-Drop™ Lite, VWR) to determine nucleic acid concentration (*see Note 3*).

### 2.2 PCR Mix

1. Genomic DNA (see above). Store at -20 °C.
2. QuantStudio™ 3D Digital PCR Mastermix (Thermo Fisher) Store at -20 °C
3. Twenty-fold premixed Primers and TaqMan®-probe mix obtained by the TaqMan® Assays-by-Design service from Thermo Fisher. Primer and probe sequences are as follows: JAK2 V617F forward primer: 5'-AAGCTTTCTCACAG-CATTGGTTT; JAK2 V617F reverse primer: 5'--AGAAAGGCATTAGAAAGCCTGTAGTT; JAK2 V617-probe (wild-type allele): VIC®-TCTCCACAGACACATAC-MGB; and JAK2 F617-probe (mutant allele) 6-FAM™-T CCACAGAACATAC-MGB. Store at -20 °C (*see Note 4*).
4. Nuclease-free water.

### 2.3 Chip-Loading, Automated

1. QuantStudio™ 3D Digital PCR 20K Chip (v2) (Thermo Fisher) (*see Note 5*).
2. QuantStudio™ 3D Digital PCR Chip Case Lid.
3. QuantStudio™ 3D Digital PCR Chip Loader (Thermo Fisher).

4. QuantStudio™ 3D Digital Sample Loading Blades (Thermo Fisher).
5. Immersion Fluid (Thermo Fisher).

#### **2.4 dPCR Run**

1. QuantStudio™ 3D Digital PCR Chip Carriers (Thermo Fisher).
2. QuantStudio™ 3D Digital PCR Thermal Pads (Thermo Fisher).
3. Dual Flat Block GeneAmp® PCR System 9700 (Thermo Fisher).
4. QuantStudio™ 3D Tilt Base.

#### **2.5 dPCR Signal**

##### **Read Out and Data Analysis**

1. QuantStudio™ 3D Instrument (Thermo Fisher).

2. QuantStudio™ 3D AnalysisSuite™ Software (Thermo Fisher).

#### **2.6 General Consumables and Other Equipment**

1. Isopropanol.
2. Pipette and pipette tips.
3. Additional syringe tips (10 and 20 µL pipette tips without filter).
4. Powder free nitrile gloves.
5. Low-lint wipes.
6. Microcentrifuge.
7. Vortexer.

### **3 Methods**

#### **3.1 gDNA Preparation**

Prepare gDNA from peripheral blood, purified granulocytes or bone marrow aspirates (*see Note 1*) using a commercial available DNA preparation kit (*see Note 2*) following instructions of the manufacturer. Accurately assess DNA concentration using a fluorometer or a spectrophotometer (*see Note 3*).

#### **3.2 Copy Number Calculation**

The procedure of determining the number of target copies per genome (genomic equivalents) is important for appropriate chip load and quality. The DNA molecule of interest must be present in a defined quantity, so that individual PCR reactions contain zero or one target molecule. To calculate DNA concentration, it necessary to estimate target copy number of genome:

1. Convert DNA concentration from ng/µL into copies/µL, assuming 3.5 pg/copy for the JAK2 gene, or other gene, present at two copies per diploid genome (*see Note 6*).

2. According to the recommendations of the manufacturer, the concentration of the target sequence in the dPCR reaction may range between 200 and 2000 copies/ $\mu$ L (corresponding to 3000–30,000 haploid genomic equivalents in the final reaction on the 20K v2 chip; *see Note 7*). In case of quantitation of JAK2 V617F allelic burden, a DNA concentration of 1000–2000 copies/ $\mu$ L in the PCR reaction shows good results (*see Note 8*).
3. Positive and negative template as well as wild type controls (if available) should be integrated into dPCR experiments. One may use plasmid constructs, which have to be linearized before use, to estimate sensitivity of dPCR application. Calculation of copy number must be adjusted to respective plasmid size.

### 3.3 dPCR Setup

Bring QuantStudio™ 3D Digital PCR Mastermix and 20-fold premixed JAK2 TaqMan® assay to room temperature and prepare nuclease free water and DNA samples. Set up dPCR reaction mix for two chips per sample as shown in Table 1.

Depending on your estimated sample concentration, add gDNA ( $y$ ) and water ( $x$ ) to obtain a final concentration between 1000 and 2000 copies/ $\mu$ L. Place on ice until use and warm it to room temperature

### 3.4 Chip Loading

Proper loading of QuantStudio™ 3D Digital PCR 20K Chip (v2) with PCR reaction mix is essential for an accurate quantitation of target gene copy number. We strongly recommend using the QuantStudio™ 3D Digital PCR Chip Loader for loading procedure since we experienced that differences in handling can influence the outcome of analysis, especially the precision of target gene estimation in technical duplicates. The procedure of automated

**Table 1**  
dPCR reaction mix for two chips per sample

Material	Volume [ $\mu$ L] Two chips
Master Mix (v2) 2×	17.4
TaqMan® Assay 20×	1.74
Nuclease-free water <sup>a</sup>	$x$
gDNA <sup>a</sup>	$y$
Total volume	34.8 <sup>b</sup>

<sup>a</sup>For a desired template concentration of 1000 copies/ $\mu$ L, use 12.2  $\mu$ L gDNA with a concentration of 10 ng/ $\mu$ L and 4.3  $\mu$ L water

<sup>b</sup>Total volume is calculated upon an application volume of 14.5  $\mu$ L per chip. The application volume might be increased if the chip quality reviewed by software is not sufficient

chip loading is given in detail in the manual of the QuantStudio™ 3D Digital PCR system (*see Note 9*). In brief, automated chip loading procedure is as follows:

1. Plug in the chip loader. Prepare immersion fluid syringe by gently pulling back the plunger to release air and then unscrew the cap and replace it by a tip by pushing it into place carefully (*see Note 10*).
2. Load the chip in the chip nest face according to the picture next to the nest. Apply the loading blade by pressing the loading blade lever with the one hand and pressing the loading blade into the loader head with the other (*see Note 11*).
3. Peel away the red protective film from the back of the chip lid (*see Note 12*) and place it with the sticky side up in the lid nest while pressing the lid nest button (*see Note 13*).
4. Transfer 14.5 µL of the dPCR reaction mixture into the sample loading port of the sample blade (*see Note 14*).
5. Press the black loading button on the top of the chip loader to start chip loading.
6. After loading, slowly apply immersion fluid on the chip surface without contact. Fifteen drops should be sufficient to cover the whole surface (*see Note 15*).
7. To apply the chip lid on the chip, rotate the arm loader and press it down for 15 s to confirm tight sealing. Each flash of the status light illustrates 1 s intervals (*see Note 16*).
8. Press the lid nest button to release the lid and put back the arm in the original position. Take care that the chip stays in the chip nest and does not remain in the lid nest.
9. To fill the chip with immersion fluid through the chip port, hold the chip in a 45° angle with the port in the top right corner. A small air bubble (<2–3 mm in diameter), slightly larger than the fill port, should remain. Rotate the chip to see whether hidden bubbles have remained (*see Note 17*).
10. For sealing, pull back the top half of the label, remove the label backing and press the label firmly over the fill port for 5 s to confirm tight sealing (*see Note 18*). Place chips in a dark and clean location until cycling. Cycling should be started within 2 h after preparation.

### **3.5 dPCR Thermal Cycling**

1. Prepare chip adapter and thermal pads and make sure that they are clean. Otherwise clean them with isopropanol sprayed on a low-lint wipe.
2. Confirm that the tilt base is installed beneath thermal cycler.

3. Install chip adapters on both samples blocks, no matter how many samples you are going to cycle, it is necessary to have installed both adapters.
4. Check that you use the correct firmware version for the appropriate thermal cycler. (see user guide)
5. Program PCR method before starting. Following cycling condition are used for JAK2 V617F dPCR thermocycling on the Gene Amp 9700 PCR System: 96 °C for 10 min, 39 cycles at 56 °C for 2 min and at 98 °C for 30 s, followed by a final extension step at 60 °C for 2 min (*see Note 19*).
6. Open the heated cover of the thermal cycler and place the chips onto the sample blocks so that the fill ports are positioned toward the front of the cycler. Confirm that chips are clean and free from contaminants (*see Note 20*).
7. Lay clean QuantStudio™ 3D Digital PCR thermal pads on the chip adapter accurately. Do not use other thermal pads.
8. Close and engage heated cover of the thermal cycler.
9. Select preprogrammed run and start cycler.
10. Unload cycler after cycling programme is finished (*see Note 21*).
11. Inspect it for leaks, let it adapt to room temperature and clean it with isopropanol before imaging. If you do not image it immediately, place it in a clean and dark location (*see Note 22*).

### **3.6 Imaging and Primary Analysis on QuantStudio™ 3D Instrument**

The QuantStudio™ 3D Instrument processes imaging of chips individually and creates a single experiment file (.eds) for every chip analyzed, which is named after the ID number on the chip, detected by the system.

1. Switch on the instrument and wait until the instrument requests chip insertion (“Insert chip to continue”)
2. You may need to specify data destination (QuantStudio™ 3D software cloud, personal network or USB stick) and well volume prior to analysis according to user guide.
3. Open the chip tray and insert the chip face up into the tray. The Chip Barcode need to be orientated toward the front of the instrument.
4. Close the tray, the detection will start automatically. You can monitor the progress on the screen. If an error occurs, inspect the chip for problems and refer to the troubleshooting guide if no obvious problems are visible.
5. You may insert another chip after imaging is done and the instrument displays “Analyzing data” (*see Note 23*).

6. If you want to reread chips afterward, make sure that you store them in a dark place until you have analyzed the data with the software (*see Note 24*).
7. The results are assessed in copies/ $\mu\text{L}$  of nucleic acid sequences targeted by VIC<sup>®</sup> or FAM<sup>™</sup> dye. Flags of different colors illustrate a first quality assessment in primary analysis. These flags allow direct estimation of chip quality after primary analysis as green colors display that data meets quality thresholds, yellow that analysis was not ideal and should be reviewed, and red that analysis failed.
8. Further information about quality and results of detection can be estimated in secondary analysis with cloud based QuantStudio<sup>™</sup> 3D analysis software.

---

## 4 Analysis

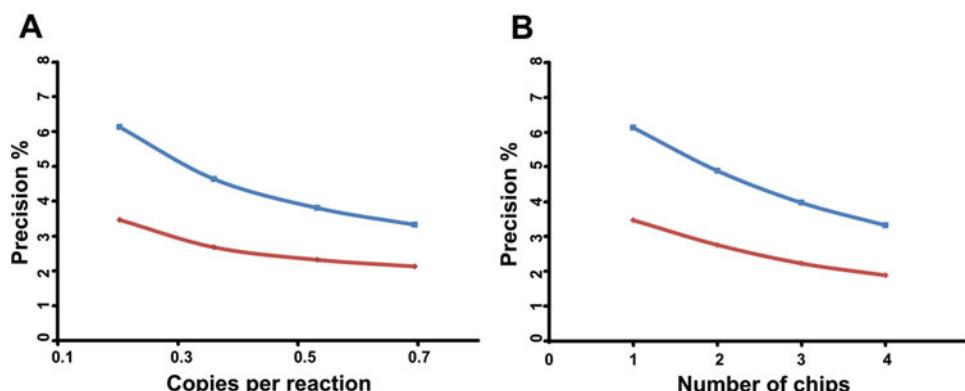
### 4.1 Chip Analysis Using QuantStudio<sup>™</sup> 3D Software

QuantStudio<sup>™</sup> 3D Digital PCR AnalysisSuite<sup>™</sup> Software is a cloud based software suite for analysis and administration of dPCR experiment data. It calculates absolute and relative number of target nucleic acid of sample and performs statistics.

1. Make sure that your computer has network connection.
2. Open a browser window and go to <https://apps.thermofisher.com/quantstudio3d/> (*see Note 25*). Sign into the software or create a new account (*see Note 26*).
3. You can create and edit projects in the Projects Home screen. Create a project and fill in a distinct project name. Import chip data (.eds files) from your computer or directly from the instrument (*see user guide*). Experiment data from up to 100 chips can be included in one project.
4. Define chips by entering sample name, assays used and fill in dilution factor of chip (*see Note 27*).

Quality of the chip can be reviewed in “Review Data” tab where all chips are displayed in the left corner of the screen. Chip of interest can be selected by double-click. Choose “colour by quality” (at Action Box) and threshold—a measure of individual well quality will be displayed from low (red) to high (green) quality on a continuous scale from 0 to 1. The default quality threshold is 0.5. The quality flag allows you to directly determine which chip to review. If a user has modified the analysis results originally generated by the instrument software, the flag appears faded (*see Note 28*). If you observe a yellow flag although the chip loading quality appears proper, separation of digital dye signals might be insufficient. In that case manual editing of dye signals appears necessary (*see step 6*).

5. To verify the uniformity of calls across the chip, review color data by call by choosing “colour by call” (at Action Box) after the quality threshold has been set. In this option each data point is displayed based on FAM™ and VIC® or ROX™ target dye signal detected.
6. In the scatter plot next to the chip view window, reporter dye signals are shown: FAM™ (blue) signals are plotted on  $\gamma$ -axis against VIC® (red) signals at the  $X$ -axis. Green signals illustrate both FAM™ and VIC® reporter dye signals while yellow ROX™ dye signal, used as internal standard, illustrates “no amplification” signals where the fluorescence intensity did not reach threshold. In the scatter plot you can now edit single data points and exclude those which are not reliable using the lasso tool (*see Note 29*).
7. Using multiple chips for one sample is required to improve precision of rare allele detection (as shown in Fig. 1). It is possible to omit chips, which did not reach standards with the “Omit Chip” option. Chosen chips are then excluded from result calculation.
8. Final results can be reviewed in the “See results” tab. Aggregate results of multiple chips and/or different dilutions of each sample (in case of identical name) are displayed in a bar graph. Ratio of mutant to wild-type target [%] or absolute quantification [copies/ $\mu$ L] is displayed at the  $\gamma$ -axis and data group (sample assay) on the  $X$ -axis.
9. The final results will give you an overview of relative and absolute results for total target and for each dye as well as confidence levels. Precision will give you information about



**Fig. 1** Impact of employed sample concentration and number of chips per sample, respectively, on precision. Blue line shows precision of FAM (mutant allele) and red line shows precision of VIC (wild type allele). **(a)** Impact of employed sample concentration on precision: Precision improves using higher template concentration, displayed as copies per reaction. **(b)** Impact of number of chips on precision: Precision improves with increasing number of chips per sample

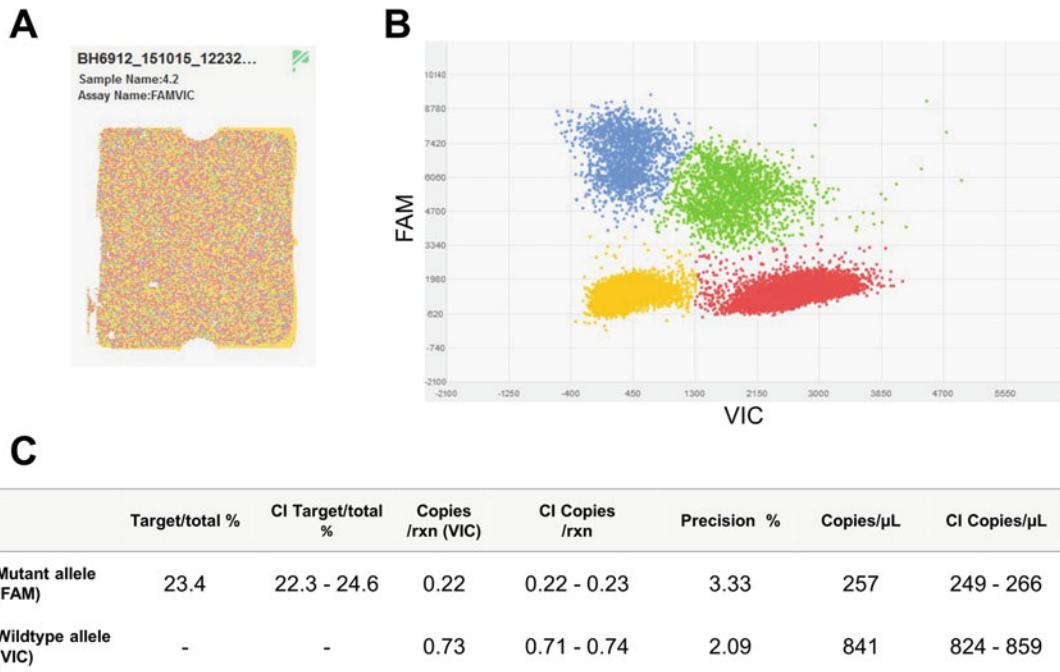
the confidence of the analysis defined as size of confidence interval for distinguishing between two sample concentrations at the same confidence level. The default confidence level is 95% and the default precision is 10%. Both can be edited by the user (*see Note 30*).

10. It is recommended to review single chip results in “replicates” tab to identify outliers and “digital calls” to identify insufficient chip load (*see Note 31*).
1. Figure 2 shows typical results for quantitation of JAK2 allele burden using QuantStudio™ 3D Digital PCR 20K Chip and QuantStudio™ 3D Digital PCR Analysis Suite™ Software. JAK2 V617F allele burden is automatically determined by the software given as target/total (%).
2. Figure 3 shows results for quantitation of JAK2 allele burden obtained from chips with different template concentrations. Both chips are of adequate quality and show comparable quantitation results for target DNA, despite differences in initial template concentration. Of note, precision is better in the chip using higher template concentrations compared to the chip using lower template concentrations in PCR.
3. Figure 4 shows chip images and results obtained from chips of different quality. Although chips exhibit different quality and a reduction of precision from high quality to low quality, estimation of target DNA delivers comparable results.

---

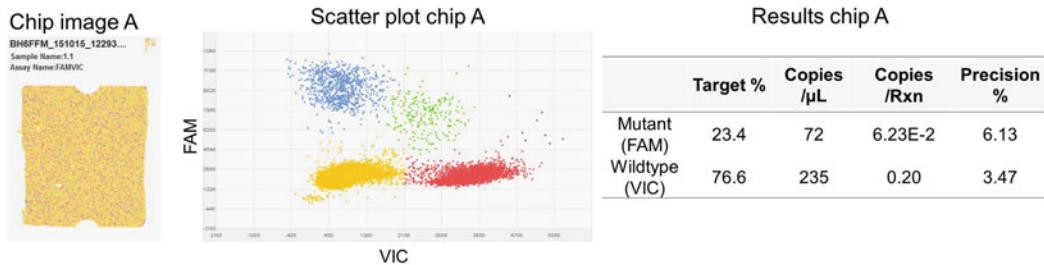
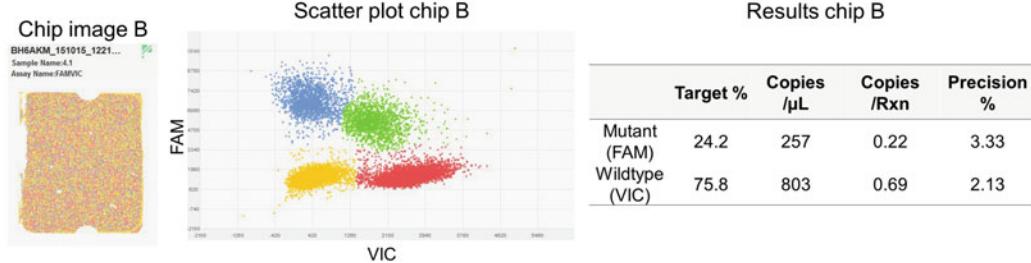
## 5 Notes

1. Detection of the JAK2 V617F mutation and quantitation of its allele burden can be performed using peripheral whole blood, purified granulocytes, or bone marrow samples. It should be noted, however, that JAK2 V617F allele burden measured in peripheral whole blood and bone marrow samples is approximately equivalent because granulocytes constitute the predominant population in both specimens [22, 23]. Furthermore, it has been shown that granulocytic isolation achieves, on average, only 15% higher JAK2 allele burden compared with that of whole white blood cells [24]. Therefore, total white blood cells from the peripheral blood are the preferred type of cell population for these analyses [25]
2. Selected DNA preparation kit should provide genomic DNA free from PCR inhibitors and with a  $A_{260}/A_{280}$  ratio between 1.7 and 1.9. Be aware that gDNA concentration is sufficient (*see Notes 7 and 8*) for subsequent dPCR analysis, especially in case of low allele burden. Reduce elution volume if necessary.



**Fig. 2** Typical results for quantitation of JAK2 allele burden using QuantStudio™ 3D Digital PCR 20K Chip and QuantStudio™ 3D Digital PCR AnalysisSuite™ Software. (a) Chip image. Green flag in upper right corner denotes adequate separation of digital dye signals. (b) Scatter plot: In the scatter plot, fluorescence signals are shown for each well of the chip. Wells with JAK2 V617F mutant alleles are represented by FAM™-signals (blue); wells with wild type alleles are represented by VIC®-signals (red); wells with both mutant and wild-type alleles are represented by FAM™ and VIC®-signals (green); and wells without any alleles (passive reference) are represented by ROX®-signals (yellow). FAM™, 6-carboxyfluorescein; VIC®, 4,7,2'-trichloro-7'-phenyl-6-carboxyfluorescein; ROX®, 6-carboxy-X-rhodamine. (c) Results obtained by QuantStudio™ 3D Digital PCR AnalysisSuite™ Software, given for displayed chip-image. Target/total % gives JAK2 V617F allele burden (23.4%). Copies per reaction (rxn) gives mean number of copies per chip partition and copies per microliter gives quantity of sample on chip for mutant and wild type allele, FAM™ and VIC®-signals, respectively. Additionally, a dilution factor can be entered into the software for assessing quantity of sample stock. Validity of results is given by predefined 95% confidence interval (CI) and precision, respectively, where the latter is defined as spread of CI around two sample concentrations at a given CI. Software shows no data for VIC concerning total percentage together with its CI

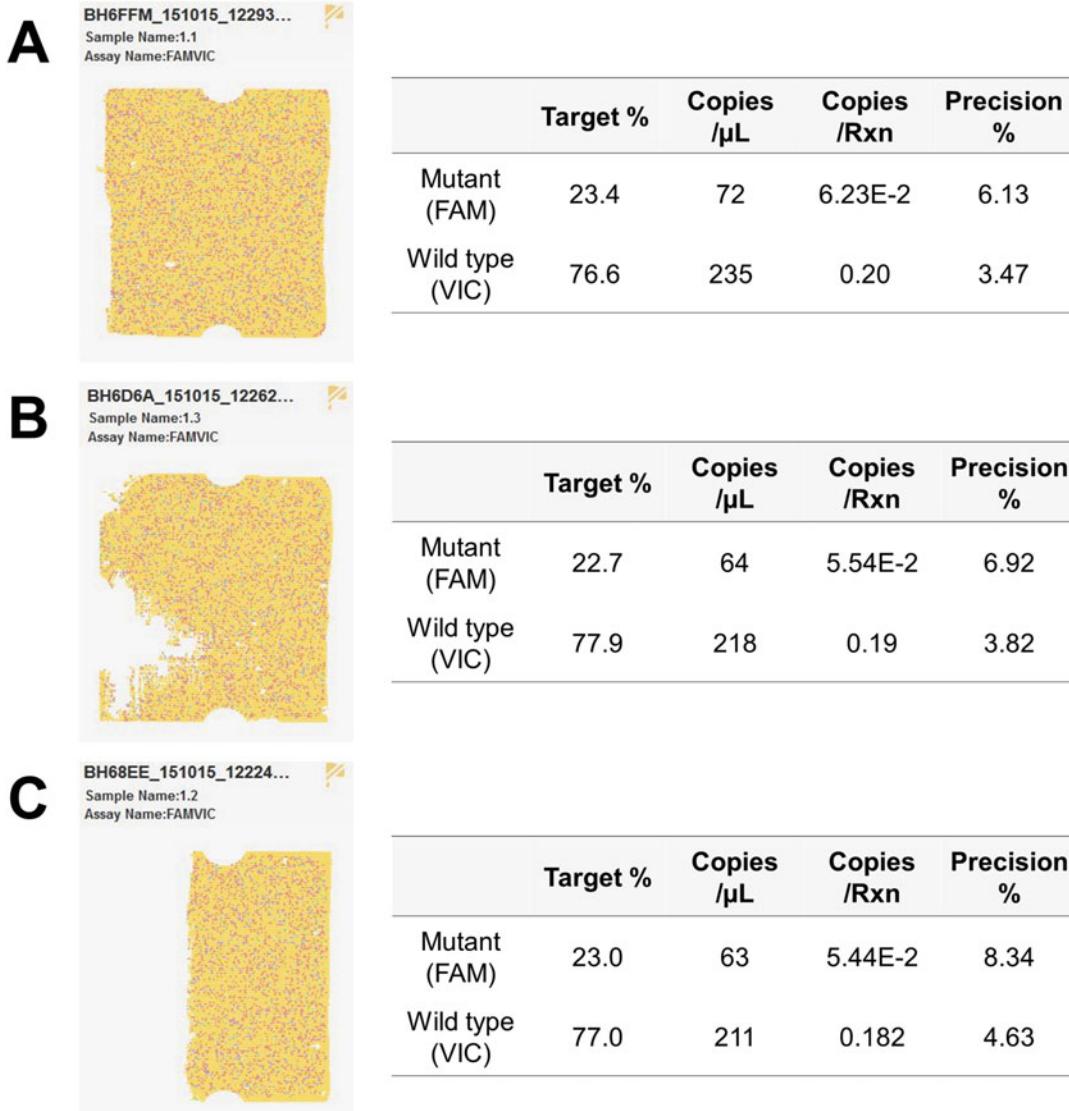
3. Accurate gDNA quantification is recommended in assessing optimal starting copy number in dPCR. A fluorescence based DNA quantification method is more accurate compared to spectrophotometry, and, therefore, should be preferred, at least in samples with a low DNA concentration.
4. Thermo Fisher offers a wet lab-validated TaqMan® SNP Genotyping Assay for JAK2 V617F quantitation, also.
5. Thermo Fisher introduced in 2015 a new generation of chips, named v2, which, in contrast to the previous v1 chip generation, enables improved UV adhesive-free sealing. Digital PCR

**A****B**

**Fig. 3** Results for quantitation of JAK2 allele burden obtained from chips with different sample dilutions. Figure shows chip images, scatter plots, and major results from two chips of the same sample, but from different DNA dilutions (**a**: 300 copies/ $\mu$ L; **b**: 1000 copies/ $\mu$ L). Both chips are of adequate quality and show comparable quantitation results for target DNA, despite differences in initial template concentration. Of note, precision is better in chip B, using higher template concentrations in PCR compared to chip A

employing v2 chips requires QuantStudio 3D™ Digital PCR Master Mix v2. However, v1 chips together with Master Mix v1 are no more available.

6. The mass of one haploid human genome is given between 3.30 and 3.51 pg in the literature, depending on assumed human genome length (3.0 billion base pairs versus 3.2 billion base pairs), the assumed average molecular weight of a nucleotide pair (650 g/mol versus 660 g/mol), and rounding differences of the avogadro constant, respectively. Differences in the conversion factor, however, will not significantly affect validity of dPCR results. Alternatively, use a copy number calculator (available for example on <http://www.thermofisher.com/at/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/dna-copy-number-calculator.html>).
7. Be aware that the specification of the gDNA concentration refers to the gDNA concentration in the dPCR reaction and not to in the gDNA stock solution. For instance, assuming a PCR-mix volume of 34.8  $\mu$ L (in case of using two chips for one sample to perform duplicates) and a desired template



**Fig. 4 (a–c)** Results for quantitation of JAK2 allele burden obtained from chips with different quality. Figure shows chip images and major results from three chips of the same sample using same template concentration in PCR, but differ in quality. Concentration was about 300 copies/µL and therefore at the lower end of optimal concentration range (200–2000 copies/µL). Although chips exhibit different quality and a reduction of precision from high quality (a) to low quality (c), estimation of target DNA delivers comparable results

concentration of 1000 copies/µL, PCR mix contains 34,800 copies or 122 ng gDNA in total.

8. In case of low allele burden (<1%) higher DNA concentrations (up to 6000 copies/µL) may be used, which results in an improved precision of calculated copy number of the rare allele.

If you have limited sample concentrations, precision can be optimized using multiple chips. Impact of employed sample concentration and number of chips, respectively, on precision is shown in Fig. 1.

9. A survey of the workflow including chip loading is demonstrated in the YouTube video “QuantStudio 3D Digital PCR Workflow Video” (<https://www.youtube.com/watch?v=PjgwDhN63Zc>).
10. To avoid spreading too much immersion fluid on the chip later on, depress the plunger and release some immersion fluid to get a feeling for pressure ratio. Note that the fluid should be used within 1 h after uncapping the syringe.
11. Confirm that the blade is properly seated by pressing firmly against the loader head and checking its position on the back-side of the loader head as well. A correct position of the loader head and blade is essential since it is responsible for the application of PCR reaction mixture on the chip. However, it is possible to install the blade after filling in the sample.
12. To avoid contact with the sticky film beneath, gently grasp only the outside margin of the lid. Remove the protective film quickly. Otherwise, it is likely that you will remove the sticker as well. V2 Chips may not fit properly in first generation chip loaders. In that case, Life Technologies provides a “chip gripper upgrade kit.”
13. Another possibility to properly apply the lid in the lid nest is to remove the sticky film after insertion of the lid into the nest. Therefore, you will need to tightly press the lid nest button while removing the film so that the lid stays fixed.
14. Filling the loading port appropriately is utterly important for chip quality. Therefore it will need precise handling and requires some practice. First of all, it is important to distribute the solution evenly in the port so that each side of the chip will contain the same amount of solution later on. Further, one should avoid discarding some solution by deflecting the blades through applying too much pressure on the tip. To avoid air bubbles in the port, do not depress the pipette to the second stop. However, if a bubble appears, try to remove it by carefully taping the port from both sides. Make sure that the port is in the right position afterward and do not tap too hard to avoid shifting the loading head or discarding solution. Note that the exact adjustment of the loading head is essential for proper chip load. If you recognize that the loading head does not apply solution properly, contact service support. Another option is to fill the loading blade prior to installation into the loading head, which might ease filling procedure, since one can choose optimal filling angle. In this particular case, take care to not discard any dPCR solution during installation of the blade.

15. It is important to apply the immersion fluid about 20 s after loading to avoid evaporation of reaction mixture. Make sure that the immersion fluid covers the surface of the chip but does not flood the whole chip case. When applying the sticky film of the lid later on, make sure that there is no immersion fluid on the edges of the chip case. In case there is immersion fluid on the edges, you can remove it with a wipe that has been sprayed with some isopropanol.
16. A pressure of >20 pounds (lb) which approximates about 9 kilograms (kg) is recommended for tight sealing. Though it is difficult to estimate the correct pressure, it can be recommended to press 20 s with moderate pressure and in a 90° angle with the chip. Make sure that the chip loader is in an appropriate position close to the operator's body and cannot tilt forward.
17. Choosing the correct angle is important to release air bubbles between the chip and the chip lid. Remaining bubbles can interfere with PCR and impair analysis of the chip later on. The 45° angle provides a good starting position although the operator will need to adjust the chip position in order to position the bubble beneath the chip port.  
It is important to pour the immersion fluid in slowly, so that the filling can be stopped easily when the chip is nearly filled (remaining a small air bubbles <2–3 mm at the port).
18. Make sure not to touch chip window too firmly after the chip is sealed. Pressure on the chip window might expel reaction mixture out of the wells.
19. For quantitation of JAK2 V617F allele burden PCR conditions are modified to 56 °C annealing temperature instead of 60 °C annealing temperature recommended by the manufacturer as default annealing temperature. The lower annealing temperature results in an improved cluster separation.
20. The elevated fill pot makes sure that small bubbles will float to the top of the case, not impairing the amplification.
21. You can remove chips after final extension step, if the cycler has cooled down to <25 °C. Avoid fast changes in temperature to prevent condensation of chips. According to specifications of the manufacturer it is possible to leave chips on 10 °C for up to 24 h (*see Note 22*, also). Chips must be brought to room temperature prior to imaging.
22. To our experience, chips can be stored several days at room temperature in a clean and dark place with consistent quality.
23. There is no need to wait until the analysis is completed. To start with the next run, just remove the prior chip and load the next one.

24. It is recommended in the manufacturer's manual, that a chip should be reread within 1 h after the initial read. In fact, the chip can be reread much longer (e.g., on the next day) with consistent quality if it is properly protected from light. Chip storage at room temperature is sufficient.
25. The software can be run using any compatible operating system and web browser but it is optimized for Microsoft® Windows® operating system and use on Google® Chrome™. Please refer to user guide to check requirement since performance may vary depending on operation system and configuration.
26. We recommend using the help-tool to get started with analysis. It contains demonstration tools, i.e., video demonstrations, helping you to make a start.
27. It is important to define rare allele dye, displaying the fluorescence dye, which detects the rare allele. Note that the identical sample names must be used in case of duplicates.
28. It is possible to include or exclude wells from analysis by adjusting the quality threshold. This assessment can be applied to all chips of the assay. The different colors of the quality flags are based on well quality threshold and other data characteristic. The chip quality assessment can help to detect problems with application of dPCR solution by visualizing the spatial distribution of data across the chip.
29. There are several tools, which help you to adjust data points. You can adjust the scatter plot axis and size of data points as well as the zoom. After you defined optimal conditions, you may use the lasso tool to choose data points to include or exclude ("Undetermined") from your data analysis. Separate data clouds (FAM™; VIC®, ROX™) accurately to improve precision of results. Especially VIC® and ROX™ clouds often overlap, concerning the huge amount of data points. Loss of some data points through exclusion is readily tolerated.
30. One can lower and therefore improve the calculated precision by measuring multiple chips for the same sample and combining them in one chip (*see Fig. 1*). The software automatically combines chips with the same name in one virtual chip. The number of chips is listed in results ("Chips"). Furthermore, it is recommended to fill in dilution factor as the calculated sample quantity (copies/ $\mu$ L) will not display your stock concentration, but the gDNA concentration in dPCR reaction.
31. Outliers can be excluded from calculation with the "omit chip" option in the "Review Data" tab. Digital calls contain data of negative, positive or undetermined calls and may help you identify problems with chip load or detection quality. For optimal output, >17,000 calls should have been qualified by quality threshold.

## References

1. Levine RL et al (2005) Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell* 7:387–397
2. Baxter EJ et al (2005) Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet* 365:1054–1061
3. Kralovics R et al (2005) A gain-of-function mutation of JAK2 in myeloproliferative disorders. *N Engl J Med* 352:1779–1790
4. James C et al (2005) A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature* 434:1144–1148
5. Vardiman JW et al (2009) The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood* 114:937–951
6. Passamonti F et al (2006) Relation between JAK2 (V617F) mutation status, granulocyte activation, and constitutive mobilization of CD34+ cells into peripheral blood in myeloproliferative disorders. *Blood* 107:3676–3682
7. Lippert E et al (2006) The JAK2-V617F mutation is frequently present at diagnosis in patients with essential thrombocythemia and polycythemia vera. *Blood* 108:1865–1867
8. Passamonti F, Rumi E (2009) Clinical relevance of JAK2 (V617F) mutant allele burden. *Haematologica* 94:7–10
9. Langabeer SE et al (2015) Molecular diagnostics of myeloproliferative neoplasms. *Eur J Haematol* 95:270–279
10. Vannucchi AM, Pieri L, Guglielmelli P (2010) JAK2 allele burden in the myeloproliferative neoplasms: effects on phenotype, prognosis and change with treatment. *Ther Adv Hematol* 2:21–32
11. Vannucchi AM, Guglielmelli P (2013) JAK2 mutation-related disease and thrombosis. *Semin Thromb Hemost* 39:496–506
12. Koren-Michowitz M et al (2012) JAK2V617F allele burden is associated with transformation to myelofibrosis. *Leuk Lymphoma* 53:2210–2213
13. Alvarez-Larrán A et al (2014) JAK2 V617F monitoring in polycythemia vera and essential thrombocythemia: clinical usefulness for predicting myelofibrotic transformation and thrombotic events. *Am J Hematol* 89:517–523
14. Tefferi A et al (2008) Low JAK2V617F allele burden in primary myelofibrosis, compared to either a higher allele burden or unmutated status, is associated with inferior overall and leukemia-free survival. *Leukemia* 22:756–761
15. Guglielmelli P et al (2009) JAK2V617F mutational status and allele burden have little influence on clinical phenotype and prognosis in patients with post-polycythemia vera and post-essential thrombocythemia myelofibrosis. *Haematologica* 94:144–146
16. Quintás-Cardama A et al (2009) Pegylated interferon alfa-2a yields high rates of hematologic and molecular response in patients with advanced essential thrombocythemia and polycythemia vera. *J Clin Oncol* 27:5418–5424
17. Kiladjian J-J et al (2008) Pegylated interferon-alfa-2a induces complete hematologic and molecular responses with low toxicity in polycythemia vera. *Blood* 112:3065–3072
18. Jovanovic J (2013) Vet al. Establishing optimal quantitative-polymerase chain reaction assays for routine diagnosis and tracking of minimal residual disease in JAK2-V617F-associated myeloproliferative neoplasms: a joint European LeukemiaNet/MPN&MPNr-EuroNet (COST action BM0902) stu. *Leukemia* 27:2032–2039
19. Lippert E et al (2009) Concordance of assays designed for the quantification of JAK2V617F: a multicenter study. *Haematologica* 94:38–45
20. Kinz E, Leiherer A, Lang AH, Drexel H, Muendlein A (2015) Accurate quantitation of JAK2 V617F allele burden by array-based digital PCR. *Int J Lab Hematol* 37:217–224
21. Majumdar N, Wessel T, Marks J (2015) Digital PCR modeling for maximal sensitivity, dynamic range and measurement precision. *PLoS One* 10:e0118833
22. Larsen TS, Pallisaard N, Møller MB, Hasselbalch HC (2008) Quantitative assessment of the JAK2 V617F allele burden: equivalent levels in peripheral blood and bone marrow. *Leukemia* 22(1):194–195. <http://www.ncbi.nlm.nih.gov/pubmed/?term=17625603>
23. Takahashi K et al (2013) JAK2 p.V617F detection and allele burden measurement in peripheral blood and bone marrow aspirates in patients with myeloproliferative neoplasms. *Blood* 122:3784–3786
24. Hermouet S et al (2007) Comparison of whole blood vs purified blood granulocytes for the detection and quantitation of JAK2(V617F). *Leukemia* 21:1128–1130
25. Gong JZ et al (2013) Laboratory Practice Guidelines for Detecting and Reporting JAK2 and MPL Mutations in Myeloproliferative Neoplasms. *J Mol Diagn* 15:733–744



# Chapter 16

## Novel Multiplexing Strategies for Quantification of Rare Alleles Using ddPCR

Miguel Alcaide and Ryan D. Morin

### Abstract

Droplet digital PCR (ddPCR) has come to be regarded as the gold standard for the ultrasensitive detection and absolute quantification of closely related DNA sequences within complex mixtures. Most ddPCR assays to date, however, rely on sets of hydrolysis probes conjugated with dyes having different emission spectra to allow independent counting of rare mutant and wild-type alleles. Here, we describe a set of novel strategies that leverage the simultaneous detection and quantification of both mutant and wild-type alleles with a single hydrolysis probe. Variants of these strategies empower multiplexing and a more cost-effective approach for concurrent screening of multiple genetic variants.

**Key words** Digital PCR, Genetic testing, Noninvasive genetic profiling, Rare allele detection, Cell-free DNA, Clonal evolution, Pathogen detection

---

### 1 Introduction

The power to detect and quantify rare genetic variants within complex DNA mixtures with both high sensitivity and specificity has profound implications both in healthcare and environmental analysis. Noninvasive genetic profiling of cell-free fetal DNA in the maternal bloodstream for prenatal testing [1, 2] and the investigation of circulating tumor DNA (ctDNA) in cancer diagnosis and monitoring by means of liquid biopsies [3–5] are among the most cited applications. Low-level pathogen detection is also critical to prevent waterborne or foodborne illnesses and has also recently shown great utility to monitor infections in anatomic fluids, the emergence of antimicrobial resistance, or the proliferation of genetically modified organisms [6–10]. Most of the standard molecular methods such as quantitative PCR, traditional Sanger sequencing, microarrays, fluorescence in-situ hybridization (FISH), immunology-based methods or even generic next generation sequencing (NGS) of PCR products cannot reliably discern between real and putatively false variants when rare allele

frequencies are around or below 1% [5]. The advent of digital PCR and novel NGS methods involving molecule barcoding strategies have pushed the limits of detection for rare alleles to one in several thousand or even hundreds of thousands [11–14].

Digital PCR involves the partition of a single PCR reaction into thousands or even millions of parallel microreactions for the confinement of single (or low numbers of) copies of the locus of interest within prefabricated chambers or nanoliter/picoliter droplets generated by water-in-oil emulsions [15, 16]. When compared to other PCR-based methods, digital PCR is less sensitive than conventional qPCR to the inhibitors that can be present, for example in crude samples such as water, soil, or anatomical fluids, thus offering a more robust testing platform [17, 18]. Digital PCR enables the straightforward and absolute quantification of rare alleles without the need of reference samples to construct standard curves and can discriminate among closely related DNA sequences with exquisite specificity. Single nucleotide polymorphisms and small indels underlie numerous inherited genetic disorders and somatic alterations of this variety have been recurrently reported as mutation “hotspots” in many human cancers. Such simple genetic alterations can similarly drive the adaptive evolution of disease-associated microorganisms, offering a variety of avenues by which sensitive genotyping can provide potential clinical utility [19–21].

Hydrolysis probes conjugated with dyes of different emission spectra allow the simultaneous detection and quantification of highly similar DNA sequences at single base pair resolution. Duplex digital PCR is an appealing technology because it reduces technical errors associated with parallel reactions and offers significant savings in time, overall costs and the amount of clinically or environmentally precious samples needed for a given assay. Even though currently available digital PCR systems are often restricted to two optical detection channels (with exceptions, for instance [22]), higher order multiplexing is nonetheless enabled in these systems by the use of different concentrations of probes labeled with the same dye but against different targets, or by the use of variable ratios of two probes labeled with different dyes targeting the same sequence, thus giving them a unique location in 2D fluorescence space [23]. The accumulation of unquenched fluorophores associated with probe hydrolysis after end-point PCR permits measuring the amount and type of fluorescence emitted by each individual partition and distinguishing those which are positive and those which are negative for a target of interest. Established statistical models relying on the Poisson distribution are then applied to undertake absolute quantification of different DNA templates in the initial sample [23]. Until recently, nondiscriminating assays to detect multiple mutations such as “wild-type negative” or “drop-off” assays required a minimum of two hydrolysis probes labeled

with different fluorophores [23–25]. Similarly, discriminating amplitude-based or ratio-based multiplexed assays specifically targeted rare and wild-type alleles with probes conjugated with different dyes [23, 26].

In this chapter, we describe assay design strategies which reduce the number of probes necessary for successful multiplex detection and quantification of both rare and wild-type alleles. These employ either discriminating approaches that exclusively use one or more mutant-specific hydrolysis probes, which also quantify the *wild-type* allele(s), or nondiscriminating approaches exclusively using one or more probes against wild-type DNA sequences (“inverted” ddPCR assays), yet which also quantify *mutant* alleles. In addition, these can be designed to detect and quantify one or several loci in a single reaction by the presence of one or two primer pairs targeting them. These cost-effective and straightforward assays are possible owing to the less efficient hydrolysis of imperfectly annealed mutant or wild-type probes—where single base mismatches between the probe and the DNA template occur—and where the consequences on each target molecule can be captured in digitized reactions as droplets with lowered fluorescence amplitude(s) [27]. The fundamentals of other multiplex digital PCR strategies for rare allele detection have been recently reviewed in detail [23] and therefore will not be addressed in the present chapter.

---

## 2 Materials

1. ddPCR™ Supermix for Probes (No dUTP) (Bio-Rad, Hercules, CA, USA; Cat. No. #186-3023, 186-3024 or 186-3025) (*see Note 1*).
2. Automated Droplet Generation Oil for Probes (Bio-Rad, Hercules, CA, USA; Cat. No. #1864110) (*see Note 2*).
3. ddPCR™ Droplet Reader Oil (Bio-Rad, Hercules, CA, USA; Cat. No. #1863004) (*see Note 3*).
4. DG32™ Automated Droplet Generator Cartridges (Bio-Rad, Hercules, CA, USA; Cat. No. #1864108 and #1864108) (*see Note 4*).
5. Pipet Tips for the AutoDG™ System (Bio-Rad, Hercules, CA, USA; Cat. No. #1864108 and #1864120).
6. Pipet Tip Waste Bins for the AutoDG™ System (Bio-Rad, Hercules, CA, USA; Cat. No. #1864108 and #1864125) (*see Note 5*).
7. ddPCR™ 96-Well PCR Plates (Bio-Rad, Hercules, CA, USA; Cat. No. #1864108 and #12001925) or Eppendorf twin. tec® PCR Plate 96, semiskirted (Eppendorf, Hamburg, Germany; Cat. No: #951020303).

8. PCR Plate Heat Seal, foil, pierceable (Bio-Rad, Hercules, CA, USA; Cat. No. #1864108 and #1814040) (*see Note 6*).
9. PCR-Cooler (Eppendorf, Hamburg, Germany; Cat. No: #022510541) (*see Note 7*).
10. Automated Droplet Generator (AutoDG™ System; Bio-Rad, Hercules, CA, USA; Cat. No. #1864101) (*see Note 8*).
11. QX200™ Droplet Reader (Bio-Rad, Hercules, CA, USA; Cat. No. #1864003) (*see Note 9*).
12. PXI™ PCR Plate Sealer (Bio-Rad, Hercules, CA, USA; Cat. No. #1814000) (*see Note 10*).
13. C1000 Touch™ Thermal Cycler with 96-Deep Well Reaction Module (Bio-Rad, Hercules, CA, USA; Cat. No. #1814000) (*see Note 11*).
14. Assay-specific hydrolysis probes: ZEN™ Double-Quenched or LNA PrimeTime® probes (Integrated DNA technologies, Coralville, IA, USA) conjugated with 6-FAM™ or HEX™ dyes and 3' Iowa Black® FQ dark quenchers (*see Note 12*).
15. QuantaSoft™ software, Regulatory edition (Bio-Rad, Hercules, CA, USA; Cat. No. #1864011).
16. Rainin (Mettler-Toledo, Giessen, Germany) or Eppendorf filter tips (Eppendorf, Hamburg, Germany; Cat. No: #951020401) (*see Note 13*).
17. ddPCR™ Buffer control for probes (Bio-Rad, Hercules, CA, USA; Cat. No. #1863052).
18. Ultra-pure water.
19. 10 mM Tris-HCl (pH = 8.0) or 10 mM Tris-HCl, 0.1 mM EDTA (pH = 8.0).
20. Buffer AVE (Qiagen, Venlo, Netherlands).
21. Buffer ATE (Qiagen, Venlo, Netherlands).
22. Restriction endonucleases (optional, *see Note 14*).
23. Uracil DNA Glycosylase (optional, *see Note 15*).
24. Wild-type DNA to be used as negative control
25. Covaris M220 ultrafocused sonicator or similar (optional; Covaris, Woburn, MA, USA).

---

### 3 Methods

#### 3.1 Sample Preparation

This protocol has been validated and optimized using genomic DNA extracts from cells or fresh tissues, prepared with various Qiagen purification kits according to the manufacturer's instructions, and eluted either in 10 mM Tris-HCl (pH = 8.0) or 10 mM Tris-HCl, 0.1 mM EDTA; Cell-free DNA (cfDNA) extracts eluted

in buffer AVE (Qiagen, Venlo, Netherlands); and FFPE-derived DNA extracts (*see Note 15*) eluted in buffer ATE (Qiagen, Venlo, Netherlands) (see examples, including FFPE samples, in [27]). Cell-free DNA does not need any treatment prior to digital PCR and can be extracted from plasma using any number of commercial kits. Preanalytical handling of blood for cfDNA processing is reviewed elsewhere [5]. Restriction endonuclease digestion or mechanical shearing is particularly recommended when the concentration of intact genomic DNA exceeds 72 ng per 22 µl ddPCR reaction in order to facilitate proper droplet formation. Enzymatic digestion can be accomplished with a restriction endonuclease that does not cut target or reference amplicons (*see Note 14*). Mechanical fragmentation can be performed using a Covaris ultrafocused sonicator (Covaris, Woburn, MA, USA). Fragmentation of sample DNA not only reduces sample viscosity but may also improve accessibility of assays to certain target regions even at lower DNA concentrations.

### **3.2 Assay Design Considerations**

#### **3.2.1 Amplicon Size**

In designing primers, the PCR amplicon size should be kept to a minimum, ideally between 60 and 80 bp, with shorter amplicons preferred. This is particularly relevant when using fragmented DNA, for example cell-free DNA, as template. Long PCR amplicons are typically observed to have lower fluorescence amplitudes and higher background noise. In addition, as reviewed elsewhere [5], longer PCR amplicons can compromise assay sensitivity in cfDNA as a substantial fraction of the typically short DNA templates could go undetected because they do not contain binding sites for both assay primers [28, 29]. Below is a list of considerations for the design of assays with good performance. However, it is noted that complying with the full list of recommendations listed can be challenging due to the sequence characteristics of the target and its genomic context. For useful assay design tools that generate candidate primer and probe sequences for the intended assays, see below (and *see Note 16*).

#### **3.2.2 PCR Primers**

1. Run BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) or BLAT ([https://genome.ucsc.edu/cgi-bin/hgBlat?com\\_mand=start](https://genome.ucsc.edu/cgi-bin/hgBlat?com_mand=start)) on potential primer sequences to avoid cross reactivity with other regions of the genome
2. The two primers must have a similar melting temperature ( $T_m \pm 2$  °C) and GC contents generally ranging between 35% and 65%
3. Primer length must not exceed 30 nucleotides
4. Avoid sequences folding into strong secondary structures, self-dimers or heterodimers as well as long homopolymer runs, particularly polyG >4 bases long

5. Ensure that there are no common SNPs overlapping the 3' ends of the primers, as these could decrease amplification efficiency.

### **3.2.3 Hydrolysis Probes**

1. Use BLAST or BLAT to screen potential probe sequences to avoid cross reactivity with other regions of the genome.
2. Avoid G bases at the 5' end as they have a well-demonstrated quenching effect over some fluorophores.
3. Probes are commonly intended to have a Tm 6–10 °C warmer than that of primers during traditional qPCR experiments. Although hydrolysis probe-based assays designed following these criteria can work well in ddPCR reactions, Tm values just 3 °C higher have been observed to perform better. Special modifications such as LNA bases (Integrated DNA Technologies, Coralville, IA, USA) or MGB modifications (minor groove binder, Applied Biosystems, Foster City, CA, USA) can be used to increase the Tm of probes without extending their length.
4. Avoid long homopolymer G runs (>4 bases), secondary structures and the formation of homodimers or heterodimers with the PCR primers
5. Place the genomic position to be interrogated as centered as possible within the sequence of the probe
6. Select the sense or antisense strand to leave at least a five nucleotide distance with respect to the 3' end of the primer that anneals to the same strand. Probes designed to one strand may be less prone than the other to form secondary structures or dimers.
7. Conjugate the probes with 6-FAM™ (Channel 1) or HEX™ (Channel 2) fluorescent dyes and add ZEN double-quenchers or Iowa Black® Dark quenchers (Integrated DNA technologies, Coralville, IA, USA). Although not evaluated for this particular application, Black Hole Quenchers® (LGC Biosearch Technologies, Novato, CA, USA) and TaqMan® minor groove binder (MGB) probes (Life Technologies, Carlsbad, CA, USA), to name a few alternative probe designs, are also expected to perform well for this particular application (*see also Note 12*).

## **3.3 Setting Up and Running ddPCR Experiments**

In this section, general instructions applicable to each of the uniplex or multiplexing strategies described in this chapter are given for reaction setup, droplet generation, thermocycling, and droplet reading/data acquisition. Refer to the sections following this one (Subheadings 3.5 through 3.9) for detailed guidance on the relevant uniplex or multiplexing approach(es) being followed.

**Table 1**

**Suggested volumes and final concentrations during the preparation of the different discriminating and nondiscriminating hydrolysis probe-based ddPCR experiments described in this chapter**

	<b>Uniplex (Single- locus)</b>	<b>Uniplex (Dual-locus)</b>	<b>Multiplex (Single locus)</b>
2× SuperMix ddPCR for Probes (no dUTP)	11 µL (1×)	11 µL (1×)	11 µL (1×)
Primers Premix—10 µM (each) stock	2.2 µL (1 µM) (1 Pair)	2.2 µL (1 µM) (2 Pairs)	2.2 µL (1 µM) (1 Pair)
6-FAM™ Hydrolysis Probe (Mutant- or Wild type Specific)—5 µM stock	1.5 µL (0.34 µM)	1.5 µL (0.34 µM)	2.2 µL (0.5 µM) Mutant-specific only
6-FAM™ Hydrolysis Probe (Mutant-Specific)—5 µM stock			0.8 µL (0.18 µM)
6-FAM™ Hydrolysis Probe (Mutant-Specific)—5 µM stock			0.33 µL (0.075 µM)
HEX Hydrolysis Probe (Mutant- or Wild type-Specific)—5 µM stock		1.5 µL (0.34 µM)	1.4 µL (0.3 µM) Mutant-specific only
Dark probes—5 µM (each) stock (optional)	0–1.5 µL (variable)	0–1.5 µL (variable) (0, 1, or 2 locus)	
DNA + Ultra Pure Water	5.8–7.3 µL	2.8–5.8 µL	4.07 µL
Total	22 µL	22 µL	22 µL

All reactions are made to a final volume of 22 µL/sample

1. Droplet digital PCR reactions must be set to a minimum final volume of 22 µL to avoid mechanical errors during automated droplet generation. This volume should contain 11 µL of 2× ddPCR™ Supermix for Probes (no dUTP). The remaining 11 µL encompass PCR primers (1 µM final concentrations), hydrolysis probes (0.075–0.5 µM final concentrations, *see Note 17*), and a variable amount of input DNA (*see Note 18* and Table 1). All reagents should be thawed at room temperature, vortexed and briefly spun to eliminate concentration gradients formed during storage. After preparing, the desired reactions according to Table 1, individual reactions are dispensed at room temperature into Bio-Rad ddPCR™ or Eppendorf twin. tec® 96-well semiskirted PCR plates.
2. For droplet generation, follow the QX100 or QX200 Droplet Generator Instruction Manual (#10026322 or 10031907, respectively) for manual droplet generation if an automated droplet generator is not available in your laboratory. If you

**Table 2**  
**Recommended cycling protocol for ddPCR reactions relying on hydrolysis probes**

Cycling step	Temperature (°C)	Time	Number of cycles
Enzyme activation	95 °C	10 min	1
Denaturation	94 °C	30 s	40
Annealing/Extension	Optimized for individual assays (52–62 °C)	1 min	
Enzyme deactivation	98 °C	10 min	1
Hold	4 °C	Infinite	1

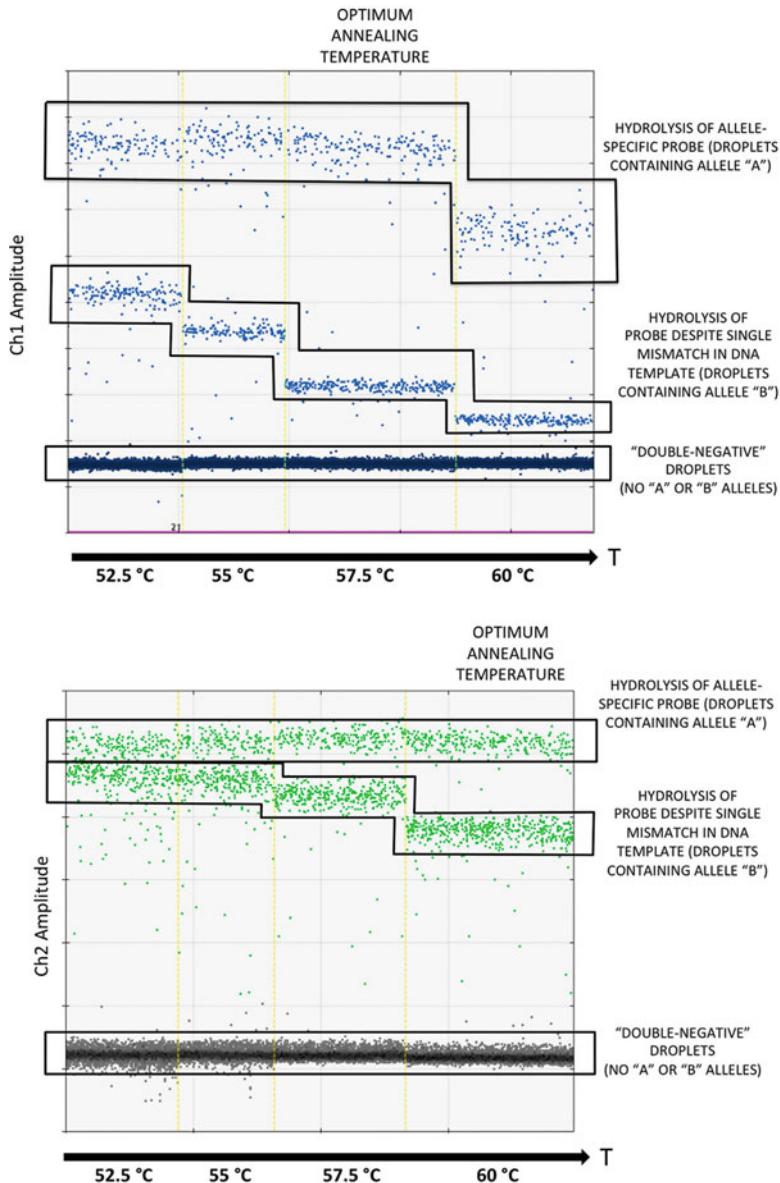
choose to assemble the reactions on ice, it is important to equilibrate reactions at room temperature for at least 3 min before droplet generation.

3. A minimum of 8 samples dispensed in one single column is required for automated droplet generation (*see Note 19*). Place your sample plate in the designated area within the automated droplet generator (left plate compartment) and a clean plate onto a plate cooler in its corresponding spot (right plate compartment) (*see also Notes 2, 4 and 5*).
  4. Click on *Configure Sample Plate* button and select the specific columns of the plate containing samples. Confirm and initiate automated droplet generation (<5 min per 8 samples or column of samples).
  5. Check for error messages in the Automated Droplet Generator screen (*see Note 20*) and proceed if the operation completed successfully. Refer to Chapter 4 of the Automated Droplet Generator Instruction Manual (<http://www.bio-rad.com/webroot/web/pdf/lsl/literature/10043138.pdf>) and/or contact Tech support for troubleshooting.
  6. After all reactions have been converted to droplets, cover the collection plate with pierceable foil and seal the foil with heat (180 °C) in the PX1™ PCR Plate Sealer (*see Note 6*).
  7. Introduce the plate inside the C1000 Touch™ Thermal Cycler and run the preloaded program described in Table 2, with the lid preheated at 105 °C, setting the volume of the reaction at 40 µL and the ramp rate for each step of the cycle at 2 °C/s (*see Note 21*).
- 3.4 Data Acquisition and Analysis**
1. Once the PCR amplification is complete, place the sealed plate containing samples in the QX200™ Droplet Digital PCR reader (*see Note 3*).
  2. Open the QuantaSoft™ software and set up a new plate layout. Double click on a particular well to open the well-editor dialog box.

3. For each well, introduce sample name, experimental type (“Absolute quantification” for target counts >100 copies per well, or “Rare Event Detection” for <100 counts per well), the type of supermix should be entered as ddPCR™ Supermix for probes (no dUTP), name of targets, optical channel associated with each target (6-FAM™ in channel 1; HEX™ or VIC™ in channel 2) and type of sample analyzed (e.g., unknown, reference, positive control, negative control, no-template control). Click “Apply” to label the wells and click “OK” when finished.
4. Click on the *Run* button to start the droplet reading process.
5. After the data collection completes, you should inspect your results by clicking on the *Analysis* tab. To perform quantification, you must first set the thresholds and clusters manually across the 2-D amplitude data plots, as guided by the cluster boundaries observed in reference samples, positive controls, negative controls, and no-template DNA controls. Specific instructions for target quantification are detailed below for each of the ddPCR assays described in this chapter.
6. Use copies/ $\mu$ l and allele ratios to estimate the abundance of targeted alleles in your initial samples.

### **3.5 Uniplex “Single-Locus” Assay Validation and Optimization Prior to Use in Multiplexing (One Pair of Primers and One Mutant Allele-Specific Probe)**

We recently demonstrated that ddPCR enables the simultaneous discrimination of mutant and wild-type alleles at single base pair resolution with the use of a unique hydrolysis probe [27]. We observed that imperfectly annealed probes can be degraded, if a single mismatch between the probe and the DNA template occurs, when mutant and wild-type probes are not competing for hybridization in the same microreaction. This phenomenon generates droplets with at least three clearly differentiated fluorescence amplitudes. The lowest fluorescence amplitude relates to the imperfect quenching of the fluorophore attached to the unhydrolyzed probe (“double-negative” droplets, i.e., droplets that do not contain either wild-type or mutant alleles for a given locus) and the highest fluorescence amplitude (“positive” droplets) can be explained by the specific and highly efficient hydrolysis of probes after a perfectly complementary match with the DNA template. Notably, we observed an intermediate fluorescence amplitude band (“low positive” droplets), clearly distinguishable from “double-negative” droplets and “positive” droplets, which we attribute to the less efficient hydrolysis of imperfectly annealed probes (*see Fig. 1*). Absolute quantification of both mutant and wild-type alleles for a given locus using assays that forego, for example wild-type specific probes, is made possible by this feature. Lower cost per assay and enhanced possibilities for multiplexing are the most direct benefits of using such single hydrolysis probes (*see Note 22*). An important preliminary step to leverage the potential of this novel approach



**Fig. 1** Effect of annealing temperature—tested in a single well—for two uniplex ddPCR assays, targeting two mutations at two different loci. The first assay (top panel) exhibits optimal separation between droplets carrying the allele specifically targeted by the probe (“positive” droplets), droplets containing alternative alleles differing in one single base pair (“low positive” droplets), and “double-negative” droplets at 57.5 °C (top, middle, and bottom boxed regions). For the second assay (bottom panel), the optimal annealing temperature is 60 °C, with lower annealing temperatures showing poor differentiation between “positive” and “low positive” droplets. Thus, if these assays were to be run together in a single well, the higher annealing temperature of 60 °C would need to be used. These assays were performed using genomic DNA extracted from fresh tumor biopsies, previously known to carry mutations at the two loci investigated, as DNA template. The first assay relies on a single hydrolysis probe labeled with 6-FAM™ (Channel 1) and the second assay relies on a single hydrolysis probe labeled with HEX™ (Channel 2)

(and the additional ones below) is to determine the best set of conditions and parameters that generate optimal fluorescence amplitude differentiation between different target alleles.

### 3.5.1 Positive and Negative Controls

Samples carrying specific mutations of interest (positive controls) and wild-type DNA (negative controls) are critical to validate uniplex ddPCR assays prior to use in multiplexing. In the absence of relevant biological samples containing the alleles of interest, synthetic double-stranded and sequence-verified DNA fragments such as gBlocks® Gene Fragments (Integrated DNA technologies, Coralville, IA, USA) can be used for control templates (e.g., spiked into wild-type background genomic DNA).

### 3.5.2 Annealing Temperature

Annealing temperature during emulsion PCR is one of the factors with the greatest impact on assay resolution. Certain annealing temperatures provide the best separation between “double-negative,” “positive,” and “low positive” droplets for a given locus (*see* Fig. 1). As an aside, where multiple probes will be present in a single assay reaction (e.g., *see* Subheading 3.7 on uniplex dual-locus reactions), if the tested probes show very different optimum annealing temperatures, it is recommended to redesign one of the them prior to multiplexing in order to achieve a more similar temperature optimum for both that gives good discrimination of clusters for each of the independent loci (*see* Note 23).

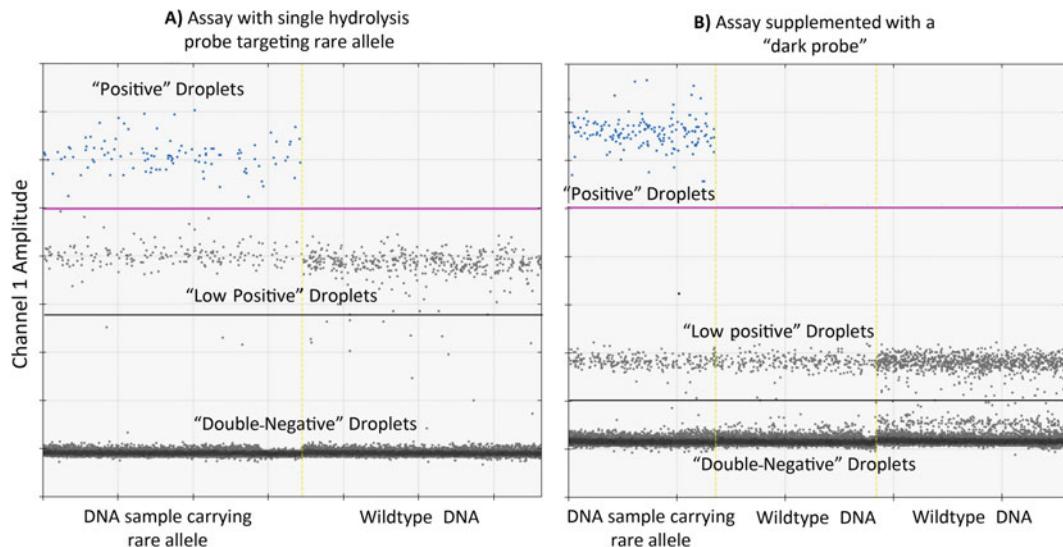
### 3.5.3 Inclusion of “Dark Probes”

The inclusion of an oligonucleotide that matches the sequence of abundant wild-type allele but lacking a conjugated fluorescent dye (i.e., “dark probes,” *see* for example [30]) can enhance the separation between the fluorescence amplitudes associated with rare and wild-type alleles (Fig. 2) (*see* Note 24).

### 3.5.4 Running and Analyzing the Assay

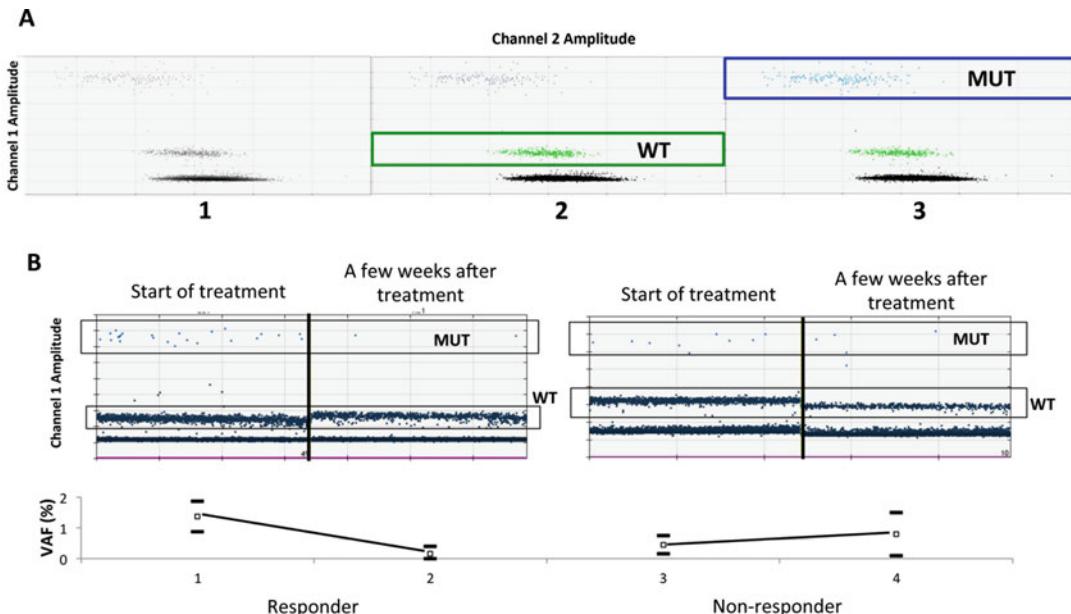
Standard setup and preparation of discriminating uniplex ddPCR assays is detailed in Table 1. After thermocycling the reactions and then reading them in the droplet reader, quantification of mutant and wild-type alleles is accomplished following manual identification of clusters in the QuantaSoft™ software (*see* for example Fig. 3, Panel a).

Under the *Setup* mode, select wells for unknown samples, positive controls, negative controls, and no-template DNA controls (if available), all run with the same assay. For the determination of allele counts and mutant fractional abundance, ensure that the two optical channels have been associated to the proper target (i.e., allele) in the *Setup* menu and display the data on the 2D *Amplitude* plot after clicking on *Analyze*. You will need to associate “positive droplets” carrying mutant alleles with the FAM channel and “low positive droplets” carrying wild-type alleles with the HEX channel. To achieve this, use the rectangular (or circular or freehand) *Multi-*

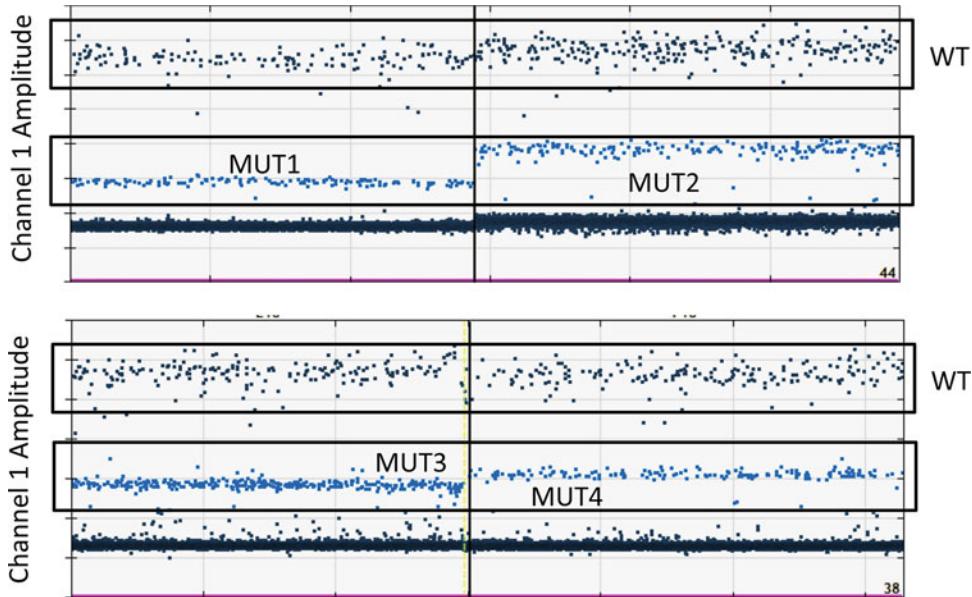


**Fig. 2** (a) The inclusion of a wild-type oligonucleotide not conjugated with a fluorescent dye (i.e., a “dark probe” as illustrated in panel b) enhances the separation of clusters of droplets carrying the rare or wild-type alleles when running a uniplex ddPCR assays. Suggested manual thresholds for the independent quantification of rare (magenta line, blue droplets) and wild-type alleles (black line, intermediate band of black droplets) are also indicated. We used a DNA extract derived from a fresh tumor biopsy sample carrying the mutation specifically targeted by the probe and a commercially available wild-type DNA sample as positive and negative controls, respectively

*well thresholding selection tool*, click on the  $-/-$  black button and enclose all droplets within the shape chosen to initially designate all droplets as “double-negative” droplets. Then, click on the  $-/+$  green button (HEX/VIC channel) and using the selection tool, enclose and define the cluster of “low positive” droplets carrying wild-type alleles. Next, click on the  $+/-$  blue button (FAM channel) to similarly define the cluster of “positive” droplets carrying mutant alleles. Hovering the mouse over the positive and negative controls in the *well-editor* (which highlights their clusters in the 2D plot) may assist you in defining the boundaries between the different fluorescence clusters when multiple samples are concurrently displayed in the 2D amplitude plots (see Fig. 3, Panel a). Absolute counts of mutant alleles per  $\mu\text{L}$  of sample are displayed in the top table of the QuantaSoft™ software and can also be visualized as exportable charts in the *Concentration* and *Ratio* tabs or exported, if desired, as a “.csv” file for additional data analysis. After cluster identification in 2D plots, data may also be displayed in 1D Amplitude plots which is particularly useful for a side-by-side comparison showing the quality of cluster separation (appearing as “bands” of differing amplitudes in 1D plots) when the same assay is performed at different annealing temperatures (Fig. 1) or supplemented with “dark probes” (Fig. 2).



**Fig. 3** (a) Steps for the calculation of mutant allele abundance and allele ratios during ddPCR experiments using uniplex single-locus assays. Under “Setup,” select all unknown samples, positive controls, negative controls, and no-template controls (if available) analyzed with the same assay. Then, in the 2D Amplitude view under *Analyze*, use the rectangular Multi-well thresholding selection tool to initially enclose and thereby designate all droplets as “double-negative” in the 2D plot (panel 1, where all clusters will appear as grey). To properly identify and define each of the positive and negative clusters, first hover your mouse over the wild-type negative control(s) in the *well-editor* which will highlight their clusters in the 2D plot. Then, after clicking on the green “-/+” button, use any of the Multi-well thresholding selection tools to enclose droplets harboring wild-type alleles (WT) within the selected shape (rectangle, circle, or freehand shape), which will cause them to be associated with HEX/VIC fluorescence (panel 2). Similarly, after clicking on the blue “+/-” button, hover your mouse over the mutant positive control(s) in the *well-editor* to help define the cluster of droplets harboring the mutant allele (MUT), which after they are enclosed within the chosen shape will be associated with FAM fluorescence (panel 3). (b) Noninvasive monitoring of somatic mutations in the plasma of cancer patients using uniplex single-locus ddPCR assays. These assays have the potential to simultaneously quantify, with high precision and sensitivity, somatic mutations (MUT) and wild-type (WT) alleles from liquid biopsies. One patient showed a significant decrease in circulating tumour DNA levels during the course of therapeutic treatment (left panel). A different patient (right panel) was refractory to the treatment and ctDNA levels did not decrease over time. VAF stands for the “variant allele frequency” of rare alleles. For the calculation of allele fractions, “positive droplets” carrying mutant alleles are assigned to the FAM channel and “low positive droplets” carrying wild-type alleles are assigned to the HEX channel in 2D Amplitude plots, as depicted in Panel (a). 1D Amplitude plots, as shown here, are nonetheless useful to compare, side-by-side, the absolute number of mutant molecules reported across different samples. Please note that the fractional abundance of each allele can be more informative in cases where different volumes of sample have been used to perform assay (see the example of the nonresponder patient)



**Fig. 4** “Inverted” ddPCR assays enable the uncovering of multiple genetic aberrations affecting a mutation hotspot by using one single hydrolysis probe matching the sequence of the wild-type allele. WT stands for droplets containing wild-type alleles of the target of interest. In the example shown below, as many as four different single nucleotide polymorphisms (MUT1 to MUT4) can be detected and quantified with one single hydrolysis probe. For this experiment, we used four different DNA samples extracted from fresh tumor biopsies. Each sample was previously known to carry a different rare allele within the same mutation hotspot. Fluorescence amplitude signals associated with each particular mutant allele may slightly vary (see for instance MUT1 versus MUT2). This finding implies that optimum annealing temperatures during thermocycling may slightly differ depending on the specific mutant allele carried by the sample. For the calculation of allele fractions “positive droplets” carrying wild-type alleles are assigned to the HEX channel and “low positive droplets” carrying mutant alleles are assigned to the FAM channel; 2D plots not shown in this figure)

### 3.6 Nondiscriminating, Uniplex “Single-Locus” Screening Reactions (One Pair of Primers and One Wild-Type-Specific Hydrolysis Probe)

Hydrolysis probes with the wild-type sequence of a given locus can be used to detect *multiple* small genetic variations such as single nucleotide polymorphisms at a hotspot or even single or dual base indels. The hydrolysis of imperfectly annealed wild-type probes generates a secondary fluorescence amplitude band that is distinguishable from the lower fluorescence amplitude band emitted by “double-negative” droplets (i.e., droplets lacking copies of the locus of interest). This approach is well suited to screening for hotspot mutations characterized by multiple genetic abnormalities when the only information required from the assay is whether a certain hotspot is mutated or not (Fig. 4). These “inverted” ddPCR assays [27] are ideal to screen for germline mutations associated with known genetic disorders as well as for recurrent somatic mutations with allele frequencies above 10%, though potentially as low as 1%, from fresh tumor biopsies (see Note 25).

### 3.6.1 Running and Analyzing the Assay

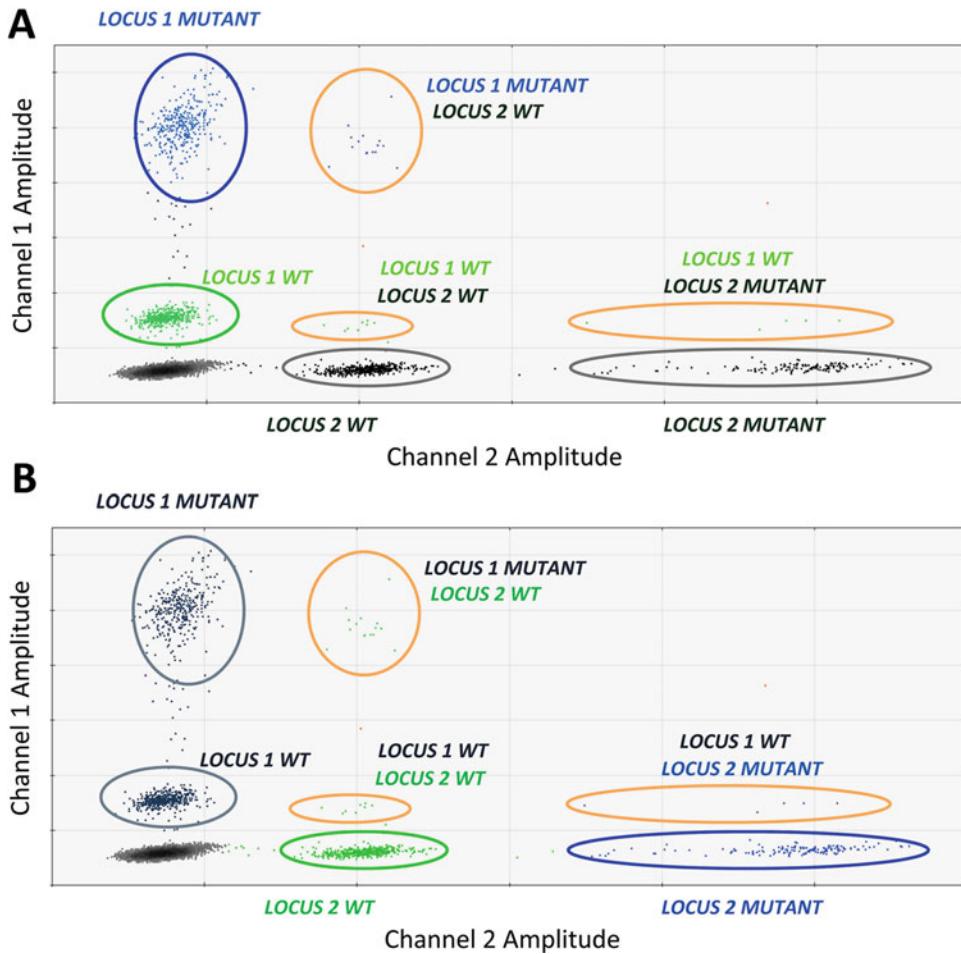
Volumes and suggested final concentrations of reagents for the preparation of nondiscriminating uniplex “single-locus” ddPCR assays are displayed in Table 1. Analogous to the procedure used for Uniplex assays with mutant-specific probe(s) in Subheading 3.5, quantification of mutant and wild-type alleles here is accomplished after manually setting thresholds in the QuantaSoft™ software. Select samples, positive controls, negative controls, and no-template DNA controls (if available) in the *Setup* mode. Ensure that the FAM optical channel has been associated to mutant alleles and the HEX/VIC optical channel has been associated to the wild-type allele in the *Setup* menu and then display the data on the *2D Amplitude* plot after clicking on *Analysis*. In essence, this approach assigns “low positive droplets” carrying mutant alleles to the FAM channel and “positive droplets” carrying wild-type alleles to the HEX/VIC channel. After clicking on the “−/−” black button, use the rectangular *Multi-well thresholding selection tool* to initially designate the entire set of droplets as “double-negative” droplets. Click on the green “−/+” button (HEX/VIC channel) to define the clusters of “positive” droplets for wild-type alleles (i.e., droplets emitting the most intense fluorescence signal). Click on the blue “+−” button (FAM channel) to define the clusters of “low positive” droplets (i.e., droplets containing rare alleles). The process for allele quantification relies in the same series of steps depicted in Fig. 3 (Panel a), with the exception that wild-type alleles linked to the HEX channel will be now represented in the cluster of highest fluorescence signal and mutant alleles linked to the FAM channel will be now represented in the cluster of intermediate fluorescence signal. Allele counts per well and estimated fractional abundances for each mutant allele will be given on the top table and can be visualized in the *Concentration* and *Ratio* charts and exported as “.csv” file.

## 3.7 Noncompeting Discriminatory Uniplex “Dual-Locus” Reactions (Two Primer Pairs and Two Rare Allele-Specific Hydrolysis Probes)

### 3.7.1 Running and Analyzing the Assay

The possibility to simultaneously quantify mutant and wild-type alleles with one single hydrolysis probe opens the possibility to use two probes conjugated with dyes of different emission spectra to interrogate two different loci in one single assay well (Fig. 5). The most direct benefit of such assay formats, beyond savings in overall costs, is that splitting precious and scarce samples to run independent assays for each individual locus can be avoided. Tracking more than one mutation has also important implications, for example, when monitoring tumor evolution during the course of therapeutic interventions in cancer patients [31, 32].

Volumes and suggested final concentrations of reagents for the preparation of discriminating “dual-locus” reactions are displayed in Table 1. The assay designs, optimization, and ddPCR workflow are carried out as described in Subheadings 3.2–3.4 above. Quantification of mutant and wild-type alleles is again accomplished after



**Fig. 5** Uniplex “Dual-locus” ddPCR reactions using two mutant only hydrolysis probes conjugated with dyes of different emission spectra to report the presence of rare alleles at two independent genomic locations. The efficient hydrolysis of mutant-specific probes generates high fluorescence amplitude either in the optical Channel 1 or Channel 2 for those droplets containing rare alleles. The hydrolysis of imperfectly annealed probes within droplets harboring wild-type alleles only (green circle) or together with mutant alleles (orange circles) allows the independent quantification of wild-type alleles. This assay was conducted on 5 ng (roughly) of a DNA sample extracted from fresh tumour biopsies previously known to harbor rare alleles for each one of the two interrogated loci. Calculation of allele copies per well and allele fractions at each independent locus, is done sequentially for the two loci where in each iteration, “positive droplets” carrying mutant alleles are assigned to the FAM channel (blue droplets) and “low positive droplets” carrying wild-type alleles (green droplets) are assigned to the HEX channel. This operation is carried out independently for each locus (illustrated in panels **a** and **b**, respectively)

manually setting thresholds in the QuantaSoft™ software. Select samples, positive controls, negative controls, and no-template DNA controls in the *Setup* or *Analyze* mode and display data on the *2D Amplitude* plot. However, in contrast to uniplex ddPCR assays, each locus in a dual-locus ddPCR assay interrogating two independent genomic positions needs to be analyzed

independently and sequentially, if using QuantaSoft™ v1.7.4.0917 or earlier versions, with the data being exported as graphs and/or as a “.csv” file after analyzing each genomic locus separately. (See Note 26 for an alternative analysis using newer software.)

Ensure that the two optical channels have been associated to their corresponding allele type (e.g., wild type or mutant) in the *Setup* menu and display the data on the *2D Amplitude* plot after clicking on *Analysis*. As with uniplex assays above, you will need to associate “positive droplets” carrying mutant alleles with the FAM channel and “low positive droplets” carrying wild-type alleles with the HEX/VIC channel, doing so iteratively for each individual locus. For the locus 1 target analysis, first set the contents of the entire plot as “−/−” double negative (i.e., grey) droplets. Use the *Multi-well thresholding selection tool* and click on the green “−/+” button (HEX/VIC channel) to define the clusters of “low positive” droplets carrying wild-type alleles of locus 1. Include here droplets with mixed content (i.e., locus 1 “low positive”/locus 2 “low positive” or locus 1 “low positive”/locus 2 “positive,” see Fig. 5). Click on the blue “+/-” button (FAM channel) to define the clusters of “positive” droplets for locus 1, also including the droplets of mixed content (i.e., locus 1 “positive”/locus 2 “low positive” or locus 1 “positive”/locus 2 “positive”). Allele counts and estimated allele frequencies for locus 1 mutations will be given on the top table and in the *Concentration* and *Ratio* charts.

After exporting locus 1 data, repeat the same set of operations to estimate the copies/well and fractional abundance of mutant and wild-type alleles of locus 2. You must redefine the new target ID associated with locus 2 in the *Setup Menu* unless generic “Mutant” or “Wild type” names are always associated with the FAM or HEX/VIC channel, respectively. Locus ID should be included in the name of the file exported from QuantaSoft™ in order not to confuse from which locus the exported “.csv” file data was derived.

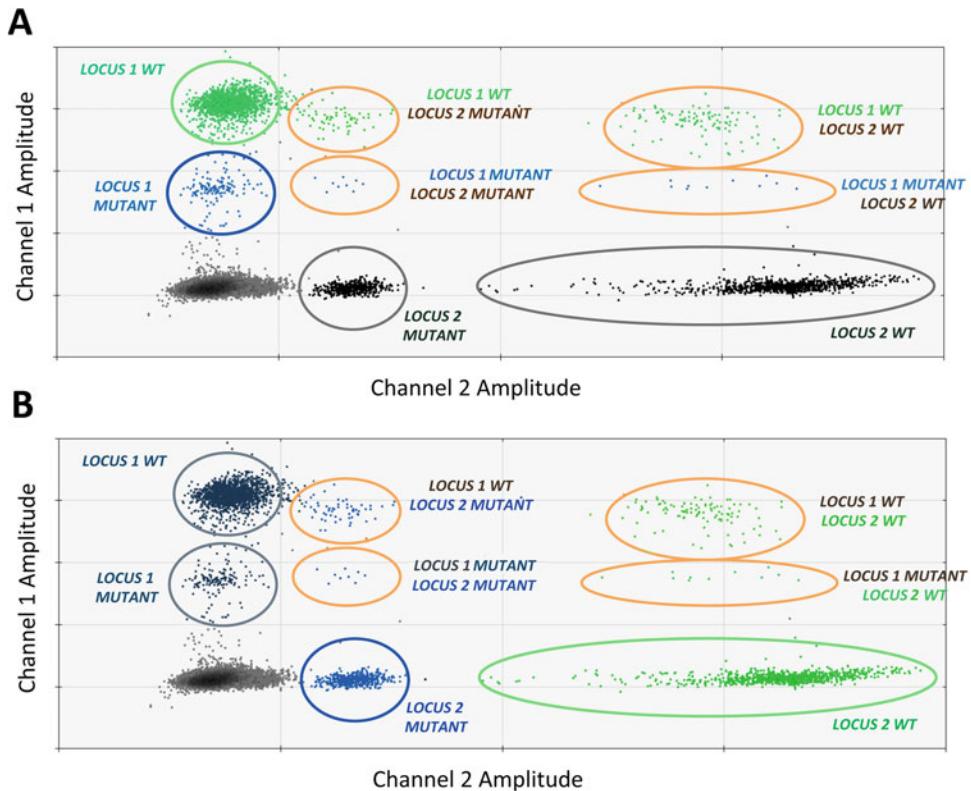
### **3.8 Inverted, Nondiscriminating Uniplex “Dual-Locus” Reactions (Two Primer Pairs and Two Wild- Type-Specific Hydrolysis Probes)**

#### **3.8.1 Running and Analyzing the Assay**

Two different mutation hotspots can be screened for single nucleotide genetic variants by using two wild-type-specific probes conjugated with dyes of different emission spectra. Hence, a substantial number of mutations can be investigated in one single assay (Fig. 6). Nondiscriminating “dual-locus” “inverted” ddPCR assays have the same limitations in terms of sensitivity and specificity as uniplex “inverted” ddPCR assays (see Note 25).

Volumes and suggested final concentrations of reagents for the preparation of nondiscriminating “dual-locus” reactions are displayed in Table 1. Assay designs, optimization, and workflow are as described in Subheadings 3.2–3.4.

Quantification of mutant and wild-type alleles is accomplished after manually setting thresholds in the QuantaSoft™ software.



**Fig. 6** Uniplex “Dual-locus” “inverted” ddPCR reactions with two hydrolysis probes matching only the wild-type alleles of two independent loci. These assays allow the simultaneous quantification of multiple mutant and the wild-type alleles using hydrolysis probes conjugated with two dyes (one per locus) detected in either the optical Channel 1 or Channel 2 in a single well. As many as nine different fluorescent clusters can be observed in the example shown in this figure. This assay was conducted on 10 ng (roughly) of a DNA sample extracted from fresh tumour biopsies with a priori known mutations for each one of the two mutation hotspots interrogated. For the calculation of allele copies and fractions at each independent locus, “positive droplets” carrying wild-type alleles (green droplets) are assigned to the HEX/VIC channel and “low positive droplets” carrying mutant alleles are assigned to the FAM channel (blue droplets). This operation is carried out independently for each locus (illustrated in panels **a** and **b**, for locus 1 and 2 respectively). Droplets containing copies of multiple alleles (for the same or different loci) are indicated by orange circles

Select samples, positive controls, negative controls, and no-template DNA controls in the *Setup* or *Analyze* mode and display data on the *2D Amplitude* plot. Nondiscriminating “dual-locus” ddPCR assays need to be analyzed for each locus independently and sequentially, when using QuantaSoft™ v1.7.4.0917 or earlier versions. For each sequentially analyzed locus, ensure that the two optical channels have been associated to its corresponding allele type (e.g., wild type or mutant) in the *Setup* menu and display the data on the *2D Amplitude* plot after clicking on *Analysis*. For this approach, assign “positive droplets” carrying wild-type alleles to the HEX channel and “low positive droplets”

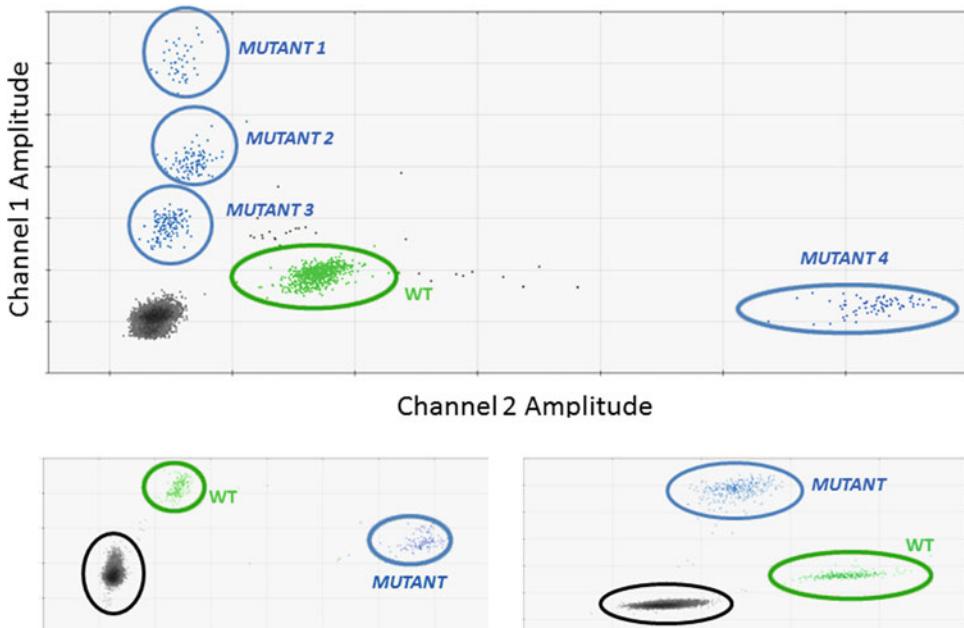
carrying mutant alleles to the FAM channel. First, however, set the contents of the entire plot as “−/−” double negative (i.e., grey) droplets. Use the *Multi-well thresholding selection tool* and click on the green “−/+” button (HEX channel) to define the clusters of “positive droplets” carrying wild-type alleles of locus 1. Include droplets with mixed content containing either wild-type or mutant alleles of locus 1 (see Fig. 6). Click on the blue “+−” button (FAM channel) to define the clusters of “low positive” droplets for locus 1 (i.e., droplets carrying rare alleles), also including the droplets of mixed content carrying mutant alleles for locus 1. Allele counts and estimated fractional abundances for locus 1 will be provided on the top table and in the *Concentration* and *Ratio* charts, as well as in the exportable “csv” file. Repeat the same set of operations to estimate the copies/well and fractional abundance of mutant and wild-type alleles pertaining to locus 2. You must redefine the new target ID associated with locus 2 in the *Setup Menu* unless generic “Mutant” or “Wild type” names are always associated with the FAM or HEX/VIC channel, respectively. Locus ID should be included in the name of the file exported from QuantaSoft™ in order not to confuse from which locus the exported “.csv” file data was derived. See Note 26 for an alternative analysis using newer software.

### **3.9 Higher Order, Single-Locus “Amplitude” Multiplexing (One Pair of Primers and Several Mutant Allele-Specific, But No Wild Type, Hydrolysis Probe)**

#### **3.9.1 Running and Analyzing the Assay**

Mutations within a given hotspot can be individually distinguished by using exclusively mutant-specific hydrolysis probes conjugated with one or the other of two dyes (here, FAM and HEX) applied at different final concentrations in the end-point PCR. Allele multiplexing in this way has advantages over conventional duplex rare mutation detection approaches due to the fact that wild-type-specific hydrolysis probes are not required. Yet as previously explained, these wild-type alleles can still be detected and quantified due to their imperfect hybridization to mutant probe(s) yielding droplets with an intermediate dual fluorescence signal (green clusters in Fig. 7).

Table 1 shows recommended starting concentration ratios for up to three mutant-specific probes conjugated with the same fluorescent dye (typically 6-FAM™). These recommendations are only suggestive and we recommend slightly altering the final concentration of each probe and conducting a temperature gradient until a satisfactory separation of each individual cluster is achieved. Quantification of mutant and wild-type alleles is accomplished after manually setting thresholds in the QuantaSoft™ software. First, select samples, positive controls, negative controls, and no-template DNA controls in the *Setup* view. Ensure that the two optical channels have been associated to their corresponding target types (i.e., mutant or wild type) in the *Setup* menu and then after clicking on *Analysis*, display the data on the *2D Amplitude* plot. All mutant alleles, regardless of the final probe concentration or dye used to



**Fig. 7** Multiplex detection of rare alleles using a single pair of primers and several mutant hydrolysis probes labeled with two different fluorophores, used at customizable final concentrations. In the upper panel, four mutant-specific probes were used together to interrogate a mixture of genomic DNA extracted from four independent tumor biopsies, each of them harboring a different mutation. Three hydrolysis probes were conjugated with 6-FAM™ and a fourth probe was labeled with HEX™. The three probes conjugated with 6-FAM™ can be associated with three well-defined fluorescent clusters (blue), according to the final concentration of each individual probe in the ddPCR reaction. The hydrolysis of the fourth probe labeled with HEX generates high fluorescence in Channel 2, but it is also assigned to FAM fluorescence for proper quantification purposes. The wild-type allele (green cluster) can be detected and independently quantified owing to the nonspecific hydrolysis of both 6-FAM™ and HEX™ probes within droplets exclusively containing wild-type alleles. When applied to a sample carrying one single mutation (bottom panels), these assays generate three different fluorescence clusters (one for “positive” droplets (blue clusters), one for “low positive,” intermediate droplets (green clusters), and one for “double-negative” droplets (grey clusters); samples negative for mutations (not shown) only exhibit two fluorescence clusters: “low positive,” intermediate droplets plus “double-negative” droplets)

target each particular allele, will be associated in the *Setup* menu with the FAM channel. Only wild-type alleles (i.e., those contained within droplets in the green cluster emitting a mixture of FAM and HEX fluorescence signals as seen in Fig. 7) will be associated with the HEX channel (in contrast to conventional QuantaSoft use where a dual-dye cluster would be labeled in orange droplets).

To define clusters, first, set the contents of the entire plot as “-/-” double negative (i.e., grey) droplets. Use the freehand lasso *Multi-well thresholding selection tool* and click on the blue “+/-” button (FAM channel) to define the cluster(s) of “positive” droplets carrying mutant alleles. (Typically, there will only be one such mutant cluster in a given sample; here, multiple samples each with a

different mutant allele were combined). Click on the green “−/+” button (HEX channel) to define the clusters of “low positive, intermediate” droplets carrying just wild-type alleles. Mutant and wild-type allele counts per well and estimated fractional abundance are given on the top table and in the *Concentration* and *Ratio charts*. Note that this assay strategy does not require iterative analysis as with the dual locus assay approaches above (Subheadings 3.7 and 3.8).

---

## 4 Notes

1. It is critical to not confuse this with the ddPCR™ Supermix for Probes (Bio-Rad, Hercules, CA, USA; Cat. No. # 186-3010, 186-3026, 186-3027, 186-3028), which contains dUTP. We have observed that the latter does not provide proper cluster resolution for the discriminating and nondiscriminating assays here described. Consistently using fresh ddPCR™ Supermix for Probes (no dUTP) is also critical for optimal assay resolution. The ddPCR™ Supermix for Probes (no dUTP) is stable, according to the manufacturer, for at least 6 months from the time of receipt when stored at −20 °C. Store at 4 °C for a maximum of 2 weeks after thawing. The manufacturer does not recommend multiple freeze/thaw cycles as this may affect assay resolution, or using reagents beyond their expiration date.
2. As with ddPCR™ Supermix for Probes (no dUTP), old or expired digital PCR droplet generation oil for probes can result in nonoptimal assay resolution. Always check the oil bottle for any signs of microbial proliferation or turbidity.
3. Spoiled or expired droplet reader oil can also generate suboptimal results.
4. Place the cartridges with the green sections facing to the right of the automated droplet generator. Ensure not to lift the DG32™ cartridges from their compartment while still in use. If lifted and placed back again, the automated droplet generator will consider it as a new cartridge and will reuse some of the sections. This will cause droplet generation errors and/or sample cross-contamination.
5. Empty waste bin when necessary.
6. Place the foil seal on top of the 96-well plate to be covered with the red line facing upward.
7. Ensure that the color of the precooled −20 °C plate cooler indicates the right temperature. These plates can maintain the temperature of samples at 0 °C for at least 1 h and then change their color when their temperature exceeds 7 °C.

8. Ensure that the levels and quality of the droplet generation oil for probes, pipette tips, and DG32™ cartridges are ready and well positioned to start the run. Be sure to use Droplet Generation Oil for Probes and, if necessary, replace any Automated Droplet Generation Oil for EvaGreen (suited for a different application) that might have been left over in the Auto DG from previous uses. It is critical to set up droplet digital PCR reactions in a final volume of 22 µL (as opposed to the 20 µL recommended in some of the manufacturer's protocols) to minimize the occurrence of errors during automated droplet generation. The later can result in low droplet counts and/or sample loss.
9. Inspect the levels of the droplet reader oil and waste bins before starting a run and replace if necessary. For example, a blinking green light indicates that the waste bin is about to be completely filled. A steady orange light indicates that the waste bin needs to be immediately replaced. If the instrument has not been used for more than 2 weeks it is recommended to run a prime-flush-prime cycle to remove old oil from the internal fluidic channels of the instrument.
10. Preheat the instrument during droplet generation and do not use it until the temperature has reached 180 °C. Remove the block from inside the instrument before the prewarming step starts.
11. Thermal cycling programs should preferably be preloaded in the instrument before starting a run. Other thermocyclers fitting the 96-well plates and specifications described above may also be compatible with this protocol, but should be tested for temperature uniformity and accuracy.
12. The methods described in this chapter have been validated with hydrolysis probes that were manufactured as described above. We recently evaluated the performance of custom ddPCR probe assays manufactured by Bio-Rad (Bio-Rad, Hercules, CA, USA; Cat. No. #1864011) and results were also satisfactory. The employ of other types of fluorescence dyes or dark quenchers has not been evaluated, nor the suitability of other fluorophore-labelled oligonucleotides such as Scorpion or Amplifluor™ primer-probes or hybridization probes such as molecular beacons.
13. When pipetting DNA extracts, it is important to use high-quality tips that do not shed microparticles. Microparticles are known to clog the microfluidic chambers of the DG32™ droplet generation cartridge and hamper droplet generation.
14. Four-base cutters and high fidelity restriction endonucleases insensitive to DNA methylation are preferred. Approximately 2–5 units of enzyme previously diluted in an appropriate

enzyme dilution buffer can often be directly added to the ddPCR reaction. Enzymatic fragmentation can be also carried out prior to ddPCR according to the manufacturer's protocol. Heat inactivation or postdigestion cleanup is not required in this case, but it is important to not heat up the sample above 65 °C and to ensure at least a tenfold dilution of the restriction enzyme buffer in the final ddPCR reaction since high salt content can inhibit the ddPCR reaction and should be avoided.

15. Uracil glycosylase treatment is recommended to reduce false positives if working with FFPE samples.
16. Some manufacturers provide useful tools for the design of hydrolysis probe-based assays, such as PrimerQuest® (Integrated DNA technologies, Coralville, IA, USA; <http://www.idtdna.com/Primerquest/Home/Index>), and OligoArchitect™ (Sigma-Aldrich, St. Louis, MO, USA; <https://www.sigmaaldrich.com/technical-documents/articles/biology/probe-design-services.html>) or Multiple Primer Analyzer (Thermo Fisher Scientific, Waltham, MA, USA). Primer3Plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/>), Oligo Calc (<http://biotools.nubic.northwestern.edu/OligoCalc.html>) and e-PCR (<https://www.ncbi.nlm.nih.gov/tools/ePCR/>) are also popular tools to evaluate oligonucleotide properties and assay specificity. It must be noted that some of these tools might be too stringent to nominate a candidate assay. A design that takes into account the vast majority (but perhaps not all) of the abovementioned recommendations could therefore arise as the only alternative in such situations.
17. Variable final concentrations of individual probes conjugated with the same fluorophore in the ddPCR reaction enable discrimination of several alleles, thus facilitating high order multiplexing.
18. There are no lower limits for the amount of input DNA used during a single ddPCR assay but this amount will determine both the precision and sensitivity of the assay. One may anticipate the expected number of copies for a given locus by considering the mass of the haploid human genome (~3.3 pg). Up to 150 ng of human genomic DNA can be added in a single well for a final concentration of ≈2000 target copies/µL but particular care must be paid to the accurate classification of a substantial number of droplets with mixed content. Assays more affected by droplet misclassification and background noise, such as inverted ddPCR assays or assays conducted over damaged FFPE samples, may benefit from lower amounts of input DNA per well and the use of multiple replicate wells to

increase the total number of genome equivalents analyzed per sample.

19. If using less than eight samples for a given run, do not leave any column wells empty or with a volume lower than 22  $\mu\text{L}$ . Fill such wells with 22  $\mu\text{L}$  of the 1  $\times$  ddPCR<sup>TM</sup> Supermix for probes (no dUTP) or 1  $\times$  ddPCR<sup>TM</sup> Buffer Control for Probes. The same applies for any multiple of 8 samples, i.e., do not leave any well empty for a given column selected for droplet generation as this will introduce air into the instrument's fluidics and immediately abort the process. Similarly, any air bubbles observed in DG8 sample wells should be dislodged and removed with a pipette tip to avoid their interfering with proper droplet formation.
20. Successful droplet generation can be visually inspected by the observation of a cloudy layer (droplets) on top of a transparent layer (oil). Failures during droplet generation translate into low numbers of droplets, small volumes in the collection plates (mostly oil) and usually result in sample loss.
21. Thermocycled emulsions of droplets can be kept at 4 °C for at least 24 h.
22. Our previous study targeting recurrent point mutations in B-cell lymphomas [27] showed a high concordance between uniplex and standard duplex ddPCR assays when applied, for example, to the same serial dilution of cell line genomic DNA. We also observed a high concordance between the allele frequencies inferred via uniplex ddPCR assays and those inferred by NGS methods like deep amplicon sequencing or targeted hybridization capture in fresh tumor biopsies. Uniplex ddPCR assays demonstrate exquisite sensitivity as well, being capable of detecting one single mutant DNA copy diluted in a background of 10,000 wild-type DNA molecules in one single well. Such extreme situations must nonetheless be confirmed by conducting assay replicates in both the interrogated sample and negative controls (see examples in ref. 27 and Chapter 3 by Tzonev, this volume). Uniplex discriminating ddPCR assays can therefore be used, for example, to screen for rare genetic variants associated with disease and hold promise for the non-invasive detection and quantification of rare genetic variants in cell-free DNA samples (Fig. 3). One of the downsides of these assays, however, relates to confidently identifying droplets containing both mutant and wild-type alleles. This phenomenon may lead to the underestimation of wild-type alleles at higher DNA input amounts due to the increased frequency of double-positive droplets which may not be distinguishable from mutant-only positive droplets. However, this should not affect the total count of mutant alleles per volume of sample. Serial

dilutions of cell line genomic DNA have also demonstrated a strong concordance between observed and expected allele frequencies during uniplex ddPCR assays, suggesting that the possible underestimation of wild-type alleles because of droplets of mixed content should not have a dramatic effect on the calculation of allele ratios when DNA input amounts are, for example, below 10 ng (equivalent to around 3333 human haploid genome equivalents per well). Another important limitation is that other rare but closely related alleles present in the sample might be confounded with wild-type alleles if there is a mismatch(es) between the mutant-specific probe and the DNA sequence of the second mutant allele.

23. It is strongly recommended to test optimum annealing temperatures for cluster differentiation on positive controls with a priori known allele ratios to determine if the allele counts associated with a certain fluorescence cluster are excessive (e.g., at lower temperatures) or deficient (e.g., at higher temperatures), possibly suggesting underly or overly stringent annealing conditions, respectively.
24. In some instances, it can be helpful to add a 3' spacer modification to the “dark probe” to avoid the emergence of additional fluoresce amplitude bands nearby the cluster of empty droplets. 3' C3 spacers (Integrated DNA technologies, Coralville, IA, USA) would represent an example of this type of chemical modifications and their main role is to avoid the extension of the probe by DNA polymerases.
25. The sensitivity and specificity of “inverted” ddPCR assays on samples with very low frequencies of mutant DNA, such as those expected in liquid biopsies, is compromised by the ddPCR “rain” effect (i.e., droplets emitting intermediate levels of fluorescence that are actually wild-type positive) and is therefore not recommended in situations where high sensitivity is needed. The quality of DNA across FFPE samples is also expected to be highly variable and associated with different levels of rain, thus precluding the use of “inverted” assays, for instance, in samples where DNA is too damaged or contains a significant amount of PCR inhibitors. Assay redesign, extended cycling regimes and DNA shearing have nonetheless been shown to lessen “rain” during ddPCR analyses [33] and the effect of these parameters on “inverted” ddPCR assays should be further explored. The analysis of positive, negative (wild type), and no-template DNA controls are also crucial here, as in Subheading 3.5.4, for setting up appropriate thresholds and manual clusters and for adequately identifying and classifying “rain” droplets. The concurrent presence of wild-type and mutant alleles within the same droplet may obscure the fluorescent amplitude signal associated with mutant alleles in these

assays, a feature that also limits the sensitivity and maximum useful DNA input per well of “inverted” ddPCR when compared to assays relying on mutant-specific probes.

26. Simultaneous quantification of more than one genomic locus (and >4 clusters at a time) can now be performed using the “Advanced User Options” of the new QuantaSoft™ Analysis Pro software, available from <http://www.bio-rad.com/en-us/sku/quantasoft-analysis-pro-v1-quantasoft-analysis-pro-software>. See Bio-Rad Bulletin #6827 for more details.

## Acknowledgments

This study was supported by the Canadian Institute for Health Research (CIHR) (New Investigator award and operating grant 300738), the Terry Fox Research Institute (projects #1043 and #1021), the Natural Sciences and Engineering Research Council of Canada (Research Tools and Instruments program EQPEQ 1501), and the BC Cancer Foundation.

## References

1. Everett TR, Chitty LS (2015) Cell-free fetal DNA: the new tool in fetal medicine. *Ultrasound Obstet Gynecol* 45:499–507
2. Liao GJW, Gronowski AM, Zhao Z (2014) Non-invasive prenatal testing using cell-free fetal DNA in maternal circulation. *Clin Chim Acta* 428:44–50
3. Bettegowda C et al (2014) Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med* 6:224ra24
4. Heitzer E, Ulz P, Geigl JB (2015) Circulating tumor DNA as a liquid biopsy for cancer. *Clin Chem* 61:112–123
5. Volik S, Alcaide M, Morin RD, Collins CC (2016) Cell-free DNA (cfDNA): clinical significance and utility in cancer shaped by emerging technologies. *Mol Cancer Res* 14(10):898. <https://doi.org/10.1158/1541-7786.MCR-16-0044>
6. Ramírez-Castillo FY et al (2015) Waterborne pathogens: detection methods and challenges. *Pathogens* 4:307–334
7. Whale AS et al (2016) Detection of rare drug resistance mutations by digital PCR in a human influenza A virus model system and clinical samples. *J Clin Microbiol* 54:392–400
8. Opota O, Jaton K, Greub G (2015) Microbial diagnosis of bloodstream infection: towards molecular diagnosis directly from blood. *Clin Microbiol Infect* 21:323–331
9. Prachayangprecha S et al (2014) Exploring the potential of next-generation sequencing in detection of respiratory viruses. *J Clin Microbiol* 52:3722–3730
10. Fu W et al (2015) A highly sensitive and specific method for the screening detection of genetically modified organisms based on digital PCR without pretreatment. *Sci Rep* 5:12715
11. Reid AL, Freeman JB, Millward M, Ziman M, Gray ES (2015) Detection of BRAF-V600E and V600K in melanoma circulating tumour cells by droplet digital PCR. *Clin Biochem* 48:999–1002
12. Mukaike M et al (2014) High-throughput and sensitive next-generation droplet digital PCR assay for the quantitation of the hepatitis C virus mutation at core amino acid 70. *J Virol Methods* 207:169–177
13. Kennedy SR et al (2014) Detecting ultralow-frequency mutations by Duplex Sequencing. *Nat Protoc* 9:2586–2606
14. Newman AM et al (2016) Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat Biotechnol* 34:547–555
15. Vogelstein B, Kinzler KW (1999) Digital PCR. *Proc Natl Acad Sci U S A* 96:9236–9241

16. Dong L et al (2015) Comparison of four digital PCR platforms for accurate quantification of DNA copy number of a certified plasmid DNA reference material. *Sci Rep* 5:13174
17. Dingle TC, Sedlak RH, Cook L, Jerome KR (2013) Tolerance of droplet-digital PCR versus real-time quantitative PCR to inhibitory substances. *Clin Chem* 59:1670–1672
18. Rački N, Dreо T, Gutierrez-Aguirre I, Blejec A, Ravnikar M (2014) Reverse transcriptase droplet digital PCR shows high resilience to PCR inhibitors from plant, soil and water samples. *Plant Methods* 10:42
19. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K (2002) A comprehensive review of genetic association studies. *Genet Med* 4:45–61
20. Chang MT et al (2016) Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity. *Nat Biotechnol* 34:155–163
21. Hawkey PM (2008) The growing burden of antimicrobial resistance. *J Antimicrob Chemother* 62(Suppl 1):i1–i9
22. Madic J et al (2016) Three-color crystal digital PCR. *Biomol Detect Quantif* 10:34. <https://doi.org/10.1016/j.bdq.2016.10.002>
23. Whale AS, Huggett JF, Tzanev S (2016) Fundamentals of multiplexing with digital PCR. *Biomol Detect Quantif* 10:15. <https://doi.org/10.1016/j.bdq.2016.05.002>
24. Bidshahri R et al (2016) Quantitative detection and resolution of BRAF V600 status in colorectal cancer using droplet digital PCR and a novel wildtype negative assay. *J Mol Diagn* 18:190–204
25. Findlay SD, Vincent KM, Berman JR, Postovit L-M (2016) A digital PCR-based method for efficient and highly specific screening of genome edited cells. *PLoS One* 11:e0153901
26. Taly V et al (2013) Multiplex picodroplet digital PCR to detect KRAS mutations in circulating DNA from the plasma of colorectal cancer patients. *Clin Chem* 59:1722–1731
27. Alcaide M et al (2016) Multiplex droplet digital PCR quantification of recurrent somatic mutations in diffuse large B-cell and follicular lymphoma. *Clin Chem* 62(9):1238–1247. <https://doi.org/10.1373/clinchem.2016.255315>
28. Sikora A et al (2010) Detection of increased amounts of cell-free fetal DNA with short PCR amplicons. *Clin Chem* 56:136–138
29. Andersen RF, Spindler K-LG, Brandlund I, Jakobsen A, Pallisgaard N (2015) Improved sensitivity of circulating tumor DNA measurement using short PCR amplicons. *Clin Chim Acta* 439:97–101
30. Miyaoka Y et al (2016) Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. *Sci Rep* 6:23549
31. Pantel K, Diaz LAJ, Polyak K (2013) Tracking tumor resistance using ‘liquid biopsies’. *Nat Med* 19:676–677
32. Barber LJ, Davies MN, Gerlinger M (2015) Dissecting cancer evolution at the macro-heterogeneity and micro-heterogeneity scale. *Curr Opin Genet Dev* 30:1–6
33. Lievens A, Jacchia S, Kagkli D, Savini C, Querci M (2016) Measuring digital PCR quality: performance parameters and their optimization. *PLoS One* 11:e0153317



# Chapter 17

## Identification and Use of Personalized Genomic Markers for Monitoring Circulating Tumor DNA

Yilun Chen, Anthony M. George, Eleonor Olsson, and Lao H. Saal

### Abstract

Digital PCR techniques are ideally suited for accurately quantifying trace amounts of target DNA sequences, such as tumor-derived mutant DNA that is present in the blood circulation of patients with cancer. Here, we describe an approach marrying low-coverage whole-genome sequencing of tumor tissues, to enumerate chromosomal rearrangement breakpoints, together with droplet digital PCR (ddPCR)-based personalized rearrangement assays to cost-effectively monitor circulating tumor DNA levels at multiple time-points during the clinical course. The method is generally applicable to essentially any cancer patient, as all cancers harbor unstable genomes, and may have uses for measuring minimal residual disease, response to therapy, and early detection of metastasis.

**Key words** Cell-free circulating tumor DNA, Personalized medicine, Liquid biopsy, Noninvasive diagnosis, Whole-genome sequencing, Droplet digital PCR

---

### 1 Introduction

Nuclear and mitochondrial DNA originating from normal and diseased cells can enter the blood stream by several processes and be found in plasma or serum as short (<200 bp) fragments, termed cell-free circulating DNA (cfDNA), at concentrations ranging from 5 to 20 ng/ml plasma in healthy persons to >1000 ng/ml plasma in advanced cancer patients. Although tumor-derived DNA (circulating tumor DNA; ctDNA) was first described decades ago, due to technological limitations, ctDNA has only recently emerged as a powerful biomarker in cancer research and clinical oncology [1–3]. The half-life of ctDNA can be measured in minutes to hours; therefore the quantity of ctDNA measured at any given time can be used as a snapshot of the cancer burden. In several cancer types, the levels of ctDNA has been shown to be associated with tumor progression [4–7] and risk of metastasis and death [8, 9]. Furthermore, quantification of ctDNA may be used to

monitor response to therapy, identify resistance mutations, and as a noninvasive companion diagnostic.

A number of features can be used to distinguish ctDNA from normal wild-type DNA in cfDNA, the most specific being somatic mutations and genome rearrangements [5, 10]. The fraction of cfDNA that originates from cancer cells can vary from less than 1% in early stage cancer to  $\gg 50\%$  in advanced metastatic disease. Irrespective of the ctDNA fraction, any given mutated or rearranged fragment may be present at exceedingly low abundance due to tumor heterogeneity/subclonality (e.g., mutant allele fractions  $<0.01\%$ ) [4, 11]. Moreover, the volume of blood plasma available for analysis may be limiting factor. For these reasons, highly sensitive and specific quantitative methods are required for accurate ctDNA analysis. Because most detection approaches use polymerase enzymes with inherent base misincorporation error rates that can introduce a single nucleotide sequence variant of interest by chance, genome rearrangements such as chromosomal rearrangements are currently the most exquisitely specific choice of ctDNA biomarker. Additionally, chromosomal rearrangements are excellent tumor-tracking biomarkers for breast cancer because many are formed early in tumorigenesis during a period of telomere crisis and breakage-fusion-bridge cycling in the originating neoplastic clone, and therefore these rearrangements are “trunk” markers that remain present in the majority of subsequent subclones [12–14].

Here, we describe an approach marrying low-coverage whole-genome sequencing of tumor tissues, to enumerate chromosomal rearrangement breakpoints, together with droplet digital PCR (ddPCR)-based personalized rearrangement assays to cost-effectively monitor circulating tumor DNA levels at multiple time-points [8].

---

## 2 Materials

### **2.1 Tissue Procurement and Preparation Components**

1. Tumor tissue (fresh-frozen or frozen after RNAlater preservation).
2. RNAlater (Ambion).
3. Disposable sterile scalpels.
4. Plastic petri dishes.
5. Normal tissue sample, e.g., peripheral blood lymphocytes.

### **2.2 Tumor and Normal DNA Isolation Components**

1. AllPrep DNA/RNA Mini kit (Qiagen).
2. QIAshredder columns (Qiagen).
3. 2-mercaptoethanol.
4. DX-Antifoaming reagent (Qiagen).

5. QIAcube instrument (Qiagen).
6. QIAcube Rotor adapters (Qiagen).
7. QIAcube Reagent bottles (Qiagen).
8. Wizard Genomic DNA Purification kit (Promega) or QIAamp DNA Blood Mini Kit (Qiagen).

### **2.3 Sequencing Library Preparation Components**

1. Covaris tubes (Covaris).
2. Covaris S220 Focused ultrasonicator (Covaris).
3. 2100 Bioanalyzer instrument (Agilent).
4. High Sensitivity DNA Analysis Kit (Agilent).
5. Illumina TruSeq DNA Sample Preparation Kit (Illumina).
6. Agencourt AMPure XP Beads (Beckman Coulter).
7. Freshly Prepared 80% ethanol in water.
8. PCR grade water.
9. Thermal cycler.
10. Magnetic stand (Life Technologies).
11. 50× Tris–acetate–EDTA Buffer (Prepare 1× TAE buffer [40 mM Tris, 20 mM acetic acid, 1 mM EDTA, pH 8.0] with distilled water before use).
12. 100 bp DNA Ladder (Life Technologies).
13. 50 bp DNA Ladder (Life Technologies).
14. Disposable sterile scalpels.
15. Ultrapure Agarose 1000 (Life Technologies).
16. 6× DNA Gel Loading Dye (Thermo Fisher Scientific).
17. MinElute Gel Extraction Kit (Qiagen).
18. Benchtop vortex mixer.
19. SyBr Gold Nucleic Acid Gel Stain (Life Technologies).
20. Qubit Fluorometer (Thermo Fisher Scientific).
21. Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific).
22. HiSeq 2000 or 2500 sequencing instrument (Illumina).
23. cBot instrument (Illumina).
24. TruSeq DNA Sample Preparation Kit (Illumina).
25. TruSeq Cluster Kit cBot HS (Illumina).
26. PE Flow Cells (Illumina).
27. TruSeq SBS Kit HS (Illumina).

### **2.4 Cell-Free DNA Collection and Extraction Components**

1. Purple K2EDTA Vacutainer (Becton Dickinson), Cell-Free DNA BCT (Streck), or other suitable blood collection tube.
2. Disposable Pasteur pipettes.
3. DPBS without Mg<sup>2+</sup>/Ca<sup>2+</sup>.

4. QIAamp UltraSens Virus Kit (Qiagen).
5. Benchtop vortex mixer.
6. One variable-speed thermomixer at 60 °C, 1.5 ml.
7. One variable-speed thermomixer at 40 °C, 2.0 ml.
8. Refrigerated centrifuge at 4 °C.
9. Centrifuge at room temperature.
10. DNA LoBind tubes, 1.5 and 2.0 ml (Eppendorf).

## **2.5 Circulating Tumor DNA Detection Components**

1. Phusion Master Mix (Thermo Fisher Scientific).
2. Caliper LabChip XT System (PerkinElmer).
3. QX100/200 Droplet Digital PCR system (Bio-Rad).

---

## **3 Methods**

### **3.1 Tissue Procurement and Preparation**

The methods for tissue handling and preparation can vary. Primary tumor tissue may be obtained from a biopsy or surgical specimen. For biopsies, fine needle aspirates may not always yield sufficient tissue and therefore core biopsy or larger is recommended. Highest quality nucleic acids are obtained from well-handled tissues with minimal ischemic time prior to fresh-freezing to at least –80 °C or preserved in RNAlater prior to freezing. We recommend obtaining paired normal tissue (*see Note 1*), such as peripheral blood lymphocytes from the buffy coat fraction of a blood collection tube. Tissue preserved by formalin fixation and paraffin embedding (FFPE) can be used, but will result in lower-yield and quality of DNA, as well as increased potential for sequencing artifacts (typically C > T base substitutions).

1. Complete details of our routine tumor tissue preparation is described elsewhere [15]. In brief, we typically use fresh-frozen (to at least –80 °C) tumor tissue or tumor specimens preserved in RNAlater prior to freezing. Using a disposable scalpel and petri dish, the tumor sample is partitioned into three parts, one 10–30 mg piece for isolation of nucleic acids, one ~10 mg piece for construction of tissue macroarray, and if remaining, the remainder piece is saved for future research. Samples are frozen at –80 °C until use.
2. For normal tissue sample, we typically use peripheral blood lymphocytes. The buffy coat fraction can be obtained following centrifugation of a standard EDTA blood collection tube at  $2000 \times g$  for 10 min at 4 °C, transfer of the top plasma fraction to separate 1.5 ml aliquots using a sterile Pasteur pipette, followed by transfer of the buffy coat fraction to a separate tube (*see Note 2*). The remaining red blood cell fraction may be discarded. Store fractions at –80 °C.

### **3.2 Tumor and Normal DNA Isolation**

Any standard method for purification of genomic DNA may be used. As described elsewhere [8, 15], we prefer the AllPrep DNA/RNA Kit (Qiagen) for simultaneous purification of genomic DNA and total RNA from tumor tissues, and the Wizard Genomic DNA Purification Kit (Promega) or QIAamp DNA Blood Mini Kit (Qiagen) for isolation of normal blood DNA from peripheral blood lymphocytes. Standard manufacturer protocols should be followed, and the isolated nucleic acids stored frozen at –80 °C until use.

### **3.3 Sample Preparation for Sequencing**

The sample library preparation for sequencing is performed using the TruSeq DNA Sample Preparation Kit (Illumina) according to the TruSeq Sample preparation guide (Part # 15005180 Rev. A) using the low-throughput (LT) protocol (*see Note 3*). We follow the standard protocol, with the following changes as detailed in Subheadings [3.3.1–3.3.5](#) below.

#### *3.3.1 Fragmentation of DNA*

1. Add 2.4 µg per sample of genomic DNA to individual Covaris tubes in a total volume of 120 µl.
2. Fragment the genomic DNA using the S220 Focused Ultrasonic instrument (Covaris). For improved physical sequence coverage, we shear the genomic DNA to ~700 bp average length using the following settings: duty cycle 5%, intensity 3, cycles per burst 200, time 30 s at 5 °C. Samples may be stored at –20 °C.
3. Analyze the distribution of fragment lengths after shearing on a 2100 Bioanalyzer instrument (Agilent) using the High Sensitivity DNA Analysis Kit (Agilent) before continuing with library preparation. Successfully sheared genomic DNA will have a size distribution from 300 to 2000 bp with peak at 900 to 1000 bp and a concentration >500 pg/µl.
4. End repair, purification, adenylation, ligation of adaptors, and purification after ligation are performed following standard Illumina TruSeq protocol (Part # 15005180 Rev. A). Libraries may be stored at –20 °C after either purification step.

#### *3.3.2 Size Separation After Ligation*

1. Ligated sequencing libraries are size separated by electrophoresis using a 2% agarose gel made with Ultrapure Agarose 1000 (Life Technologies) and 1 × Tris-acetate-EDTA (TAE) buffer. Heat gel mixture in microwave, let cool for 5 min, and add 1 µl Sybr Gold Nucleic Acid Gel Stain (Life Technologies) per 10 ml gel and swirl to mix before pouring; let gel set.
2. Prepare samples, 20 µl after purification, by adding 4 µl of 6× DNA Gel Loading Dye (Thermo Fisher Scientific).
3. Prepare ladder: for each desired lane (typically two or three lanes, the first and last and one in the center), mix 3 µl of 1:10 diluted 50 or 100 bp DNA Ladder (Life Technologies), 17 µl

resuspension buffer RSB (TruSeq kit; Illumina), and 4 µl of 6× DNA Gel Loading Dye.

4. Load 24 µl of each sample and ladder mix with at least one empty lane between each sample/ladder lanes.
5. Run gel at 120 V constant voltage for 120–145 min.
6. View the gel on a UV transilluminator. Using a fresh scalpel blade for each sample, excise a band from the gel spanning the width of the lane and ranging in size from 550 to 950 bp and place in a tube (*see Note 4*). Use the DNA ladders as a guide. Minimize time under transillumination to avoid DNA damage. Determine weight of each gel piece by measuring tube without and with gel piece. Gel pieces may be stored at –20 °C.

### **3.3.3 Purification After Size Separation**

1. Follow the protocol for the MinElute Gel Extraction Kit (Qiagen) to purify each sample. Incubate the gel slices in the QG solution at room temperature (a change from the Qiagen protocol, as recommended by Illumina) until the gel slices have completely dissolved (10–20 min), while vortexing every 2 min.
2. Elute the samples in 25 µl Buffer EB (if more than one column is needed per sample, the eluted DNA should be pooled after purification).
3. Store the samples at –20 °C or continue with PCR enrichment.

### **3.3.4 PCR Enrichment of DNA Fragments**

1. Thaw the PCR Master Mix and PCR Primer Cocktail at room temperature. Once thawed, keep the tubes on ice.
2. Prepare PCR reactions using 12 µl of DNA (~1 µg) per sample, 8 µl of H<sub>2</sub>O (PCR grade), 5 µl of PCR Primer Cocktail, and 25 µl of PCR Master Mix (total volume is 50 µl per reaction).
3. Use the following program for PCR enrichment (with heated lid):
 

98 °C for 30 s  
 12 cycles of: (modification of Illumina protocol)  
     98 °C for 60 s (modification of Illumina protocol)  
     60 °C for 30 s  
     72 °C for 30 s  
     72 °C for 5 min  
     Hold at 4 °C
4. Purification after PCR is performed using AMPure XP Beads (Beckman Coulter) following the Illumina protocol (*see Note 5*).
5. Analyze the distribution of post-PCR fragment lengths by diluting 1 µl library in 49 µl water and running on a 2100

Bioanalyzer instrument (Agilent) using the High Sensitivity DNA Analysis Kit (Agilent) and the concentration should be measured using the Qubit Fluorometer (Thermo Fisher Scientific). Successful libraries will have a fragment peak at approximately 800–900 bp and an undiluted concentration >10 ng/μl (typically 30–50 ng/μl).

### 3.3.5 Cluster Generation and Sequencing

Sequencing clusters can be generated on a cBot instrument (Illumina) using TruSeq Cluster Kit cBot HS (Illumina) and PE Flow Cells (Illumina).

Paired-end sequencing of 2 × 50 bp, 2 × 100 bp, or 2 × 150 bp plus index read can be performed on a HiSeq 2000 or HiSeq 2500 sequencer using TruSeq SBS Kit HS. For enumeration of chromosomal rearrangements using our SplitSeq bioinformatics pipeline, >60 million read-pairs (2 × 50 bp or 2 × 100 bp) to >9× physical coverage (>2× sequence coverage) is generally sufficient.

## 3.4 Sequencing Bioinformatics

Bioinformatics steps are described in detail elsewhere [8]. In short:

1. The paired-end reads are aligned to a reference human genome, e.g., GRCh37, using Novoalign (Novocraft Technologies) with soft-clipped read alignments (option –o Softclip) (*see Note 6*).
2. Potential chromosomal aberrations are first identified with high sensitivity using BreakDancer [16] with default options for discordant read-pairs.
3. The rearrangement-supporting discordant read pairs are realigned to the reference genome using Novoalign with the 1000 top-scoring alignments above a moderate alignment score being reported (options –r Exhaustive 1000 –t 250). Initially discordant read pairs that become concordant after this step should be excluded. The purpose of this step is to reduce the false-positive rate due to misalignment of paralogous sequences.
4. DNA copy number is determined across the genome in windows of 50 kb using FREEC version 5.6 with default parameters [17]. DNA copy number gains at putative rearrangements may be utilized to prioritize candidates.
5. Identified rearrangement breakpoint ends are annotated with RefSeq genes, sequence gaps (gaps track), and repetitive elements (RepeatMasker track) for the human reference genome (hg19), all obtained from the UCSC Table Browser (<http://genome.ucsc.edu/cgi-bin/hgTables>), and with entries from the Database of Genomic Variants (<http://dgv.tcag.ca/>).
6. To deplete the list of predicted chromosomal rearrangements for potential nonspecific rearrangements and false-positives

before experimental validation, the following filtering criteria are applied:

- (a) At least two discordant read-pairs supporting the rearrangement.
- (b) No satellite DNA (RepeatMasker class “Satellite”) present within 1 kb of any of the two breakpoint ends of the rearrangement.
- (c) No sequence gap (UCSC track “gaps”) present within 1 kb of any of the two breakpoint ends.
- (d) No matching rearrangement in other tumor samples within 1 kb (both breakpoint ends matched).
- (e) No matching rearrangement in other normal samples within 1 kb (both breakpoint ends matched).
- (f) For intrachromosomal rearrangements, size of the rearrangement (distance between the two breakpoint ends) greater than 1 kb.
- (g) Both breakpoint ends on chromosomes 1–22 or X (i.e., no involvement of nonstandard sequence contigs present in the human genome reference).

7. SplitSeq software was developed to reconstruct the exact sequence of breakpoints from low-coverage whole-genome sequencing data. The software and code will be described in a forthcoming manuscript. Briefly, SplitSeq uses discordant read-pairs as anchors and searches the regions around the breakpoints for split-reads with soft-clipped bases that do not match the reference (soft-clipping in Novoalign can occur either due to mismatches or due to low base qualities at any end of a read). The same regions are also searched for reads with unmapped mates (i.e., read-pairs where only one read of the pair is aligned), and alignment of the unmapped mates to the breakpoint regions is attempted. Alignment is done with a reduced gap extension penalty (parameter -x 1) starting with the full read length, followed by two rounds of trimming the read to 2/3 of its previous length on either end. Any mapped split-read identified with this procedure is added to the list of split reads. Next, all soft-clipped reads from this list are sorted by their clipping position in the reference genome in order to identify putative exact breakpoint positions. Reference genome positions with support of less than a total of  $X$  clipped bases (typically 6) may be discarded. For all other positions, starting with those with the highest number of clipped-base support, all pairs connecting the two breakpoints (defined by the discordant read-pair anchors) are searched for perfect matches of all clipped bases when aligned to a fusion sequence reconstructed from the two breakpoint positions.

### **3.5 Cell-Free Plasma DNA Extraction**

#### *3.5.1 Fractionation of the Whole Blood*

1. Fill blood collection tube.
2. Centrifuge the blood samples at  $2000 \times g$  for 10 min at  $4^\circ\text{C}$  (*see Notes 2 and 7*).
3. Transfer the upper plasma fraction with a Pasteur pipette into 1.5 ml aliquots, and process immediately or store at  $-80^\circ\text{C}$ .

### **3.6 Preparation for Isolation of cfDNA**

1. Before starting the cell-free DNA extraction protocol, prepare the following reagents and devices. The reagents are all from the Qiagen UltraSens Virus Kit unless otherwise specified.
  - (a) Add 310  $\mu\text{l}$  Buffer AVE to one tube of lyophilized carrier RNA to obtain a 1  $\mu\text{g}/\mu\text{l}$  solution. Use 5.6  $\mu\text{l}$  of this solution per extraction. Store the carrier RNA aliquots at  $-20^\circ\text{C}$ .
  - (b) Add 96–100% ethanol to the Buffer AB, AW1, and AW2 according to the manufacturer’s instructions.
  - (c) In heating block, bring Buffer AR to  $60^\circ\text{C}$  (330  $\mu\text{l}$  is required per sample).
  - (d) If working with frozen stored plasma, thaw on ice.

### **3.7 Isolation of cfDNA**

Cell-free DNA extraction is performed using the UltraSens Virus Kit protocol with modifications:

1. Centrifuge the plasma samples at  $10,000 \times g$  for 10 min at  $4^\circ\text{C}$ .
2. Transfer the supernatant to 2 ml DNA LoBind tubes, and store any excess aliquots at  $-80^\circ\text{C}$  (*see Note 8*).
3. If isolating cfDNA from less than 1 ml plasma, use DPBS (without  $\text{Mg}^{2+}/\text{Ca}^{2+}$ ) to bring the volume of each plasma sample to 1 ml (*see Note 9*).
4. Pipet 5.6  $\mu\text{l}$  carrier RNA (prepared in Subheading 3.6 above) into the dry lid of each sample tube (*see Note 10*).
5. Add 800  $\mu\text{l}$  Buffer AC directly to each sample.
6. Close the lid, invert tube five times and mix thoroughly by vortexing for 10 s.
7. Incubate the tubes at room temperature for 15 min to allow protein-nucleic acid complexes to form. Regular mixing by inverting the tube several times every 5 min or less during the period of incubation is suggested.
8. Centrifuge the samples at  $1200 \times g$  for 3 min at room temperature (*see Note 11*).
9. Remove and discard all but ~50  $\mu\text{l}$  of the supernatant from the samples (*see Note 12*).
10. Flick the tubes forcefully to loosen the pellets.
11. Add 20  $\mu\text{l}$  Proteinase K to the lid of each tube (*see Note 13*).

12. Add 300  $\mu$ l of preheated Buffer AR to each tube and invert and flick the tube directly after adding the buffer. Do not use the vortex mixer here.
13. Incubate the tubes for 10 min at 40 °C in a thermomixer at 400 revolutions per minute or until all pellets are completely dissolved. A brief centrifuge is helpful to remove drops from the lid.
14. Add 400  $\mu$ l Buffer AB, mix thoroughly with vortex mixer, and centrifuge briefly.
15. Carefully apply 700  $\mu$ l lysate to a QIAamp spin column, centrifuge at  $2000 \times g$  for 3 min, and discard the flow-through (*see Note 14*).
16. Place the QIAamp spin column in a new 2 ml collection tube, add 500  $\mu$ l Buffer AW1, and centrifuge at  $6000 \times g$  at room temperature for 1 min.
17. Place the QIAamp spin column in a new 2 ml collection tube, add 500  $\mu$ l Buffer AW2, and centrifuge at the maximum speed (13,000–16,000  $\times g$ ) at room temperature for 3 min.
18. Place the QIAamp spin column in a new 2 ml collection tube, and centrifuge at the maximum speed for 1 min to completely dry the membrane (*see Note 15*).
19. Elute cfDNA using two rounds of Buffer EB. Place the QIAamp spin column in a clean 1.5 ml DNA LoBind tube, apply 50  $\mu$ l elution buffer directly to the membrane, incubate 1 min at room temperature, and centrifuge at  $6000 \times g$  at room temperature for 1 min. Repeat this step and combine the eluate to obtain ~100  $\mu$ l cfDNA solution (*see Note 16*).

### **3.8 Detection of Tumor-Specific Genomic Rearrangements in cfDNA**

#### **3.8.1 Rearrangement Selection**

1. Select a subset of the predicted fusion sequences for assay design. It is advisable to inspect the breakpoint positions with a genome browser such as IGV. For example, inspect individual reads that lay across the breakpoint (split-reads). We typically select 10 rearrangements per tumor, from which ~5 will be analyzed in cfDNA samples per patient.

### **3.9 ddPCR Assay Design**

1. Design primers and probes. The forward and reverse primers should be on opposite sides of the breakpoint location with a hydrolysis probe positioned between them. The primers should be positioned as close to the probe as possible, but the primers and probe should not overlap each other or the breakpoint position (*see Note 17*). The amplicons should be designed to be as short as possible due to the highly fragmented nature of cfDNA.

2. The  $T_m$  and sequence parameters for the primers and probes should adhere to the Applied Biosystems Primer Express Software guidelines on “Quantification TaqMan MGB Probe Design Guidelines and Considerations” and “Primer Design Guidelines and Considerations” as well as the guidelines from the Bio-Rad “Droplet Digital PCR Applications Guide”.
  - (a) Primer Guidelines:
    - $T_m$  between 50 and 65 °C, preferably between 58 and 60 °C and should have similar  $T_m$  (e.g., within 1–2 °C of each other).
    - GC content between 30 and 80%, preferably between 50 and 60%.
    - Avoid polynucleotide repeats, in particular avoid repeats of four or more consecutive Gs or Cs.
    - The terminal 3' base should be a G or C if possible, the final three bases should not all be Gs or Cs, and the final five nucleotides should include no more than three G and/or C bases.
    - The primers should be positioned as close to the probe as possible without overlapping.
    - Avoid hairpin loops, self-dimerization, and cross-dimerization.
  - (b) Probe Guidelines:
    - $T_m$  of the probe should 68–70 °C, and greater than 3 °C higher than the  $T_m$  of the primers, preferably 8–12 °C higher.
    - Probe length should be 13–30 bases.
    - GC content between 30 and 80%.
    - The terminal 5' end should not be a G, and when using a FAM dye-labeled probe, the second base from the 5' should also not be a G.
    - Avoid polynucleotide repeats, in particular avoid repeats of four or more consecutive Gs or Cs. Also avoid repeats of six or more consecutive As.
    - Nonfluorescent quenchers, such as Black Hole Quencher, are recommended.
    - Consider selecting probes with more Cs than Gs to minimize fluorescence quenching.
    - Avoid hairpin loops, self-dimerization, and cross-dimerization.
3. Label the hydrolysis probe with fluorescence group of FAM, HEX, or VIC and we recommended quenching using an internal ZEN and 3'-IBHQ molecule (Integrated DNA Technologies).

### **3.10 Rearrangement and Primer Validation**

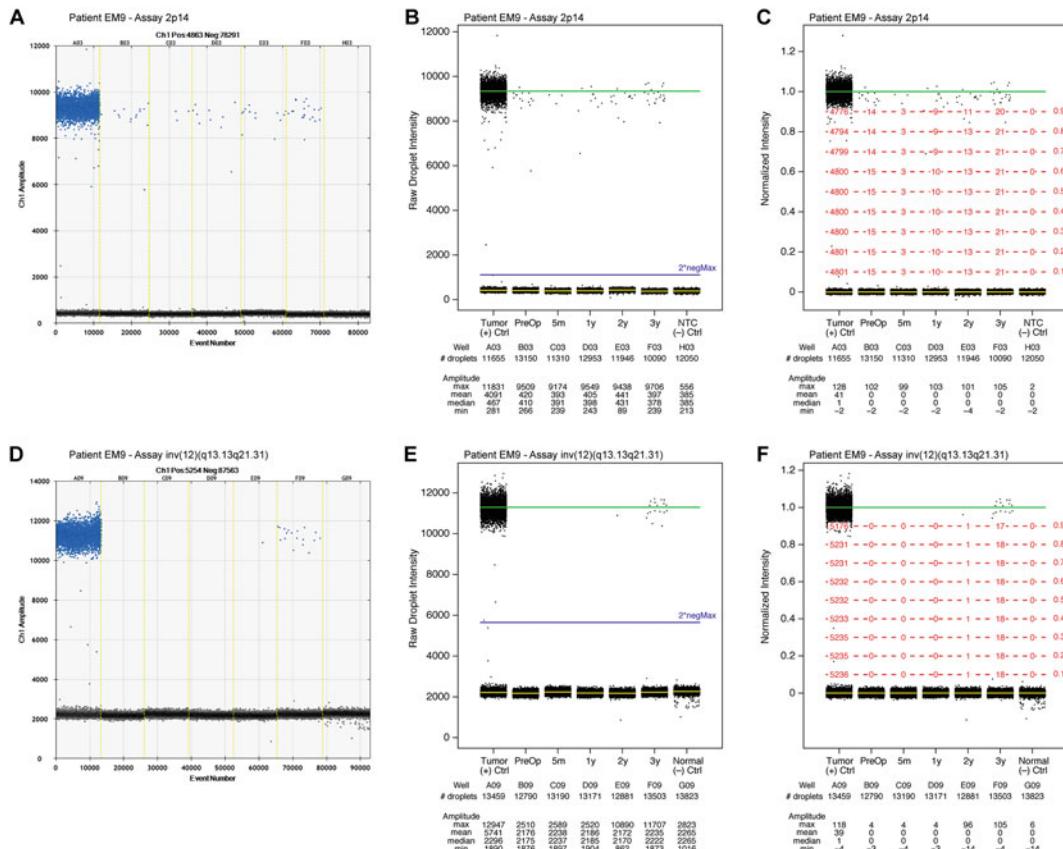
1. Order the primer pairs (but not the hydrolysis probes) and perform standard PCR using tumor DNA as positive control template to validate the structural variants of interest, and using matched normal DNA to distinguish tumor-specific somatic rearrangements from germline variants. Any standard PCR set-up may be employed here; typically we perform:
  - (a) PCR Reactions (each 10 µl total volume):
 

Phusion Master Mix 1× (Thermo Scientific)  
250 nM of each primer  
2% DMSO  
10 ng of template DNA
  - (b) PCR Cycling Conditions:  
Initial denaturation: 98 °C for 2 min  
11 cycles of:  
    98 °C for 10 s  
    70 °C (−1 °C/cycle) for 30 s (70–60 °C in 1 °C intervals)  
    72 °C for 15 s  
29 cycles of:  
    98 °C for 10 s  
    60 °C for 30 s  
    72 °C for 15 s  
Final elongation at 72 °C for 5 min
2. Analyze the PCR products to validate the fragment size and the germline or somatic status of the alterations, e.g., using the Caliper LabChip XT System (PerkinElmer) (*see Note 18*).
3. For the structural variants that were confirmed to be somatic by touchdown PCR, order the TaqMan probesv for the corresponding assays.

### **3.11 ddPCR Assay Validation**

1. Validate the complete assay (probe and primer pair) in droplet digital PCR using the QX100/200 Droplet Digital PCR System (Bio-Rad) with tumor DNA as the template for positive control and matched normal DNA as the template for negative control. In addition to negative control DNA template, a no-template control (NTC) consisting of PCR-grade water can be run. The following ddPCR setup was optimized for this protocol:
 

ddPCR Reactions (each 20 µl total volume) (*see Notes 19–21*):  
1 × digital PCR supermix for probes (Bio-Rad)



**Fig. 1** Example personalized genomic marker analysis of circulating DNA at multiple time points during clinical follow-up. Panels **a–c** illustrate ddPCR results for the 2p14 wild-type assay for breast cancer patient EM9 and panels **d–f** show results for the inv(12)(q13.13q21.31) assay for the same patient. Wells are ordered, from left to right: tumor positive control (+ Ctrl), blood plasma prior to surgery (PreOp), plasma at 5-months (5m), 1 year (1y), 2 years (2y), and 3 years (3y) after surgery, and negative control (− Ctrl). QuantaSoft 2D droplet plots are shown in **a** and **d** with wells separated by vertical yellow lines. Panels **b** and **e**, and **c** and **f** illustrate the transformation steps for assay-specific data normalization and thresholding on the exported intensity measurements (see Subheading 3.12). In **b** and **e**, the blue line “2<sup>negMax</sup>” is drawn for a given ddPCR assay at two times the maximum droplet intensity of the negative control (no-template control, NTC; or matched normal DNA, Normal) reaction well, defining the data from the given assay into an upper and lower portion. Yellow horizontal lines are drawn at the median value for the droplets in the lower portion within each well (“lowerMed”), and the green line is drawn across at the median value for the upper portion of the positive control primary tumor DNA well (“posUpperMed”). Data are then scaled to 0 and 1 (**c**, **f**) corresponding to these lowerMed and posUpperMed values, respectively. Red dashed lines are drawn for nine example threshold values, from 0.1 to 0.9, and numbers given within the lines indicate the count of droplets above the threshold for the respective well. Summary statistics for each well is provided below each plot. Figure adapted from Olsson et al. [8]

250 nM of each primer

250 nM of probe

10 ng of template DNA

2. Mix the ddPCR reactions well before transferring 20  $\mu$ l to the wells of the sample row (middle row) of a DG8 cartridge (the DG8 cartridge should already be loaded into the DG8 cartridge holder before transfer) (*see Note 22*). Empty sample wells should be filled with 1  $\times$  ddPCR buffer control. Do not mix the reactions in the DG8 cartridge.
3. Add 70  $\mu$ l of droplet generation oil into the wells of the oil row (bottom row). Always add the sample before the oil, as the opposite can result in fewer generated droplets.
4. Put the rubber gasket into place on the DG8 cartridge holder and then carefully load it into and start the QX100/200 Droplet Generator. The two liquids combined in the microfluidic channels and an emulsion is produced in the droplet row (top row) of the cartridge.
5. Slowly and carefully (to avoid shearing of the droplets) pipette 40  $\mu$ l from the droplets row to a 96-well PCR plate (Eppendorf twin.tec PCR plates 96, semiskirted). Consider using an electronic multichannel pipette adjusted to the slowest aspirate and dispense settings.
6. Place the 96-well PCR plate in a thermal cycler (e.g., Bio-Rad C1000 Touch Thermal Cycler) and run the ddPCR cycling protocol as specified here:

ddPCR Cycling Conditions (*see Note 23*):

Initial denaturation/enzyme activation: 95 °C for 10 min

10 touchdown cycles of:

94 °C for 30 s

65 °C for 60 s (adjusted –0.7 °C/cycle)

35 cycles of:

94 °C for 30 s

58 °C for 60 s

Enzyme deactivation: 98 °C for 10 min.

Hold at 10 °C.

7. Transfer the 96-well PCR plate to the Bio-Rad (QX100/200) Droplet Reader and set the QuantaSoft (Bio-Rad) software to ABS (absolute quantification) mode and to interrogate the appropriate fluorescent signals (FAM, HEX, or VIC) (*see Note 24*).
8. For ddPCR assays that validate successfully, ~5 assays can be used to analyze circulating DNA isolated from patient blood plasma (Fig. 1). Follow the same ddPCR setup and instructions as above (*see Note 25*). For the template input volume and concentration, consider the number of assays you wish to run, the number of replicates, and, without exceeding 100,000

haploid genome copies/well (<330 ng/well), we recommend using the maximum amount of template per ddPCR reaction that can be afforded per reaction well (typically, 8  $\mu$ l of reaction volume is available for the template).

### 3.12 Droplet Digital PCR Intensity Thresholding

1. Determine a threshold for each assay based on the intensity values from positive and negative control samples (Fig. 1). The threshold is a fixed amplitude value that reliably separates true-positive droplets from true-negative droplets. Even with a good quality assay, however, some droplet intensities will fall somewhere between the positive and negative clouds. Because of this, careful consideration should be taken when defining exactly where to place the threshold. If the threshold is set too low, the frequency of potential false-positives is higher whereas alternatively if the threshold is set too high, the frequency of potential false-negatives increases. When detecting a target molecule that occurs as an extremely rare event, as can be the case with circulating tumor DNA, it is critical that both false-positives and false-negatives are avoided.
2. Various methods can be used to define where to draw the threshold. One method is to export the droplet intensity values from QuantaSoft in order to be able to apply automatic, unbiased, reproducible, and operator independent thresholding that is based on normalized intensity signals from positive and negative controls (Fig. 1). Normalizing the data facilitates employing one intensity threshold for all samples in order to define negative droplets (below threshold) and positive droplets (above threshold). Using centering and scaling, this procedure normalized each individual assay's droplet fluorescent intensity data to a new scale where 0 corresponds to no intensity (at the median intensity of negative control droplets) and 1 corresponds to full intensity (at the median intensity of positive control droplets) and was carried out as described below. For each rearrangement assay:
  - (a) Determine  $negMax$  = maximum intensity value in the negative control well.
  - (b) For each ddPCR reaction well, define a lower group of droplets with intensity  $<=2 * negMax$  and an upper group of droplets with intensity  $>2 * negMax$ .
  - (c) For each well, determine  $lowerMedian$  = median intensity value of the lower group of droplets.
  - (d) For each well, subtract  $lowerMedian$  from each intensity value (bringing all medians to zero).
  - (e) Determine  $posUpperMedian$  = median intensity value in the upper group of droplets in the positive control well.

- (f) For each well, divide each intensity value by *posUpperMedian* (setting the median of the upper droplets in the positive control to 1).
  - (g) Set threshold *t* to 0.5 to define negative droplets (below threshold) and positive droplets (greater than or equal to threshold).
3. We typically measure the concentration of a normal nonrearranged 132 bp region of chromosome 2p14 as a surrogate for total cfDNA. This 2p14 region very infrequently undergoes copy number change in breast cancer. The assay sequences are, 2p14-Fwd GCTGAATTGCTTGAGCTTCAG; 2p14-Rev CATGTCACCTCCAGTTACAGGGAA; 2p14-Probe TGTCACCAGTTGTCTGCTGTTCTTCGGG.
  4. The number of fragments per  $\mu\text{l}$  input purified circulating DNA ( $C_{V_i}$ ) is calculated from the number of positive droplets  $P$ , total number of droplets analyzed  $T$ , droplet volume  $V_d$  ( $0.91 \times 10^{-3} \mu\text{l}$  or  $0.85 \times 10^{-3} \mu\text{l}$  depending on hardware/software versions; we recommend checking with the manufacturer, Bio-Rad, if unsure), ddPCR reaction volume  $V_r$  ( $20 \mu\text{l}$ ; includes PCR mix, primers, probe, input DNA), and volume of purified circulating DNA input into the reaction  $V_i$ , using the formula: 
$$C_{V_i} = \frac{-\ln(1 - \frac{P}{T})}{V_d} \left( \frac{V_r}{V_i} \right).$$
  5. Quantification may be expressed as copies/ml plasma by multiplying by a conversion factor accounting for starting plasma volume and the cfDNA elution volume. Tumor-specific rearrangements may also be expressed as percent ctDNA by dividing by the counts of the normal control fragment of 2p14.

#### 4 Notes

1. In order to filter germline events, both tumor and normal DNA from the same patient should be analyzed. In the event that matched normal DNA sample is unavailable, unmatched normal DNAs may be used to filter the most common germline events in the population.
2. Fractionate blood as soon as possible after collection. For standard EDTA tubes, within 2 h is ideal. Streck Cell-Free BCT tubes may be kept at room temperature or  $4^\circ\text{C}$  for up to 14 days (refer to manufacturer instructions).
3. If more than 48 samples will be prepared for sequencing, the high-throughput (HT) protocol should be used.
4. Cutting a band between 550 and 950 bp will result in an insert size of approximately 450–850 bp, accounting for the size of

the adapters. Adapters add approximately 120 bp to each fragment.

5. The AMPure XP beads must be acclimated to room temperature before use (at least 30 min).
6. Data from samples run on multiple lanes of a flow cell should be merged into one BAM file using Novosort (Novocraft Technologies). Duplicate read-pairs in the BAM files can be flagged with Picard MarkDuplicates and ignored in subsequent analyses.
7. Whole blood contains approximately 50% plasma. Plasma is the preferred source of cfDNA since, during the formation of serum, normal cells lyse and release their genomic DNA into solution, thereby increasing background and obscuring true cfDNA quantities.
8. Free circulating DNA in blood plasma is sensitive to multiple freeze and thaw cycles; therefore aliquot your plasma samples.
9. Optionally, a volume of exogenous DNA reference fragments, with the size between 100 and 200 nucleotides and at known concentration, can be added to the sample prior to compensation with DPBS. Reference exogenous DNA fragments may be used to calculate the extraction efficiency.
10. It is important that the carrier RNA does not mix with Buffer AC or with the sample until the two have been mixed together as this can affect the yield.
11. A pellet should be at the bottom of each tube containing the cfDNA and the supernatant should be mostly clear. Centrifuge force and time may be adjusted, but prolonged and/or a high centrifuge force can compact the pellet and make resuspension difficult.
12. Leaving a small volume of supernatant greatly aids resuspension.
13. We recommend keeping Proteinase K solution at 4 °C.
14. Excessive centrifuge force will reduce binding of nucleic acids to the membrane. If the total volume of the lysate is larger than 700 µl, repeat this step with the same column until all lysate has been loaded.
15. Residual Buffer AW2 often remains on the plastic “lip” above the membrane. This may be avoided by removing excess solution around the lip using a pipette prior to this second centrifugation.
16. The elution buffer and volumes may be changed depending on the downstream application, however the total yield may decrease if the elution volumes are small.

17. There are two generations of Droplet Digital PCR systems from Bio-Rad. The QX100 Droplet Digital PCR system supports hydrolysis probe assays (TaqMan) while the QX200 Droplet Digital PCR system supports both hydrolysis probe and DNA binding dye (EvaGreen) assays. This protocol is optimized to use hydrolysis probe assay method.
18. Analysis of the PCR products can be done with a number of instruments including the 2100 Bioanalyzer (Agilent) or by standard gel electrophoresis. The Caliper system is ideal for high throughput.
19. Prepare reactions at a slightly higher volume than 20  $\mu$ l, i.e., 22  $\mu$ l. This is to ensure that a full 20  $\mu$ l is available for droplet generation.
20. For the positive control reactions, the amount of template tumor DNA used can vary but should be less than 66 ng. For reference, 10 ng of human genomic DNA (which is roughly 3000 haploid genome copies) will result in thousands of positive reaction droplets if the assay is working properly, so much less tumor template can be used if desired.
21. For the negative control reactions, it is advisable to run a large amount of nonspecific DNA template (e.g., >20 ng matched normal DNA). Since circulating tumor DNA can exist at extremely low concentrations in extracted plasma, it is important to have very specific assays since the number of true-positive droplets can be low. Ideally no false-positive droplets should occur even after analyzing very many thousands of normal genomes as template. An advantage of structural variants as tumor biomarkers, as compared to other types of variations such as point mutations, is that they result in very unique sequences, and thus very specific assays should be attainable.
22. Plan ahead the ddPCR cartridge and plate layout. For a given assay, it is recommended to place test samples first, followed by the positive and then negative controls (with regards to the order in which the droplets are generated and the plate is read).
23. If possible, the ramp rate between any two consecutive steps should be set to ~2.5 °C/s to ensure reliable thermal control.
24. The droplet reader measures the fluorescence intensity, size, and shape of each droplet, excludes droplets that do not pass quality metrics, and then outputs the intensities of the passing droplets for each of the measured dyes. For an assay to be considered validated, the following criteria should be met:
  - (a) In the positive control wells, there should be two clusters (appear as clouds or bands in 1D plot) of droplets that are clearly separate from one another. The cluster with the

higher amplitude represents droplets that contained the target template and the cluster with the lower amplitude represents droplets that contained no target template. There two bands should be relatively compact (tight around their average intensity).

- (b) In the negative control wells, there should be a single cluster that has the same average intensity as the lower amplitude cluster of the positive control reactions for the same assay. The negative control cluster should be a tight band with no stray droplets in the positive amplitude direction. If this is observed, the assay should be excluded as it could result in false-positives.
25. In breast cancer, metastases typically retain 89% (range 61–100%) of the chromosomal rearrangements present in their matched primary tumor [14]; using the lower 61% figure, the odds are 99/100 that one or more rearrangements out of five will be retained in a metastatic clone and thus be informative for minimal residual disease, tracking response to therapy, or for early detection of occult metastatic disease.

## Acknowledgments

We thank members of the Translational Oncogenomics Unit, Division of Oncology and Pathology for assistance, and in particular to Christof Winter and Robert Rigo for bioinformatics work. This work was supported by the Swedish Cancer Society, Swedish Research Council, Swedish Foundation for Strategic Research, Knut and Alice Wallenberg Foundation, VINNOVA, and Governmental Funding of Clinical Research within National Health Service, Swedish Breast Cancer Group, Crafoord Foundation, Lund University Medical Faculty, Gunnar Nilsson Cancer Foundation, Skåne University Hospital Foundation, BioCARE Research Program, King Gustav Vth Jubilee Foundation, Kräpparup Foundation, and the Mrs. Berta Kamprad Foundation.

## References

1. Haber DA, Velculescu VE (2014) Blood-based analyses of cancer: circulating tumor cells and circulating tumor DNA. *Cancer Discov* 4:650–661
2. Ignatiadis M, Dawson SJ (2014) Circulating tumor cells and circulating tumor DNA for precision medicine: dream or reality? *Ann Oncol* 25:2304–2313
3. Schwarzenbach H, Hoon DS, Pantel K (2011) Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer* 11:426–437
4. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, Bartlett BR, Wang H, Luber B, Alani RM, Antonarakis ES, Azad NS, Bardelli A, Brem H, Cameron JL, Lee CC, Fecher LA, Gallia GL, Gibbs P, Le D, Giuntoli RL, Goggins M, Hogarty MD, Holdhoff M, Hong SM, Jiao Y, Juhl HH, Kim JJ, Siravegna G, Laheru DA, Lauricella C, Lim M, Lipson EJ, Marie SK, Netto GJ, Oliner KS, Olivi A, Olsson L, Riggiens GJ, Sartore-Bianchi A, Schmidt K, Shih

- I M, Oba-Shinjo SM, Siena S, Theodorescu D, Tie J, Harkins TT, Veronese S, Wang TL, Weingart JD, Wolfgang CL, Wood LD, Xing D, Hruban RH, Wu J, Allen PJ, Schmidt CM, Choti MA, Velculescu VE, Kinzler KW, Vogelstein B, Papadopoulos N, Diaz LA Jr (2014) Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med* 6:224ra224
5. Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, Chin SF, Dunning MJ, Gale D, Forshaw T, Mahler-Araujo B, Rajan S, Humphray S, Becq J, Halsall D, Wallis M, Bentley D, Caldas C, Rosenfeld N (2013) Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N Engl J Med* 368:1199–1209
  6. Jung K, Fleischhacker M, Rabien A (2010) Cell-free DNA in the blood as a solid tumor biomarker—a critical appraisal of the literature. *Clin Chim Acta* 411:1611–1624
  7. Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, Thornton K, Agrawal N, Sokoll L, Szabo SA, Kinzler KW, Vogelstein B, Diaz LA Jr (2008) Circulating mutant DNA to assess tumor dynamics. *Nat Med* 14:985–990
  8. Olsson E, Winter C, George A, Chen Y, Howlin J, Tang M-HE, Dahlgren M, Schulz R, Grabau D, van Westen D, Ferno M, Ingvar C, Rose C, Bendahl PO, Ryden L, Borg A, Gruvberger-Saal SK, Jernstrom H, Saal LH (2015) Serial monitoring of circulating tumor DNA in patients with primary breast cancer for detection of occult metastatic disease. *EMBO Mol Med* 7:1034–1047
  9. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ, Cheang M, Osin P, Nerurkar A, Kozarewa I, Garrido JA, Dowsett M, Reis-Filho JS, Smith IE, Turner NC (2015) Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Sci Transl Med* 7:302ra133
  10. Leary RJ, Kinde I, Diehl F, Schmidt K, Clouser C, Duncan C, Antipova A, Lee C, McKernan K, De La Vega FM, Kinzler KW, Vogelstein B, Diaz LA Jr, Velculescu VE (2010) Development of personalized tumor biomarkers using massively parallel sequencing. *Sci Transl Med* 2:20ra14
  11. Diehl F, Li M, Dressman D, He Y, Shen D, Szabo S, Diaz LA Jr, Goodman SN, David KA, Juhl H, Kinzler KW, Vogelstein B (2005) Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proc Natl Acad Sci U S A* 102:16368–16373
  12. Chin K, de Solorzano CO, Knowles D, Jones A, Chou W, Rodriguez EG, Kuo WL, Ljung BM, Chew K, Myambo K, Miranda M, Krig S, Garbe J, Stampfer M, Yaswen P, Gray JW, Lockett SJ (2004) In situ analyses of genome instability in breast cancer. *Nat Genet* 36:984–988
  13. Alkner S, Tang MH, Brueffer C, Dahlgren M, Chen Y, Olsson E, Winter C, Baker S, Ehinger A, Ryden L, Saal LH, Ferno M, Gruvberger-Saal SK (2015) Contralateral breast cancer can represent a metastatic spread of the first primary tumor: determination of clonal relationship between contralateral breast cancers using next-generation whole genome sequencing. *Breast Cancer Res* 17:102
  14. Tang MH, Dahlgren M, Brueffer C, Tjitrowirjo T, Winter C, Chen Y, Olsson E, Wang K, Torngren T, Sjostrom M, Grabau D, Bendahl PO, Ryden L, Nimeus E, Saal LH, Borg A, Gruvberger-Saal SK (2015) Remarkable similarities of chromosomal rearrangements between primary human breast cancers and matched distant metastases as revealed by whole-genome sequencing. *Oncotarget* 6:37169–37184
  15. Saal LH, Vallon-Christersson J, Hakkinen J, Hegardt C, Grabau D, Winter C, Brueffer C, Tang MH, Reutersward C, Schulz R, Karlsson A, Ehinger A, Malina J, Manjer J, Malmberg M, Larsson C, Ryden L, Loman N, Borg A (2015) The Sweden Cancerome Analysis Network - Breast (SCAN-B) Initiative: a large-scale multicenter infrastructure towards implementation of breast cancer genomic analyses in the clinical routine. *Genome Med* 7:20
  16. Chen K, Wallis JW, McLellan MD, Larson DE, Kalicki JM, Pohl CS, McGrath SD, Wendl MC, Zhang Q, Locke DP, Shi X, Fulton RS, Ley TJ, Wilson RK, Ding L, Mardis ER (2009) Break-Dancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat Methods* 6:677–681
  17. Boeva V, Zinovyev A, Bleakley K, Vert JP, Janoueix-Lerosey I, Delattre O, Barillot E (2011) Control-free calling of copy number alterations in deep-sequencing data using GC-content normalization. *Bioinformatics* 27:268–269



# Chapter 18

## Single Color Multiplexed ddPCR Copy Number Measurements and Single Nucleotide Variant Genotyping

Christina M. Wood-Bouwens and Hanlee P. Ji

### Abstract

Droplet digital PCR (ddPCR) allows for accurate quantification of genetic events such as copy number variation and single nucleotide variants. Probe-based assays represent the current “gold-standard” for detection and quantification of these genetic events. Here, we introduce a cost-effective single color ddPCR assay that allows for single genome resolution quantification of copy number and single nucleotide variation.

**Key words** Digital PCR, Single color, Mutation detection, Single nucleotide variant (SNV), Copy number variation (CNV), Circulating cell free DNA (ccfDNA)

---

### 1 Introduction

This novel molecular assay utilizes BioRad’s QX200 Droplet Digital PCR (ddPCR) system to detect and quantify clinically actionable mutations within poor quality and limited quantity samples such as FFPE DNA and circulating tumor DNA (ctDNA). This new molecular assay technology has been optimized for quantitatively measuring clinically actionable mutations, is highly sensitive, and is capable of detecting the targeted mutant fraction at an individual DNA molecule resolution. DNA shed from tumors can be extracted from the plasma fraction of routine blood draws; the resulting ctDNA is extremely variable and low in concentration, typically yielding well below 5 ng DNA per  $\mu\text{L}$  of plasma [1], thus making the ctDNA a prime candidate for this assay. The ddPCR single-color system allows for accurate and absolute quantification of the mutant and wild-type allelic fractions within a given patient sample using easy to design standard PCR primers; this means that any cancer mutation can have a customized assay generated and tested efficiently.

Allele-specific mutation quantitation relies on two DNA primer sets that are identical with the exception of the mutant or wild type

specific base at the 3' end of the “detecting” primer sets, to amplify the genomic region of interest. Through the addition of artificial 5' noncomplementary tails to our mutant and wild type-specific “detection” primers, we are able to consistently differentiate between droplets that contain the mutant or wild type alleles based upon their differential amplicon lengths. The artificial tails minimize bias within the PCR reaction by allowing our primers to target an identical region of genomic DNA, with the exception of the SNP specific base at the 3' end of the detecting primer. The ddPCR technology allows the standard PCR reaction to be partitioned into approximately 20,000 individual emulsions. These can be effectively treated as 20,000 unique reactions that are cycled in parallel within a single well of a PCR plate. We can then assay each individual droplet to assess whether the individual DNA molecule partitioned into the droplet is “wild-type” or the target “mutant”—giving an absolute count of mutant and wild-type templates in a given patient sample. This assay is capable of detecting seven mutant genome equivalents among a total of 7000 genome equivalents using standard nonmodified DNA primers. The sensitivity of this assay also allows it to handle the low concentration samples typically gathered from the extraction of circulating cell free DNA from the plasma fraction of a standard blood draw.

---

## 2 Materials

Pre- and post-PCR workspaces should be in physically separate locations with designated equipment and consumables (\*Denotes an item that should be kept exclusively in post-PCR areas). All reagents/primers should be divided into smaller one-time-use volumes to reduce the risk of contamination.

### 2.1 Instruments/ Equipment

1. QX200™ Droplet Generator (Bio-Rad, 1864002).
2. QX200™ Droplet Reader\* (Bio-Rad, 1864003).
3. QuantaSoft™ Software (Bio-Rad).
4. Veriti® 96-well Thermal Cycler\* (Applied Biosystems, 4375786).
5. DG8™ Cartridge Holder (Bio-Rad, 1863051).
6. PX1™ Plate Sealer (Bio-Rad, 1814000).
7. AirClean 600 Combination PCR Workstation.
8. Centrifuge.
9. Tabletop centrifuge.
10. Promega Maxwell® RSC.

**2.2 Consumables**

1. QX200™ Droplet Generation Oil for EvaGreen (Bio-Rad, 1864005).
2. DG8™ Cartridges for QX200™/QX100™ Droplet Generator (Bio-Rad, 1863008).
3. DG8™ Gaskets for QX200™/QX100™ Droplet Generator (Bio-Rad, 1863009).
4. ddPCR™ Droplet Reader Oil (Bio-Rad, 1863004).
5. Eppendorf 96-Well twin tec PCR Plates; 250 µL semiskirted (Eppendorf, 951020389).
6. Pierceable Foil Heat Seal (Bio-Rad, 1814040).
7. Eppendorf DNA LoBind Microcentrifuge Tubes; Safe-Lock Tube 1.5 mL; DNA; PCR Clean; (Eppendorf, 22431021).
8. Rainin P20 Low-Retention Filter Tips (Rainin, RT-L10FLR).
9. Rainin P200 Low-Retention Filter Tips (Rainin, RT-L200FLR).
10. Rainin P1000 Low-Retention Filter Tips (Rainin, RT-L1000FLR).

**2.3 ddPCR Reagents**

1. QX200™ ddPCR™ EvaGreen Supermix (Bio-Rad, 1864034).
2. Qx200™ 2× Buffer Control for EvaGreen (Bio-Rad, 1864052).
3. UltraPure™ DNase/RNase-Free Distilled Water (Invitrogen, 10977-023).
4. Elution Buffer, 10 mM Tris-HCl pH 8.0: Prepare elution buffer solution by diluting UltraPure™ 1 M Tris-HCl, pH 8.0 (Invitrogen, 15568-025) to a final concentration of 10 mM with UltraPure™ Water (Invitrogen, 10977-023).
5. Restriction endonuclease (TBD by user, *see* Subheading 3.1).
6. Primers: Resuspend all primers in elution buffer to a final concentration of 100 µM; vortex primers thoroughly and briefly spin down. Dilute primer to a final working concentration of 10 µM in UltraPure™ Water and store resuspended primers at -20 °C).

**2.4 Sample Extraction Reagents**

1. Maxwell® RSC DNA FFPE Kit (Promega, AS1450).
2. EDTA Blood Collection Vials.
3. Maxwell® RSC ccfDNA Plasma Kit (Promega, AS1480).

---

**3 Methods**

All reagent preparation and droplet generation should be carried out in a PCR workstation to reduce the risk of contamination (*see Note 1*). Use Rainin Low Retention Filter-Tips when preparing all reagents and master mixes.

### **3.1 Sample DNA Extraction**

#### *3.1.1 Extraction of DNA from FFPE Tissue*

1. Obtain 5–10 µm sections of FFPE block containing tissue of interest.

#### *3.1.2 Extraction of ccfDNA from Blood*

1. Collect two EDTA vials of blood, keep at 4 °C until ready to process. **Note: Processing must take place within 2 h of collection.**
2. Spin blood containing vials for 10 min at 2000 × *g* in a centrifuge.
3. Transfer plasma fraction to a fresh 1.5 or 2.0 mL LoBind Eppendorf tube.
4. Spin the plasma containing tube for 10 min at 2000 × *g*.
5. Transfer supernatant a fresh 1.5 or 2.0 mL LoBind Eppendorf tube using care not to disturb the pellet.

*Proceed to step 6 for long-term storage or skip step 6 and proceed to step 7 to proceed and extract ccfDNA.*

6. Flash freeze the tube containing plasma in liquid nitrogen until frozen, store at –80 °C until ready to extract ccfDNA.
7. Proceed with the standard protocol included in the Maxwell® RSC ccfDNA Plasma Kit.
8. Store extracted ccfDNA at –20 °C for short term (1–2 days) or at –80 °C for long-term storage. Avoid excessive freeze–thaw cycles.

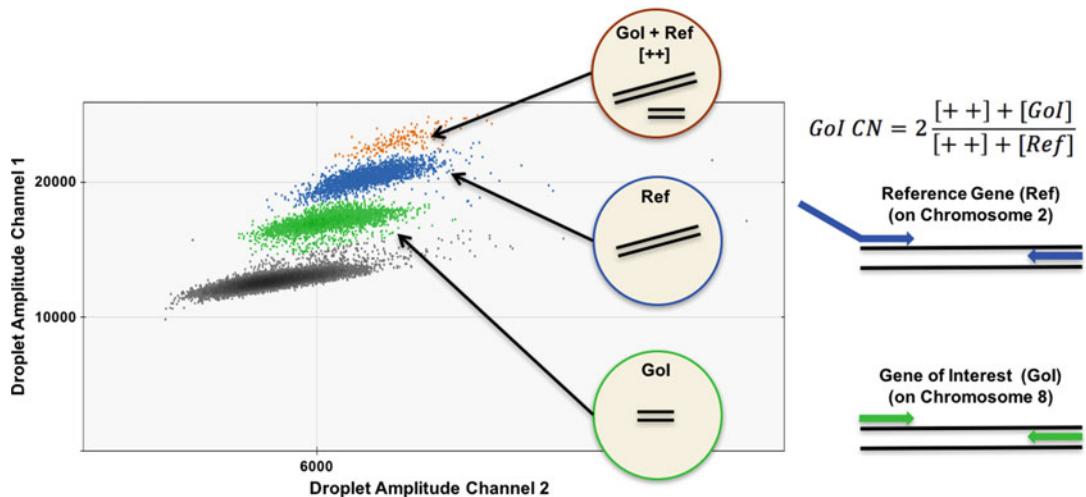
### **3.2 Designing SNV Primers**

#### *3.2.1 Copy Number Variation (CNV) Assay Primer Design*

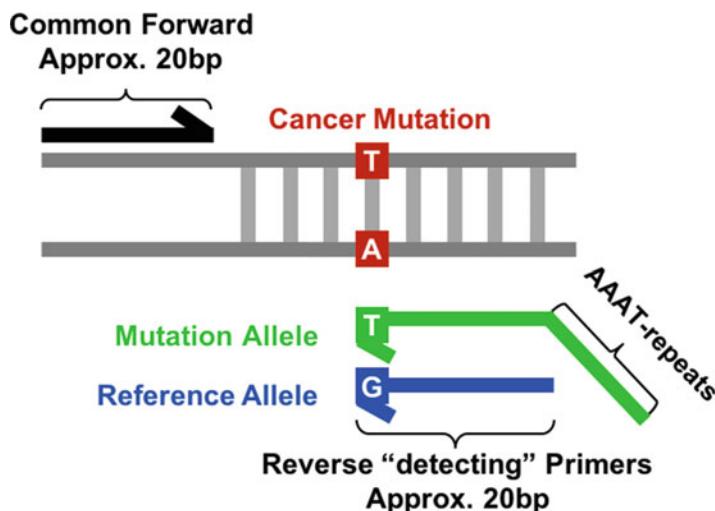
CNV assay primers should target approximately 20 bp of genomic sequence and may be designed using Primer3. The primers' specificity can be confirmed using the UCSC Genome Brower Blat tool [2]. CNV assay primers targeting the genomic region of interest should be multiplexed with an off target reference gene as a copy number control (multiplexing suggestions Subheading 3.2.3). To determine the copy number of the gene of interest (GoI), you will count the number of double positive droplets, or droplets that contain both the reference and GoI, plus the droplets that contain the GoI and dividing by the double positive droplets plus the droplets containing the reference. Then multiply the resulting value by 2 (Fig. 1).

#### *3.2.2 Single Nucleotide Variation (SNV) Assay Primer Design*

Three primers are designed for each SNV genotyping assay: wild type and mutant genotyping primers and one common primer upstream or downstream of the genotyping primer. The genotyping primers should be approximately 20 bp and identical to one another except for the one with the SNV specific position at the last 3' base position (Fig. 2). IDT's primer quest is a useful tool for designing these genotyping assays. The specificity of any primer



**Fig. 1** Example CNV Assay: Quanta Soft generated 2D plot of a copy number variation assay targeting one gene of interest (smaller amplicon) and one reference gene that has been artificially tailed in order to generate a larger amplicon. Gene of Interest copy number (Gol CN) is determined by counting the number double positive droplets plus the droplets that contain the Gol and dividing by the double positive droplets plus the droplets containing the Ref then multiplying by 2



**Fig. 2** SNV Genotyping primer design and assay structure. Black arrow represents the shared genomic primer, the blue arrow represents a wild type allele specific nontailed primer, and the green arrow represents an artificially tailed mutant allele specific primer

designed without PrimerQuest should be verified using the Blat tool [2]. Artificial 5' noncomplementary tails to one genotyping primer allows for differentiation between droplets that contain the mutant or wild type alleles based upon their differential amplicon

lengths (and consequent difference in fluorescence amplitude due to the difference in amount of dye binding per target molecule). The artificial tails minimize bias within the PCR reaction by allowing our primers to target an identical region of genomic DNA (*see* Subheading 3.2.3 for multiplexing suggestions).

### **3.2.3 Modifying Single-Color SNV Assays for Multiplexing**

In contrast to typical single-color multiplexing that involves two unique amplicons [3], we utilize the addition of noncomplementary 5' tails artificially increases amplicon length. This allows us to target a single genomic region that only differs if the targeted mutation is present; the presence of the mutation will result in 2 PCR products with unique lengths. Long tail sequences should be “AT” rich while short tails should be “GC” rich in order to minimize melting temperature variation. To multiplex the SNV assays, artificial tails should be added to either the wild type or mutant genotyping primer (*see* table below for examples). For example, to optimize a multiplexed SNV genotyping assay the user should create multiple iterations of tailed primers by adding repeats of “AAAT” to the 5' end of the mutant or reference specific primer in order to achieve 12, 24, 36, 48, 60, 80, and 100 non-complementary bases. All seven possible tailed primers sets (common forward + tailed reverse) should be individually multiplexed with the nontailed reference allele-specific primer set in a ddPCR assay. When designing multiplexed assays testing a variety of increasing tail lengths helps to ensure that the resulting reference and mutant amplicons provide a distinct ddPCR signals. Figure 3 shows an improvement on distinct signal between the mutant and reference populations that results from a 12 bp and a 100 bp tail added to the mutant specific primer. Adding long tails can significantly increase the melting temperature of primers; the user may add short “G” or “C” rich tails to the common genomic primer and the non-long tailed detecting primer to bring the melting temperatures of each individual primer within 3 °C of one another while maintaining amplicon length uniqueness.

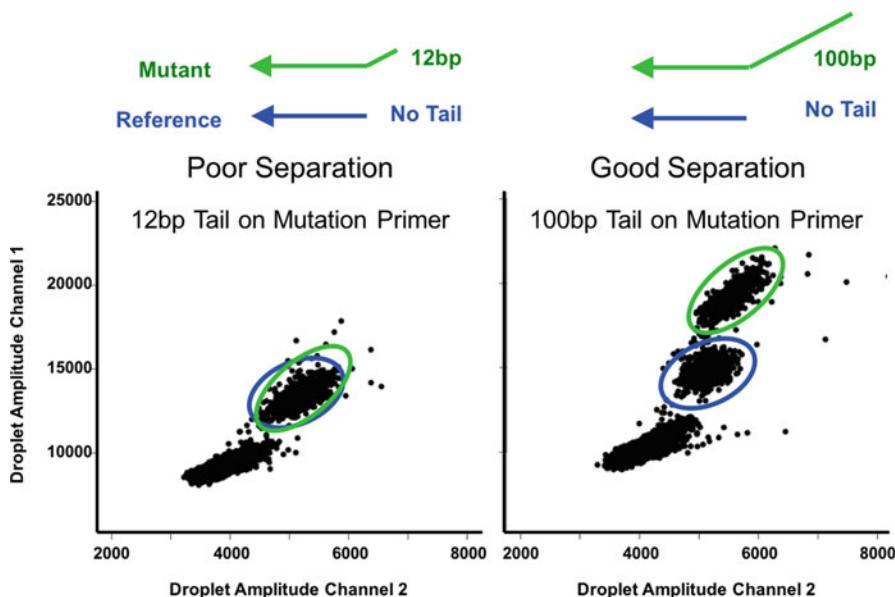
## **3.3 Preparing Primers and Template DNA for ddPCR Assays**

### **3.3.1 Preparation of Primers**

### **3.3.2 Restriction Digest Prior to ddPCR**

Forward and reverse primers may be combined when diluting to 10 µM and then added to the PCR master-mix. Ensure that both primers exist at 10 µM each in the final primer mix.

Template DNA should be digested using a restriction endonuclease according to the manufacturer’s protocol prior to use in ddPCR. Ensure that the cut site of the enzyme is outside of any targeted amplicons. High quality DNA and DNA extracted from FFPE tissues should be digested with a restriction digest; however ccfDNA extracted from plasma does not require restriction



**Fig. 3** Separation of SNV populations for multiplexed mutation detection. KRAS G12V multiplexed genotyping assay of a 50% KRAS G12V mutant DNA template. Repeats of “AAAT” are used to artificially “tail” the 5’ end of the mutant detecting primer (reverse) with either 12 or 100 bp. The overall amplitude of the mutant population increases in channel 1 (y-axis) as the length of the 5’ artificial tail increases (left versus right) resulting in better mutant versus reference population separation. Representative green and blue ovals placed over mutant and wild type populations respectively

digestion prior to use. The digest should be performed *prior to* ddPCR so that the digested template may be quantified before running SNV or CNV assays. This quantification allows the user to create a standard curve at a concentration equivalent to that of the patient sample before each run.

### 3.4 ddPCR Assay Preparation

(Adapted from Bio-Rad Product Insert # 10028376 Rev C.)

1. All High Quality DNA and DNA extracted from formalin-fixed, paraffin-embedded tissue (FFPET) should be digested with a restriction endonuclease that does not cut within the targeted amplicon.
2. Digested template DNA may be added to the ddPCR master-mix directly, but the final amount should not exceed 20 ng per 22  $\mu$ L reaction. See Note 2 regarding optimum template concentrations for SNV and CNV reactions.
3. Thaw all reagents at room temperature. Vortex all primers and spin down in a table-top microcentrifuge for 5–10 s. Mix by pipetting the QX200™ ddPCR™ EvaGreen Supermix 15 times with a P1000; spin down briefly.
4. Prepare the ddPCR PCR master-mix by combinin QX200™ Evagreen Supermix, shared (common) primer, reference allele

**Table 1**  
**Example SNV assay master-mix preparation**

Reagent	1 reaction	Final concentration
QX200 ddPCR™ EvaGreen Supermix	11.0 µL	1 ×
10 µM Common Primer	Variable	100–400 nM
10 µM Reference Allele Primer		50–100 nM
10 µM Mutant Allele Primer		50–100 nM
Template DNA	Variable	Up to 20 ng
UltraPure™ Water	Remaining volume to 22 µL	–

primer, mutant allele primer, and template DNA and water to a final volume of 22 µL per reaction as suggested in Table 1 (see Note 3 about multiple samples and buffer controls).

- (a) The final concentration of the common primer in the reaction should be equivalent to the concentration of the reference and wild type allele primers. (i.e., 100 nM common primer, 100 nM reference allele primer, 100 nM wild type allele primer).

### 3.5 Generating Droplets

(Adapted from Bio-Rad Manual #10031907)

1. Add 20 µL of the thoroughly mixed master-mix to the middle row of wells on the DG8™ cartridge that is snapped into the cartridge holder. Pipette only to the first stop to ensure that no bubbles form in the well (see Note 4).
2. Pipette 70 µL of QX200™ Droplet Generation Oil for Eva-Green into the wells of row 1 into the wells labeled “oil.”
3. Affix gasket over the cartridge holder and place into Droplet Generator, press the button to close door and begin droplet generation.
4. Using a P50 multichannel pipette aspirate 41 µL of droplets from the outlet well row using the slowest aspiration rate possible. This should take between 5–10 s.
5. Dispense 41 µL of the droplets into a eppendorf 96-well twin-tech semiskirted plate using the slowest dispense rate possible. This should take between 5 and 10 s.
6. Place one pierceable foil heat seal on top of the plate with the red strip facing up.
7. Place plate with heat seal in the plate sealer set to 180 °C for 5 s.
8. Carefully transfer the sealed plate to a thermal cycler.

**Table 2**  
**Thermal cycling conditions for ddPCR reactions**

Cycling step	Temperature (°C)	Time (mm:ss)	Ramp rate	Number of cycles
Set lid temperature to 105 °C				
Enzyme activation	95	5:00	2°/s	1
Denaturation	95	0:30		40
Annealing/extension	Variable <sup>a</sup>	1:00–2:00		
Signal stabilization	4	5:00		1
	90	5:00		1
Hold	4	∞		1

<sup>a</sup>Each primer set's annealing/extension temperature should be individually optimized using a thermal gradient within the primer's expected melting temperature. The annealing/extension time may also be adjusted to improve ddPCR signal. Additional information may be found in Bio-Rad's *Droplet Digital PCR Applications Guide*

### 3.6 Thermal Cycling Conditions

1. The thermal cycling conditions for both the SNV and CNV assays retain a majority of the same steps, however there is variation from primer set to primer set based on primer annealing temperatures. Table 2 outlines a general ddPCR protocol.

### 3.7 Reading Droplets and Data Analysis

#### 3.7.1 Preparing Template File and Reading Droplets

(Adapted from Bio-Rad Manual #10031906)

1. Create a new template using QuantaSoft. For both CNV and SNV assays be sure to select SuperMix: QX200 Evagreen ddPCR Supermix then select Assay Type: CNV1, and ensure that data will be collected for both channel 1 and channel 2 by selecting their individual boxes. Apply these settings for each well that contains an assay. Add experiment specific information into the provided boxes before saving the template.
2. After opening the Droplet Reader door, unsnap the top securing plate and insert the cycled-plate onto the QX200™ Droplet Reader in the correct orientation. Return the top securing plate and snap in place before closing the droplet reader door.
3. Open the appropriate template and click “Run” on the left hand side of the QuantaSoft program. Specify the preferred data collection direction (across rows or down columns) for the range of wells specified in step 1 to be read.

#### 3.7.2 QuantaSoft Analysis and Exporting Data

(Bio-Rad Manual #10031906)

1. After the run has completed select “OK” to return to the template main page.
2. Select wells to be analyzed and click “Analyze” on the left hand side of the screen.
3. The automated clustering provided by the QuantaSoft software will not differentiate between multiple positive populations for

the single-color assays. All Clustering on the program should be done manually, and you may use the provided selection tools to cluster populations.

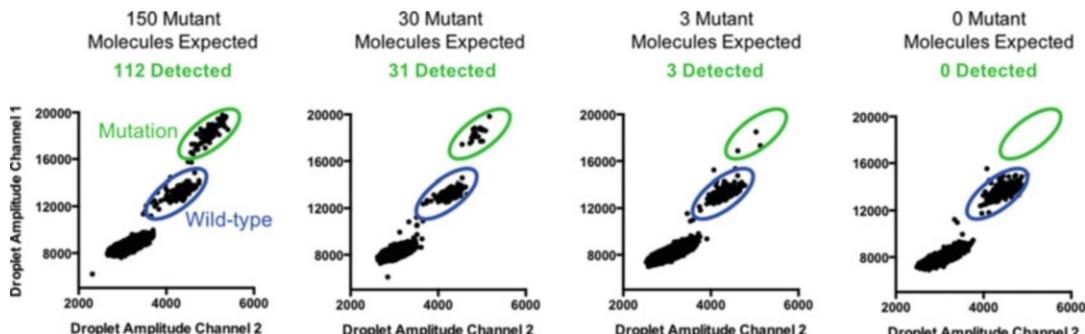
- To use any other clustering algorithms, you may export raw amplitude data select all wells with data and return to the template main screen. Select **options** and then **export cluster and amplitude data**. The resulting file will be a “.csv” that contains the channel 1 and channel 2 amplitudes as well as the cluster that was either automatically or manually assigned in the QuantaSoft software to each droplet.

### 3.7.3 Analyzing CNV Data

Copy number is determined by counting the total number of positive droplets amplified by the region of interest primer set and comparing this to the total number of positive droplets amplified by the reference primer set. The clustering algorithm used to determine gene copy number analyzes each data point individually in order to locate the center of a cluster [4]. The user may then define a threshold that is used by the algorithm to assign a cluster number to each data point.

### 3.7.4 Analyzing SNV Data

The multiplexed single-color ddPCR genotyping assay can be used to determine the absolute quantity of reference and variant alleles present within a given sample. Similar to the CNV assay, the number of reference or variant alleles can be determined by feeding the raw amplitude data into a standard clustering algorithm that accounts for Poisson distribution. The user may then count the number of positive events in either the mutant or wild-type positive clusters. In order to generate confidence intervals for calling mutant and wild type events, we run a series of standard curves parallel to each patient sample containing assay (Fig. 4).



**Fig. 4** Example In-Line Standard Curve. 2-Dimensional plot of raw channel 1 and channel 2 amplitude data for the KRAS G12V standard curve assay at mutant fractional amounts of 50, 10, 1, and 0%

---

## 4 Notes

1. In addition to the use of a PCR workstation and to reduce the risk of contamination pre and post-PCR. Pre- and post-PCR workspaces should be in physically separate locations with designated equipment and consumables. All workspaces and instruments should be thoroughly decontaminated once a week with a 10% bleach solution for 10 min and a subsequent wipe down with 70% ethanol solution.
2. We have found that the positive populations have the best resolution when using approximately 10 ng (~3000 genome equivalents) of template DNA per reaction in CNV assays, and 1–20 ng (~300–6000 genome equivalents) of template DNA for SNV assays. For SNV assays multiple replicates should be run in order to increase the total number of templates being assayed. This will help boost signal and reduce noise.
3. Reaction mixes may be made in batches depending on the variable being tested. Pool as many common reagents as possible to reduce variability of results. If <8 samples are to be run, fill all remaining sample wells with 20 µL of 1× QX200™ ddPCR™ Buffer control for EvaGreen.
4. Gently dislodge any bubble with a clean P20 pipette tip prior to starting droplet generation.

---

## Acknowledgments

This work was supported by the National Institutes of Health (NIH) grant NHGRI P01 HG00020526 (H.P.J., C.M.W.B.).

## References

1. Gray ES, Rizos H, Reid AL, Boyd SC, Pereira MR, Lo J et al (2015) Circulating tumor DNA to monitor treatment response and detect acquired resistance in patients with metastatic melanoma. *Oncotarget* 6(39):42008–42018. <https://doi.org/10.18632/oncotarget.5788>
2. Miotke L, Lau BT, Rumma RT, Ji HP (2014) High sensitivity detection and quantitation of DNA copy number and single nucleotide variants with single color droplet digital PCR. *Anal Chem* 86(5):2618–2624. <https://doi.org/10.1021/ac403843j>
3. McDermott GP, Do D, Litterst CM, Maar D, Hindson CM, Steenblock ER et al (2013) Multiplexed target detection using DNA-binding dye chemistry in droplet digital PCR. *Anal Chem* 85(23):11619–11627. <https://doi.org/10.1021/ac403061n>
4. Rodriguez A, Laio A (2014) Machine learning. Clustering by fast search and find of density peaks. *Science* 344(6191):1492–1496. <https://doi.org/10.1126/science.1242072>



# Chapter 19

## A Universal Droplet Digital PCR Approach for Monitoring of Graft Health After Transplantation Using a Preselected SNP Set

Julia Beck, Michael Oellerich, and Ekkehard Schütz

### Abstract

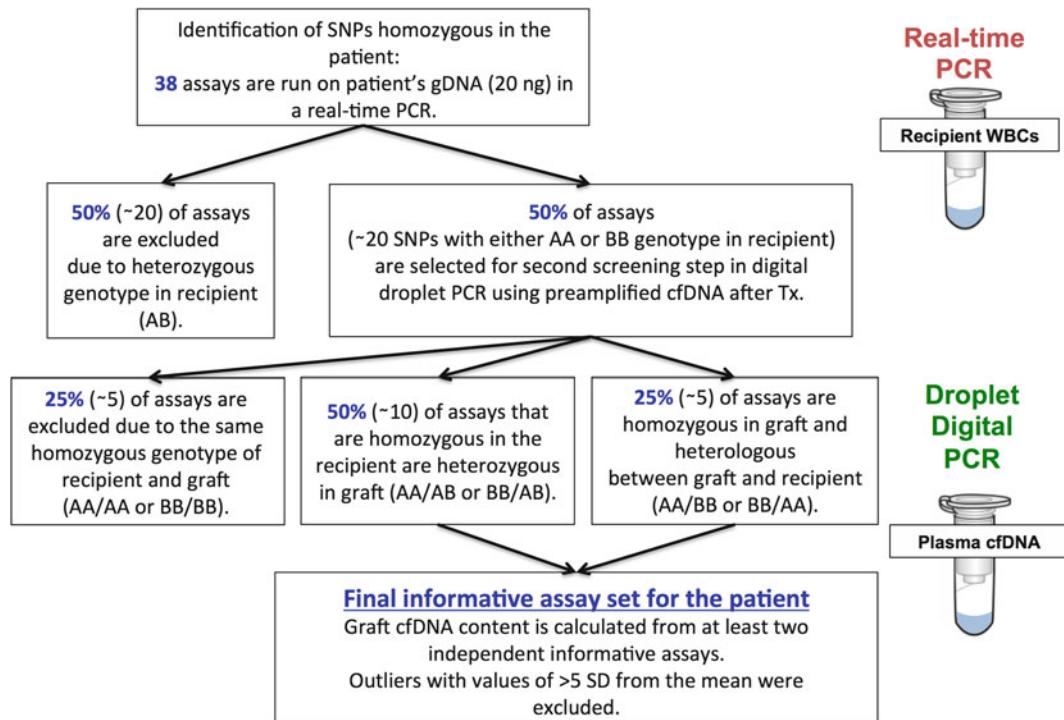
Transplanted organs release cell-free DNA into the bloodstream of the recipient. This graft-derived cell-free DNA (GcfDNA) is a sensitive biomarker for organ health, since higher GcfDNA levels are indicative of increased cell-death in the graft. This protocol describes a method to measure relative GcfDNA concentrations by ddPCR assays. The method uses a set of preselected SNP assays from which the informative SNPs for each recipient–donor combination are selected in a straightforward two-step procedure that requires only one blood draw. Sampling of donor tissue and separate genotyping is not required, rendering the technique applicable also to patients, whose transplantation was not recent. In these patients there will be mostly no access to donor DNA anymore.

**Key words** Transplantation, Rejection monitoring

---

### 1 Introduction

Graft-derived cell-free DNA is present in the circulation of patients, who have received organ transplants and it has gained high interest as specific biomarker for graft health and injury [1]. Besides the low amount of the total cell-free DNA that is extracted from the plasma, the GcfDNA is only a minor fraction with as little as ten genomic copies per mL of plasma. Because of the said minute amounts of the analyte, the accurate quantification of GcfDNA by any method relies on the specific detection of graft-specific genetic markers and was first restricted to the detection of Y-chromosome sequences in gender-mismatched donor–recipient pairs [2]. The use of HLA-mismatches enables GcfDNA testing [3], but requires upfront knowledge of the HLA genotypes and laborious assay optimization for every donor–recipient pair [4] for several hundred known HLA genotypes. With the advance and wide availability of single molecule counting techniques, such as high-throughput



**Fig. 1** The first screening step is performed in a real-time PCR using the recipient's genomic DNA extracted from buffy coat. In this step all SNP assays for which the recipient has a heterozygous genotype are eliminated, because they cannot be used in the quantification ddPCR. The next assay selection step uses preamplified cfDNA as template in ddPCR defining the final informative assay set for the individual patient. An informative assay detects an SNP that is homozygous in the recipient and for which the graft carries the heterologous allele, either in heterozygous or in (preferred) homozygous state. The percentages and numbers of assays given for each selection step are calculated for a minor allele frequency of 0.5 and vary between individual patients

sequencing (HTS) and droplet digital PCR (ddPCR), GcfDNA quantification became feasible as biomarker for graft health [5–7] under routine conditions. Both aforementioned techniques target informative single nucleotide polymorphisms (SNPs) (i.e., SNPs with different alleles in donor and recipient; Fig. 1), for which the two distinct alleles are directly counted, resulting in a percentage measure of GcfDNA. Higher levels indicate increased cell death in the donor organ. The thresholds below which transplanted organs are considered stable are about 10% for livers [7], 0.25–0.5% for hearts [5, 8], and about 1% for kidneys [8]. Compared to HTS, the quantification by ddPCR is less expensive (less than \$200 per test), enables shorter turnaround times (about 24 h) and can be performed with even one single sample at the same costs. The ddPCR method described in Beck et al. [7], which is explained in more detail herein, is based on a set of predefined and optimized assays, which are targeting SNPs that were preselected for having high

minor allele frequencies (MAF). Based on Hardy–Weinberg equilibrium calculations a SNP with a MAF between 0.4 and 0.5 has a nearly equal distribution of both alleles in the given population and therefore has a chance of 11.5–12.5% of having a different (homozygous) genotype in two individuals of this population (e.g., donor and recipient) (Fig. 1). To identify at least three such informative SNPs for each donor–recipient combination a set of only 30–35 need to be tested once upfront. By using such a preselected set of SNPs in conjunction with the described screening approach for the individual patient, it is possible to interrogate the informative SNP set using just one blood-draw from the recipient. Sampling of donor tissue and separate genotyping is not required, rendering the technique applicable also to patients, whose transplantation was not recent. In these patients there will be mostly no access to donor DNA anymore. The method has already been proven useful to confirm the lower therapeutic tacrolimus ranges in a set of 10 liver transplanted patients [9] and to closely monitor a marginal donor liver [10] and has been used in more than 500 samples from 107 liver transplant recipients as part of a prospective multicenter trial [11]. The described method includes a preamplification of the total extracted cfDNA has been shown to result in CVs < 15% for GcfDNA concentrations of 2% [7]. Also, the workflow for informative assays selection is described.

---

## 2 Materials

### 2.1 Sample Collection

1. Either: EDTA-anticoagulated blood in collection tubes (any vendor) Or: Cell-Free DNA BCT Tubes (Streck, Omaha, USA) (*see Note 1*).

### 2.2 Extraction of cfDNA/Cell Associated DNA

1. 15 mL Falcon® tubes (any vendor).
2. High Pure Viral Nucleic Acid Large Volume Kit (Roche Applied Science, Mannheim, Germany).
3. DNeasy Blood and Tissue Kit (QIAGEN, Hilden, Germany).
4. DNA LoBind Tubes (Eppendorf, Hamburg, Germany).
5. 0.1× Tris–EDTA, (Low TE) pH 8.0.

### 2.3 Real-Time PCR on Buffy Coat DNA (Recipient Germline)

1. FastStart Taq DNA Polymerase, dNTP Pack (includes: 10× PCR reaction Buffer with 20 mM MgCl<sub>2</sub>, MgCl<sub>2</sub> stock solution (25 mM) and PCR grade dNTP mix; Roche Applied Science, Mannheim, Germany).
2. Primers and probes acc. to Table 1 (any vendor).
3. Real-time PCR instrument (any vendor).

**Table 1** Primer and probe sequences

S84	CCTAATCACTCGTGAGGAGTG	ATCCACCATGATGCTCACAA	cccaCgggaggatgtcttgg	cccAtggggactctgtggcc	Yes
S85	GGAAACATCTAGACGCCA	CGGAATCGGGAGGCCAC	acacaAaggccctcccgat	acaGaggccctcccgat	Yes
S86	GTCCTCCCTCCCCAAAGGTGC	GCCAACCTCAAGGGCAGTT	agggaaaaacccatcAgatrcagt	tgaggattaaatgcacatCgaaatgt	Yes
S87	GGCATCTGAATTCAAGCTTGGTC	TTCTTCTAGITGGTCTGGTAGGCT	aggcttgatcacatCtccecc	acatggatggggaaAgt	Yes
S88	TGGTTATTGTTACTAGGTCCCACC	AGATAAAGCAAGATGTTGGCAGTGAG	aggactttatggggggCtgac	ctggaaaggccaatcgCcctc	Yes
S90	AAGGAAAAGGAAATTCTGGC	CCTGCCTATGCTCAGGCA	cgtTgcctctgcaggaa	gggcCtgtctgagcatag	Yes
S92	TTTATTTAAATGACTGTCAGGTC	TTTACAGACCTTCAAACCAC	cccgccatgtgcacatgtg	atrcgAggccacaatgtg	Yes
S94	CTGGGGCAGAGTGGAGATC	ATCCACCTCTGAACCCAGCC	aggacAtcgatcgatgg	cag/gtcctcgatcgatctt	No
S96	TCCCAGGCTCAGGTCAAGAT	GGATCAAATGGGCTGCTCCCT	tcgcgccttcgtggatgc	agggcAagacttggact	No
S97	AGCCCTGCACACTCACTTAC	TGGCATTCAGATCATCAGGCCTCT	ccatcgatgtcgActc	tgcaggaaatgcgg	Yes
S99	GGCAAAGTGGCAAGGGTCT	GCCTCTAAAGCTTGAGCCACA	ttggggccaGgtaccctgg	tggggccaAgtaccctgtt	No
S102	AACAGTGGCAGCCCTCTGT	ACACTGGTICATGGGGTTGTG	tggccratcttgcctraaCatg	aggcacatccatCttaggcc	Yes
S103	TTCTATATGAATTCTCTTCTAA	AAGCAGTCAGGAAGTATCC	ccctggggccatcaGtt	ccttggggccatcaAgttt	Yes
S105	ACCCCAAGAGGCTTTATAGGG	CCTTCCAAACGGGTTTGACC	ccactgggtggCccctc	atggaggaggAccgc	Yes
S108	ACACTCCTGCTGCGTCTG	TTCTCTCCCCACACTCCCAT	gatcccagtgCgtgg	atgcctccacAaccagt	Yes
S110	GGTCCCTACCGAGGTGGGTGA	CATTGCCAAGGACAGAGGGAGA	tttgttagggaaaggactCcaat	atcaggccatgtgatgtcc	No

Capital letters in probe sequences indicate the position of the inferred SNP

## **2.4 Preamplification of cfDNA**

1. NEBNext Ultra II DNA Library Prep Kit (New England Biolabs, Ipswich, USA).
2. NEBNext Singleplex Oligos for Illumina (New England Biolabs, Ipswich, USA).
3. 80% ethanol (freshly prepared).
4. Nuclease-free water.
5. 0.1× Tris-EDTA, (Low TE) pH 8.0.
6. EvaGreen® dye 20× (any vendor).
7. P5-Primer-Transplant (5'-CCTACACTCTTCCCTACACG ACGCTCTTCCGATCT-3').
8. P7-Primer-Transplant (5'-GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT-3').
9. AMPure XP Beads (Beckman Coulter, Brea, USA).
10. DNA LoBind Tubes (Eppendorf, Hamburg, Germany).
11. Magnetic rack/stand (any vendor).
12. Real-time cycler (any vendor).
13. PCR cycler (any vendor).
14. PCR plates according to recommendations of instrument vendor.

## **2.5 Droplet Digital PCR**

1. QX200 ddPCR System (Bio-Rad, Hercules, USA).
2. PCR Plate Sealer (any vendor).
3. PCR Machine (any vendor).
4. ddPCR Supermix for Probes (Bio-Rad, Hercules, USA).
5. Primer and probes acc. to Table 1 (any vendor).
6. Droplet Generation Oil for Probes (Bio-Rad, Hercules, USA).
7. DG8 Cartridges for QX200 Droplet Generator (Bio-Rad, Hercules, USA).
8. DG8 Gaskets for QX200 Droplet Generator (Bio-Rad, Hercules, USA).
9. Piercable Foil Heat Seal (Bio-Rad, Hercules, USA).
10. 96-well PCR plates (twin.tec semiskirted, Eppendorf, Hamburg, Germany).

## **2.6 SNP Probe Hydrolysis Assays**

1. A set of assays targeting 38 different SNPs is used.
2. All SNPs comply with the following criteria:
  - (a) Known and validated MAF of  $\geq 43\%$  in Caucasian whites and over all reported ethnicities (public databases: Hapmap or 1000Genomes).

- (b) Not located within or directly adjacent to an annotated repetitive element.
3. SNP-specific probes were designed using thermodynamic nearest-neighbor model calculations with a desired Tm of 65 °C at standard PCR buffer conditions (e.g., 0.18 mol/L monovalent cation equivalents [12] and 500 nmol/L DNA-/Primer) and were selected for maximum Gibbs free energy difference between allele binding (match/mismatch) at the given conditions. For each SNP one allele-specific probe carries a FAM fluorophore while the other SNP carries a HEX fluorophore in conjunction with BHQ1 as quencher. Respective PCR primers were designed to exhibit a melting temperature of 68 °C and a binding efficacy of ~95% at 60 °C, thermodynamically assessed using published formulas [13]. All probes and primers are given in Table 1.

## 2.7 Analyses

1. Quantasoft software (Bio-Rad, Hercules, USA).
2. Excel (Microsoft, Seattle, USA).

---

## 3 Methods

### 3.1 In informative Assay Selection

Figure 1 depicts an overview of the procedure employed to determine the informative assays, personalized for each recipient–donor combination. To select those SNPs with a homozygous genotype in the recipient, the patient’s germline DNA—extracted from white blood cells—is tested for all SNPs in real-time PCR format. The following step uses preamplified cfDNA as input in a ddPCR for testing only those assays, in which a homozygous recipient genotype was detected in the first step. Every assay showing an allele that is heterologous to the recipient is an informative assay. SNPs with heterozygous alleles in the graft show half of the percentages as compared to homozygous alleles. The assay selection procedure is applied only once for each patient, informative assays are stored in a database (Excel, Filemaker, SQL, etc.), and all subsequent samples are tested only with the patient’s personal informative assay set.

### 3.2 Sample Collection and Storage

1. Collect 10–20 mL blood per patient in either EDTA-anticoagulated blood tubes or Cell-Free DNA BCT tubes referring to respective protocols for proper handling. For EDTA blood, the plasma should be separated from blood cells within 1 h after collection. Blood samples collected in Cell-Free DNA BCT for cfDNA analysis are stable for 14 days (according to manufacturer), when stored between 6 and 37 °C.

2. Separate plasma from blood cells by centrifugation at  $2500 \times g$  for 10 min at 4 °C. Remove the plasma carefully without disturbing buffy coat layer.
3. Plasma should be aliquoted in 1–2 mL portions.
4. Collect buffy coat for extraction of recipients' germline DNA. Collection of buffy coat is needed for first time contact patients, only.
5. Collected plasma and buffy coat should be stored frozen at –20 °C if not used immediately.

### ***3.3 Extraction of cfDNA/Cell Associated DNA***

1. Subject 1–2 mL plasma aliquots to a second centrifugation at  $4000 \times g$  for 20 min at 4 °C to remove any remaining cells/cell debris.
2. Transfer supernatant to a fresh 15 mL tube. Extract cfDNA using the High Pure Viral Nucleic Acid Large Volume kit according to the manufacturer's instructions, but without the use of poly-A carrier RNA.
3. Make sure to collect final eluate (48–50 µL) in DNA Lopbind tubes with low TE buffer. Store at –20 °C until further processing (see Note 2).
4. For first-time contact patients extract the buffy coat DNA using the DNeasy Blood and Tissue Kit (QIAGEN, Hilden) according to the manufacturer's protocol.

### ***3.4 Real-Time PCR on Buffy Coat DNA***

1. Prepare real-time PCR in a total volume of 20 µL per reaction.
2. Each reaction should contain 2 µL FastStart 10× buffer with MgCl<sub>2</sub>, 200 µmol/L each dNTP, 2 units FastStart Taq, 900 nmol/L each primer, 250 nmol/L each probe (Table 1; if the use of MgCl<sub>2</sub> is indicated add 1.8 µL of the 25 mM PCR grade MgCl<sub>2</sub> stock solution to a final total concentration of 4.3 mmol/L).
3. Cycling conditions using a Roche LightCycler 480 are: 95 °C for 10 min, 50 × (95 °C for 30 s, 65 °C for 1 min) with fluorescence detection (FAM/HEX) in each cycle at 80 °C and ramp rates of 4.4 °C/s for heating and 2.2 °C for cooling (see Note 3).

### ***3.5 Preamplification of cfDNA***

Use the NEBNext II Ultra DNA Library Prep Kit for Illumina according to the manual with adherences to the following instructions:

1. Since cfDNA is already fragmented, no initial fragmentation of the extracted DNA is required. Use the 50 µL cfDNA elution directly for library preparation.
2. Follow the protocol for 5 ng–1 µg fragmented input DNA.

3. For ligation DO NOT dilute the adaptor tenfold, although this is suggested for input amounts of <100 ng. Use the NEBNext Singleplex adaptor (NEB, #E7350).
4. Perform the option 1.3B “Cleanup of Adapter-ligated DNA without Size Selection” for purification after adapter ligation step.
5. To the PCR amplification reaction add 1.5  $\mu$ L EvaGreen dye 20 $\times$  and 500 nmol/L of each primer P5-Primer-Transplant and P7-Primer-Transplant instead of the Universal PCR Primer/i05 and Index Primer/i07 (*see Note 4*). Perform PCR amplification in a real-time PCR instrument while closely monitoring the amplification using the FAM (green) channel of a real-time PCR instrument. Stop PCR just after the amplification curves leave the linear phase (*see Note 5*) and proceed to “Cleanup of PCR” according to NEBNext protocol.
6. Determine concentration by any appropriate method. Concentrations should be between 20 and 100 ng/ $\mu$ L (*see Note 6*).

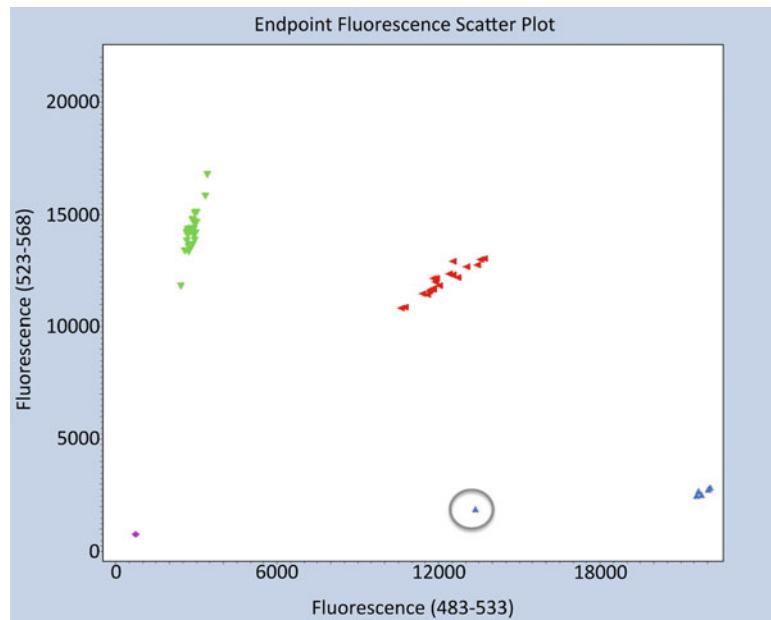
Preamplification of library can be completed within <3.5 h.

### **3.6 Droplet Digital PCR**

1. Prepare ddPCR reactions using 11  $\mu$ L of the 2 $\times$  ddPCR Supermix for Probes (Bio-Rad).
2. Each reaction should contain 30–100 ng preamplified cfDNA as template, 900 nmol/L of each primer, and 250 nmol/L of each probe in a total volume of 22  $\mu$ L.
3. Generate droplets using the QX200 droplet generator or automated droplet generator.
4. Cycling conditions using a Biometra 3000 Gradient Cycler are: 95 °C for 10 min, 50  $\times$  (94 °C for 30 s, 61 °C for 1 min), then 98 °C for 10 min, ramp rate 2.5 °C/s (*see Note 3*).
5. Some assays are supplemented with additional 2.3 mmol/L MgCl<sub>2</sub> (*see Table 1*).
6. Assays can be grouped according to their respective annealing temperature, so that several assays can be run on the same 96-well plate.
7. Positive controls with known heterozygote genotypes as well as nontemplate controls for each assay should be included in every run (*see Note 6*). After cycling droplets are read in a QX200 droplet reader instrument (Bio-Rad) selecting “rare event detection” and “FAM/HEX assay format.”

### **3.7 qPCR Analysis of Candidate Assays**

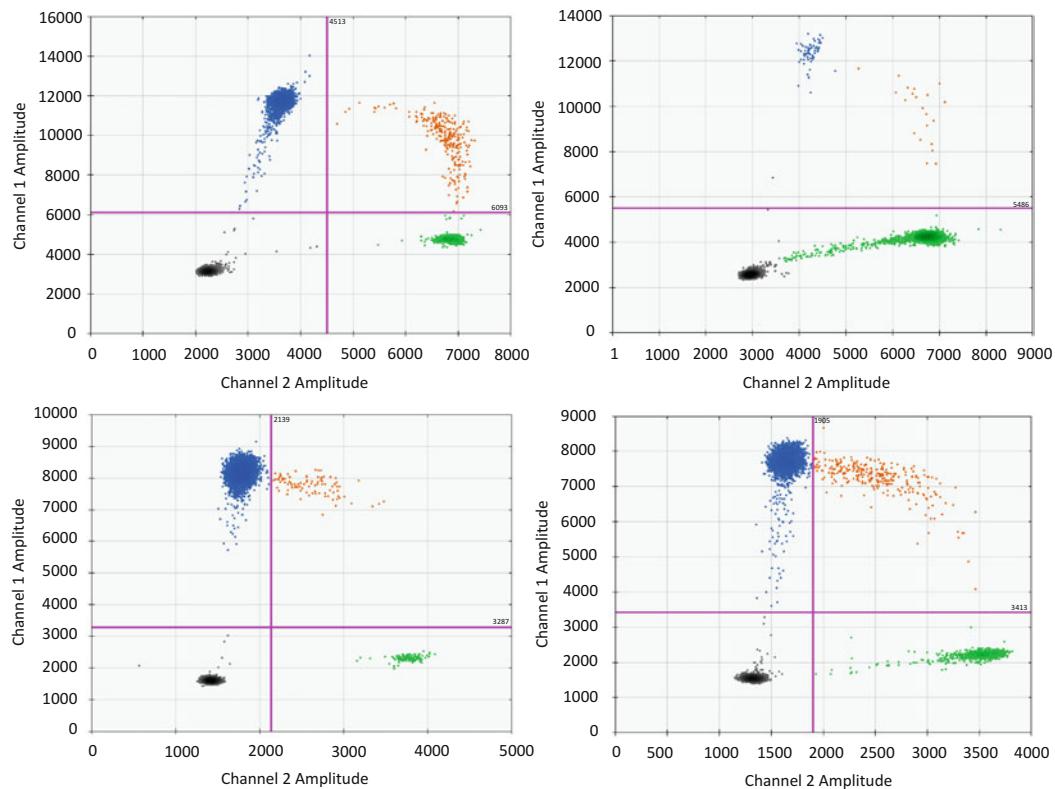
1. Genotypes are called using the “Endpoint Genotyping” module of the LightCycler 480 software.
2. Manual inspection and recalling might be necessary for some reactions.
3. Refer to Fig. 2 for example.



**Fig. 2** Exemplary qPCR assay run on a LightCycler480. A scatterplot of the “Endpoint Genotyping” analysis module (LightCycler480 Software, version: 1.5.0 SP3) is shown on the left (x-axis: FAM signal intensity, y-axis: HEX signal intensity). Genotypes called as AA are shown in blue, genotypes called as BB in green and AB genotypes in red. One of the samples (grey circle) shows lower FAM signal intensity, but is still unequivocally called as AA genotype

### 3.8 Droplet Digital PCR Analysis

1. Primary analysis is conducted using the Quantsoft instrument software.
2. Set the proper thresholds for each assay ensuring the best separation of the four droplet fractions FAM+/HEX-, FAM-/HEX+, FAM+/HEX+, and FAM-/HEX- (*see Note 7, Fig. 3*).
3. Check nontemplate controls for negativity (*see Note 6*). Select “Events” tab and check the box for display of total droplets.
4. Reactions should achieve a minimum of 10,000 total droplets.
5. Check the box for display of positive droplets.
6. Reactions should contain at least 100 positive droplets for the minor fraction in order to achieve a maximum CV of ~10% (*see Note 6*).
7. Select the “Ratio” tab in Quantasoft and check the box for display of “Fractional Abundance.”
8. Check the heterozygous positive controls for achieving the expected ratio of 50/50 for the two alleles (*see Note 6*).
9. Export results to comma separated values (.csv) file.



**Fig. 3** Four examples for ddPCR assays as 2D-plots (Quantasoft, version: 1.7.4) from upper left in clockwise order: S59 with informative B-allele (FAM channel) (17%); S63 with informative A-allele (FAM channel) (1.43%); S43 with informative B-allele (HEX channel) (8%); S38 with informative B-allele (HEX channel) (19%)

10. Open .csv file in Microsoft Excel and group all assays per sample.
11. All fractional abundances (FA) are displayed as  $A/(A + B) \times 100$ .
12. Correct the assays for which the informative allele is  $B$  by calculating  $100 - FA$ .
13. Correct the heterozygote informative alleles by calculating  $FA \times 2$ .
14. Refer to Table 2 for example.
15. Calculate the mean and standard deviation over all corrected FA for each sample.

#### 4 Notes

1. Cell-free DNA BCT Tubes contain, in addition to EDTA, a preservative reagent that prevents DNA release from blood cells during storage for up to 14 days, when stored at

**Table 2**  
Example calculation for initial ddPCR screen for informative assays

FA as exported from Quantasoft	100-FA for informative B allele	FA*2 for heterozygous informative alleles	Final FA	Tested recipient genotype	Inferred graft genotype
0			Not informative	BB	BB
0			Not informative	BB	BB
5.5		<i>11.0</i>	<i>11.0</i>	BB	AB
<i>12.5</i>			<i>12.5</i>	BB	AA
86.3	<i>13.7</i>		<i>13.7</i>	AA	BB
86.9	<i>13.1</i>		<i>13.1</i>	AA	BB
88.5	<i>11.5</i>		<i>11.5</i>	AA	BB
92.4	7.6	<i>15.2</i>	<i>15.2</i>	AA	AB
93.1	6.9	<i>13.8</i>	<i>13.8</i>	AA	AB
94.5	5.5	<i>11.0</i>	<i>11.0</i>	AA	AB
94.7	5.3	<i>10.6</i>	<i>10.6</i>	AA	AB
100			Not informative	AA	AA
Mean			12.5		
STDEV			1.6		

Every row depicts the values for one assay. Final FA at different calculation steps are shown in bold italics. Recipient genotype results are obtained from real-time PCR using DNA extracted from buffy coat. Note that informative SNPs have a higher chance of being in heterozygous rather than being in homozygous state in the graft

FA fractional abundance

6–37 °C (according to manufacturer). These tubes should be used if samples need to be shipped or cannot be centrifuged within 1 h after blood draw.

2. cfDNA concentrations in the eluate show wide individual variability with high concentrations seen for example soon after surgery. The median yield from 1 mL plasma was 30 ng (min 0.8 ng/mL, max: 280 ng/mL) in 270 samples obtained from liver transplanted patients. Due to the low cfDNA concentrations adequate techniques; for example, droplet PCR or at least fluorometric quantification should be used for accurate measurements. The protocol described here, however, does not require routine quantification of the cfDNA elutions, instead the samples can be subjected to preamplification directly. After a maximum amplification for 15 cycles a sample should then yield >20 ng library material (also see Note 6).

3. Annealing/extension temperature and/or ramp rates may need to be adjusted for different thermocyclers. 2 °C/s is recommended by Bio-Rad.
4. Shorter primers than in the original sequencing library protocol are used, since the libraries are not intended for sequencing.
5. Amplification reaction should be stopped at the end of the linear PCR phase in order to avoid over-amplification. Library amplifications with different amounts of input DNA will have different stopping points. Discontinue the real-time PCR when the first libraries reach their inflection point and then estimate the number of additional cycles for the lower concentration libraries. Remove the first libraries from the PCR plate and add more cycles to the others (*see Note 6*).
6. Quality Controls throughout the whole procedure are:
  - (a) Amplification curves and final concentration of preamplified libraries: less than 15 cycles needed; yield >20 ng/µL.
  - (b) Number of total droplets and positive droplets in ddPCR: >10,000 total; >100 positive for donor allele.
  - (c) Fractional abundance of heterozygote control DNA in ddPCR: target: 50%, accepted range: 47–53%.
  - (d) Positive droplets in ddPCR NTC control: <2 events.
  - (e) The analysis of at least four informative assays per patient sample is recommended.

If sample fails quality control checkpoint #1 it is recommended to reextract the plasma cfDNA and repeat library preparation. The initial extracted plasma volume might be raised. In case of sample failure for checkpoints #2–#4, repeat ddPCR.

7. The positions of droplet clusters in the 2D plot of the heterozygous positive control can help to set the thresholds in the patient samples. Heterozygous genotypes can be identified from the real-time PCR on patients' buffy coat DNA and then used as controls in the ddPCR.

## Acknowledgments

The authors thank Sarah Bierau and Stefan Balzer for their excellent technical assistance.

## References

1. Gielis EM, Ledeganck KJ, De Winter BY, Del Favero J, Bosmans JL, Claas FH, Abramowicz D, Eikmans M (2015) Cell-Free DNA: an upcoming biomarker in transplantation. Am J Transplant. <https://doi.org/10.1111/ajt.13387>

2. Zhang J, Tong KL, Li PK, Chan AY, Yeung CK, Pang CC, Wong TY, Lee KC, Lo YM (1999) Presence of donor- and recipient-derived DNA in cell-free urine samples of renal transplantation recipients: urinary DNA chimerism. *Clin Chem* 45(10):1741–1746
3. Gadi VK, Nelson JL, Boespflug ND, Guthrie KA, Kuhr CS (2006) Soluble donor DNA concentrations in recipient serum correlate with pancreas-kidney rejection. *Clin Chem* 52(3):379–382. <https://doi.org/10.1373/clinchem.2005.058974>
4. Lo YM (2011) Transplantation monitoring by plasma DNA sequencing. *Clin Chem* 57(7):941–942. <https://doi.org/10.1373/clinchem.2011.166686>
5. De Vlaminck I, Valentine HA, Snyder TM, Strehl C, Cohen G, Luikart H, Neff NF, Okamoto J, Bernstein D, Weisshaar D, Quake SR, Khush KK (2014) Circulating cell-free DNA enables noninvasive diagnosis of heart transplant rejection. *Sci Transl Med* 6(241):241ra277. <https://doi.org/10.1126/scitranslmed.3007803>
6. Snyder TM, Khush KK, Valentine HA, Quake SR (2011) Universal noninvasive detection of solid organ transplant rejection. *Proc Natl Acad Sci U S A* 108(15):6229–6234. <https://doi.org/10.1073/pnas.1013924108>
7. Beck J, Bierau S, Balzer S, Andag R, Kanzow P, Schmitz J, Gaedcke J, Moerer O, Slotta JE, Walson P, Kollmar O, Oellerich M, Schütz E (2013) Digital droplet PCR for rapid quantification of donor DNA in the circulation of transplant recipients as a potential universal biomarker of graft injury. *Clin Chem* 59(12):1732–1741. <https://doi.org/10.1373/clinchem.2013.210328>
8. Beck J, Oellerich M, Schulz U, Schauerte V, Reinhard L, Fuchs U, Knabbe C, Zittermann A, Olbricht C, Gummert JF, Shipkova M, Birschmann I, Wieland E, Schütz E (2015) Donor-derived cell-free DNA is a novel universal biomarker for allograft rejection in solid organ transplantation. *Transplant Proc* 47(8):2400–2403. <https://doi.org/10.1016/j.transproceed.2015.08.035>
9. Oellerich M, Schütz E, Kanzow P, Schmitz J, Beck J, Kollmar O, Streit F, Walson PD (2014) Use of graft-derived cell-free DNA as an organ integrity biomarker to reexamine effective tacrolimus trough concentrations after liver transplantation. *Ther Drug Monit* 36(2):136–140. <https://doi.org/10.1097/FTD.0000000000000044>
10. Kanzow P, Kollmar O, Schütz E, Oellerich M, Schmitz J, Beck J, Walson PD, Slotta JE (2014) Graft-derived cell-free DNA as an early organ integrity biomarker after transplantation of a marginal HELLP syndrome donor liver. *Transplantation* 98(5):e43–e45. <https://doi.org/10.1097/TP.0000000000000303>
11. Schütz E, Fischer A, Beck J, Harden M, Koch M, Wuensch T, Stockmann M, Nashan B, Kollmar O, Matthei J, Kanzow P, Walson PD, Brockmöller J, Oellerich M, Singal A (2017) Graft-derived cell-free DNA, a noninvasive early rejection and graft damage marker in liver transplantation: a prospective, observational, multicenter cohort study. *PLoS Med* 14(4):e1002286
12. von Ahsen N, Wittwer CT, Schütz E (2001) Oligonucleotide melting temperatures under PCR conditions: nearest-neighbor corrections for Mg<sup>2+</sup>, deoxynucleotide triphosphate, and dimethyl sulfoxide concentrations with comparison to alternative empirical formulas. *Clin Chem* 47(11):1956–1961
13. Schütz E, von Ahsen N (2009) Influencing factors of dsDNA dye (high-resolution) melting curves and improved genotype call based on thermodynamic considerations. *Anal Biochem* 385(1):143–152



# Chapter 20

## Detection and Quantification of HDR and NHEJ Induced by Genome Editing at Endogenous Gene Loci Using Droplet Digital PCR

Yuichiro Miyaoka, Steven J. Mayerl, Amanda H. Chan,  
and Bruce R. Conklin

### Abstract

Genome editing holds great promise for experimental biology and potential clinical use. To successfully utilize genome editing, it is critical to sensitively detect and quantify its outcomes: homology-directed repair (HDR) and nonhomologous end joining (NHEJ). This has been difficult at endogenous gene loci and instead is frequently done using artificial reporter systems. Here, we describe a droplet digital PCR (ddPCR)-based method to simultaneously measure HDR and NHEJ at endogenous gene loci. This highly sensitive and quantitative method may significantly contribute to a better understanding of DNA repair mechanisms underlying genome editing and to the improvement of genome editing technology by allowing for efficient and systematic testing of many genome editing conditions in parallel.

**Key words** Genome editing, TALEN, CRISPR/Cas9, HDR, NHEJ, ddPCR

---

### 1 Introduction

Genome editing allows for manipulation of genomes at will in any cell type. This can be used to precisely alter a single base, to insert a sequence of interest at a precise location or to introduce random insertions and deletions which can disrupt gene function. The tools for genome editing are sequence-specific nucleases that induce a double strand break (DSB) or nicks at targeted genomic regions. The induced DNA damage activates two major DNA repair pathways to repair the damage. One is nonhomologous end-joining (NHEJ), in which the broken DNA ends are joined together without any templates. NHEJ often generates random insertions and deletions at the site of repair. The other pathway is homology-directed repair (HDR), in which cells use homologous DNA as a repair template to precisely repair DNA. By delivering homologous donor DNA with a new DNA sequence along with editing

reagents, the genome can be precisely modified via donor-dependent repair [1].

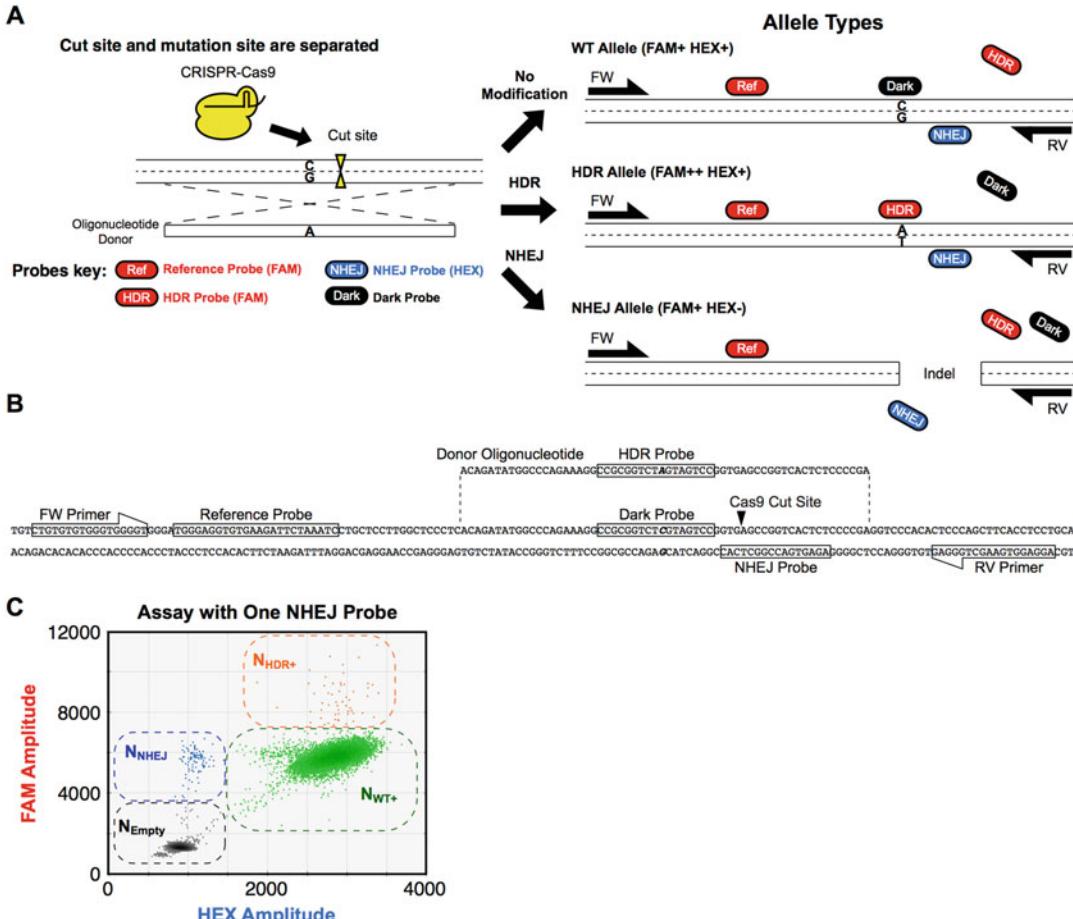
When precise donor-dependent HDR is required, it is important to measure genome editing outcomes—HDR and NHEJ—for successful application of genome editing, since NHEJ often predominates. This task has typically been done by using artificial reporter systems, deep sequencing, or gel-based assays, or by detecting only NHEJ as a surrogate for HDR because of technical limitations [2–5]. None of these methods are suitable for high-throughput screening. To overcome this hurdle, the method we describe here utilizes a combination of allele-specific hydrolysis probes and a new digital PCR technology called Droplet Digital™ PCR (ddPCR™) [6]. Our original ddPCR-based assay with probes for the wild-type (WT) and HDR alleles was designed to detect only HDR-mediated point mutagenesis without measuring NHEJ [7]. (In this protocol, we describe the original allele before genome editing as “WT,” but the same strategy can be applied to correct mutations.) The method we describe here [8] contains additional NHEJ probes to allow simultaneous quantification of both HDR-mediated point mutagenesis and NHEJ-mediated introduction of insertions and deletions (Figs. 1 and 2). ddPCR partitions a reaction into 20,000 nanoliter-scale water-in-oil droplets, so that each droplet contains from zero to ~10 copies of the genome targets which are analyzed by HDR and NHEJ allele-specific hydrolysis probes in individual droplet reactions. The allele-specific combination of fluorescent signals results in droplets that contain either HDR, NHEJ, or WT alleles, or combinations of these; these droplets in turn occupy distinct locations on the two-dimensional plot, allowing absolute quantification of discrete alleles (e.g., Fig. 1c). Therefore, HDR and NHEJ events can be detected in a highly sensitive and quantitative manner.

Measuring genome editing outcomes is critical in evaluating genome editing tools (ZFN, TALEN, CRISPR/Cas9, etc.) and conditions (concentration of nucleases, target sequences, donor types, etc.). This protocol provides researchers with a rapid way to measure HDR- and NHEJ-inducing activities of genome editing conditions.

---

## 2 Materials

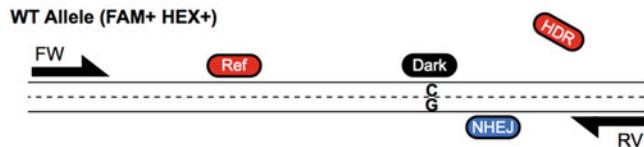
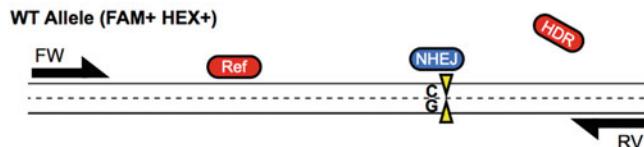
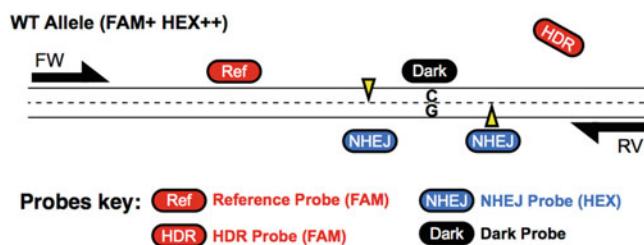
Prepare all solutions using ultrapure water (prepared by purifying deionized water to attain a sensitivity of  $18\text{ M}\Omega\text{ cm}$  at  $25\text{ }^{\circ}\text{C}$ ) and analytical grade reagents. Use filtered tips and compatible pipetters (such as RANIN) to avoid contamination of reagents in sample preparation and droplet generation. Prepare and store all reagents at  $-20\text{ }^{\circ}\text{C}$  (unless indicated otherwise). Diligently follow all waste disposal regulations when disposing of waste materials.



**Fig. 1** Separated cut and mutation sites: Assay design, example target sequence, and 2D droplet plot of edited genomic DNA. Depending on the editing strategy and the relative positions of cut site and edit site, the positions of assay probes and the need for a competitive blocking (dark) probe vary. The reference and HDR probes are constant regardless of the location or number of cut sites. (a) Assay design when cut site and mutation site are separated. Outcomes for three different alleles are shown (retention of WT sequence, alteration of a single base via HDR, or indel via NHEJ). The HDR and NHEJ probes are located on nonoverlapping mutation and cut sites, respectively. To prevent nonspecific binding of the HDR probe to the original “WT” sequence, a nonfluorescent (dark) “WT” probe is included that competes with the HDR probe for WT allele binding (see Figs. 2, 4, and 5 for other cases). (b) Assay design when cut site and mutation site overlap. See Fig. 4 legend for details. (c) Assay design when two cut sites are introduced. See Fig. 5 legend for details

## 2.1 Reagents

1. 20× FAM and HEX hydrolysis probe and primer assay mixtures in water. Components are shown in Table 1.
2. WT genomic DNA in water adjusted to 100–150 ng/μl.
3. Synthetic double-stranded DNA (gBlocks, Integrated DNA Technologies) that contains the point mutation at the desired edit site (HDR control) or a 2-bp deletion at the predicted nuclease cut site (NHEJ control). Lyophilized gBlocks are

**A Cut site and mutation site are separated****B Cut site and mutation site overlap****C Two cuts are introduced**

**Fig. 2** Comparison of three ddPCR assay designs for three different editing designs. **(a)** Assay design when cut site and mutation site are separated. See Fig. 1a legend for details. **(b)** Assay design when two cut sites are introduced. See Fig. 4 legend for details. **(c)** Assay design when cut site and mutation site overlap. See Fig. 5 legend for details

**Table 1**  
**Components of 20× assay mixtures**

Component	Concentration ( $\mu\text{M}$ )
Forward primer	18
Reverse primer	18
Reference probe (FAM)	5
HDR probe (FAM)	5
NHEJ probe(s) (HEX/VIC)	5 (each, if multiple probes are used)
Dark probe (3' phosphate is added)	10 (depending on the assay)

resuspended in 250  $\mu\text{l}$  TE + 100 ng/ $\mu\text{l}$  polyA carrier. Two additional 200-fold dilutions in TE + polyA result in a master stock of around 40,000 copies/ $\mu\text{l}$  that is maintained in LoBind tubes (Eppendorf) (see Note 1). Dilute the master stock by 20-fold to make a solution of around 2000 copies/ $\mu\text{l}$  for use.

Note that these are concentrations estimates and can be off by tenfold in either direction.

4. Restriction enzyme that does not cut within the amplicon. HindIII-HF, CviQI, MseI, AluI, HaeIII restriction enzymes (NEB) have all been validated to work directly in the ddPCR Supermix. 2–4 U restriction enzyme can be directly added to a 20  $\mu$ l ddPCR reaction without additional incubation.
5. ddPCR™ Supermix for Probes (No dUTP) (Bio-Rad 186-3024).
6. ddPCR™ Buffer Control Kit (Bio-Rad 186-3052).
7. Droplet Generation Oil for Probes (stored at room temperature) (Bio-Rad 186-3005).
8. DG8™ Gaskets for QX200™ Droplet Generator (Bio-Rad 186-3009).
9. DG8™ Cartridges for Droplet Generator (Bio-Rad 186-4008).
10. Pierceable Foil Heat Seal (Bio-Rad 181-4040).
11. Eppendorf twin.tec 96-well plates, semiskirted (Fisher 951020346).
12. Rainin filter pipette tips (*see Note 2*).
13. Genomic DNA in distilled water or TE isolated from cells treated with genome editing tools to induce HDR and/or NHEJ.

## 2.2 Equipment

1. QX100™ or QX200™ Droplet Digital™ PCR system (Bio-Rad 186-4001).
2. PX1™ PCR Plate Sealer (Bio-Rad 181-4000).
3. DG8™ Cartridge Holder (Bio-Rad 186-3051).
4. 96-well thermocycler (such as Bio-Rad C1000 that has a uniform block with deep wells, 2 °C/s ramp rate, and temperature gradient capability).
5. Rainin 20  $\mu$ l eight-channel pipette.
6. Rainin 50  $\mu$ l eight-channel pipette (*see Note 3*).

---

## 3 Methods

To detect HDR and NHEJ at the same time, four different kinds of probes are designed within a single amplicon [8]. The first is a FAM (reference) probe that is nonoverlapping with the cut site and which always binds to the genomic DNA. It provides a positive FAM amplitude for counting total genome copies in the sample (contributing to a FAM+ or FAM++ signal, Fig. 1). The second is a

HEX (NHEJ) probe that binds at the cut or nick site of the genomic locus. It has a wild type sequence, so in the case where NHEJ occurs and results in either insertion or deletion of sequences at the cut site, the probe will no longer be able to bind resulting in a loss of HEX signal. Loss of HEX signal with retention of the reference probe FAM signal (FAM+, HEX-) identifies molecules that have undergone NHEJ (Fig. 1b, c). The third probe is another FAM probe (HDR probe), which will bind to the DNA when precise edits have occurred indicating an HDR event. In the case of HDR, the FAM amplitude will further increase separating the HDR population from the wild type (FAM++, HEX+). The fourth probe is a nonfluorescent (“dark”) probe having “WT” sequence which prevents non-specific-binding of the HDR probe to the unaltered “WT” allele (*see* Fig. 2, assay designs a and c).

### **3.1 Design of Hydrolysis Probes and Primers**

1. Modify the settings of Primer3Plus (<http://primer3plus.com>) compatible with the master mix: 50 mM monovalent cations, 3.0 mM divalent cations, 0 mM dNTPs, SantaLucia 1998 thermodynamic and salt correction parameters (*see* ref. 9 and further guidance below).
2. For the target region of interest, position the predicted nuclelease cut sites mid-amplicon, with 75–125 bp flanking either side up to and including the primer binding sites (*see* Note 4).
3. Set melting temperatures for primers, reference probes, NHEJ probes, HDR probes, and Dark probes as  $55 \pm 1$  °C, 60 °C,  $57 \pm 1$  °C, 55 °C, and 57 °C, respectively.
4. At least one primer must be positioned outside the donor molecule sequence to detect integrated edits (Fig. 1b).
5. Reference probe and primers are designed distant from the cut site to avoid loss of binding sites by NHEJ.
6. In some cases, a dark, nonextendible oligonucleotide (3' phosphorylation) is designed to block cross-reactivity of the HDR probe and the WT sequence. This is generally required if the HDR edit site and nuclease cut site do not closely overlap (Figs. 1a and 2a, c). If they do directly overlap, the NHEJ probe can serve as the competitive blocker (Fig. 2b). When measuring NHEJ induced by dual nuclease systems, two NHEJ probes should be designed for the two cut/nick sites (Fig. 2c).
7. Probe position and number vary depending on the relative positions of the cut site(s) and edit site (Figs. 1 and 2). Our assay for RBM20 is shown as an example (Fig. 1b).

### **3.2 Assay Validation**

1. Mix the reagents shown in Table 2 in a well of a 96-well plate to make a 25-μl reaction (*see* Note 5).

**Table 2**  
**Reagents for assay validation**

Reagent	Amount per well
ddPCR Supermix for Probes (no dUTP)	12.5 µl
20× assay mixture	1.25 µl
Restriction enzyme of choice	2–4 U
WT genomic DNA in water	Up to 150 ng
HDR and NHEJ gBlock control solutions with 2000 copies/µl	1 µl each
Water	To give 25 µl final reaction volume

2. Carefully apply 20 µl of the mixture into each of the 8 “sample” wells of a DG8 Droplet Generator Cartridge (*see Note 6*).
  3. Apply 70 µl of Droplet Generation Oil for Probes into each of the 8 “oil” wells of the DG8 Droplet Generator Cartridge (*see Note 7*).
  4. Hook a DG8 Gasket for QX200 Droplet Generator onto the DG8™ Cartridge Holder.
  5. Put the holder in the droplet generator to generate droplets in 8 wells (in ~2 min).
  6. Transfer droplets into a semiskirted Eppendorf twin.tec 96-well plate by using an eight-channel pipette set to 45 µl (*see Note 8*).
  7. Seal the plate with a Pierceable Foil Heat Seal using the PX1™ PCR Plate Sealer set to 180 °C, 5 s.
  8. For assay validation, a two-step thermal cycling with a 50–60 °C gradient is recommended (Table 3). If assay amplicon exceeds 150 bp in length, a three step thermal cycling protocol with a discrete 72 °C extension step can be helpful (Table 4).
  9. Analyze the droplets by the Droplet Reader.
  10. To determine the limit of detection of an assay, analyze 100–150 ng of WT genomic DNA samples spiked with different amounts of HDR or NHEJ gBlock. The limit of detection is the lowest amount of gBlock that gives positive signal significantly higher than the background noise by comparison of results with and without gBlock.
- 3.3 Digital PCR Detection of HDR and NHEJ Events in Genomic DNA Samples**
1. Dilute genomic DNA samples to 100–150 ng/µl in distilled water or TE (*see Note 9*).
  2. Assemble the master mix below on ice by mixing the reagents shown in Table 5 (volumes shown are for one reaction to which DNA must still be added). Multiply these by the number of

**Table 3**  
**Two-step thermal cycling for ddPCR (all the steps ramped by 2 °C/s)**

Step	Temperature (°C)	Duration and repeat
1	95	10 min
2	94	30 s
3	59	1 min, (repeat steps 2 and 3, 39 more times)
4	98	10 min
5	12	Hold

**Table 4**  
**Three-step thermal cycling for ddPCR (all the steps ramped by 2 °C/s)**

Step	Temperature (°C)	Duration and repeat
1	95	10 min
2	94	30 s
3	58	1 min
4	72	2 min, (repeat steps 2, 3, and 4, 39 more times)
5	98	10 min
6	12	Hold

**Table 5**  
**Reagents for quantification of HDR and NHEJ in genomic DNA samples ( $X \mu\text{l}$  (100–150 ng) of a genomic DNA sample later)**

Reagent	Volume per well
ddPCR Supermix for Probes (no dUTP)	12.5 $\mu\text{l}$
20× assay mixture	1.25 $\mu\text{l}$
Restriction enzyme of choice	2–4 U
Water	Up to 25— $X \mu\text{l}$

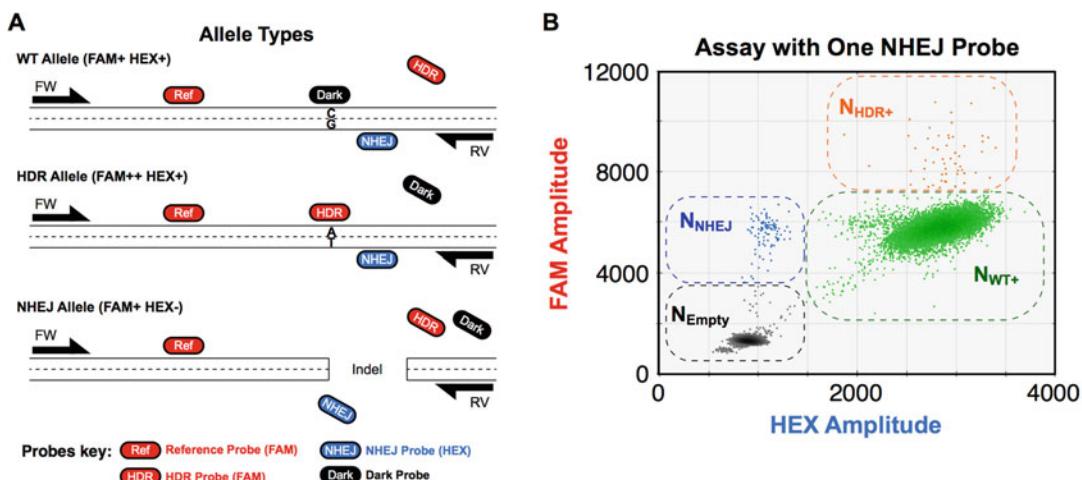
samples, a distilled water only negative control, and the positive control mixture (WT genomic DNA with synthetic alleles as made in Subheading 3.2, above).

- Aliquot 25— $X \mu\text{l}$  of master mix into eight-tube PCR strips or a 96-well plate (any 96-well plate).
- Add  $X \mu\text{l}$  (for a total of 100–150 ng) of genomic DNA per sample. Also, add controls to tubes or wells designated for the controls.

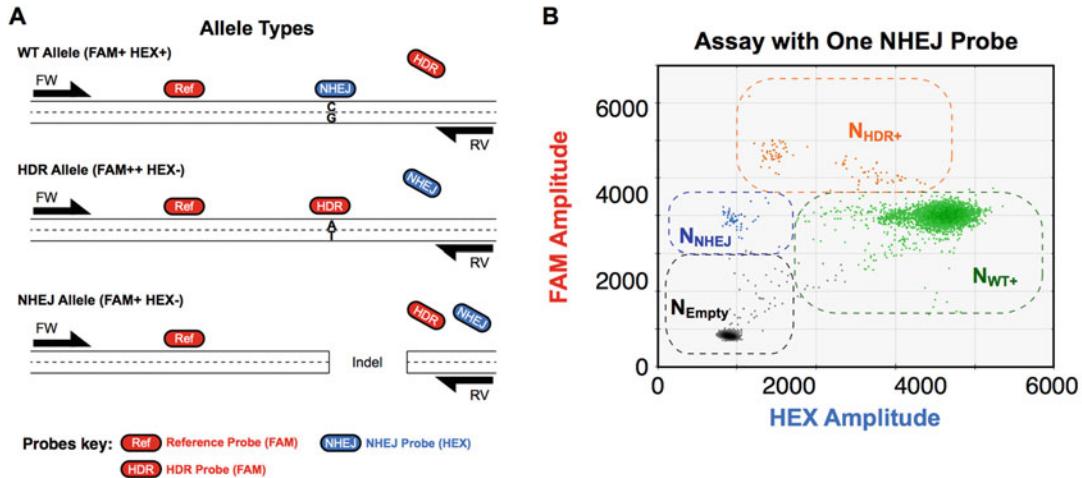
5. If the total number of samples to be analyzed is not a multiple of 8, the remaining empty tubes of the PCR strip must be filled with 25  $\mu$ l of ddPCR™ Buffer Control Kit diluted to 1× by distilled water.
6. Briefly spin the PCR tubes or plates down.
7. Gently pipette up and down 7–8 times to mix reactions by a 20  $\mu$ l eight-channel pipette and carefully apply the mixtures into “sample” wells.
8. Generate droplets as in Subheading 3.2, steps 2–6.
9. Repeat steps 4–8 until droplets are generated for all the samples.
10. Seal the plate with foil using the PX1™ PCR Plate Sealer set to 180 °C, 5 s.
11. Perform a thermal cycling with the best temperature found in Subheading 3.2.
12. Analyze the droplets by the ddPCR analyzer in the same way as Subheading 3.2.

### 3.4 Analysis of Droplet Digital PCR Data

1. Go to “Analyze” and then “2-D Amplitude” in the QuantaSoft software.
2. In a successful assay, distinct negative, WT, NHEJ, and HDR populations should be seen in the 2D-plot (Figs. 3, 4, and 5).
3. If four distinct populations cannot be detected, the probes and/or primers must be redesigned.
4. By using the lariat tool, gate the droplets without any template ( $N_{Empty}$ ), droplets only with NHEJ allele ( $N_{NHEJ}$ ), droplets



**Fig. 3** Separated cut and mutation sites, with one NHEJ probe: Assay design and 2D droplet plot with droplet group definitions for analysis. Details as in Fig. 1a, c



**Fig. 4** Cut site and mutation site overlap. (a) Assay design when cut site and mutation site overlap. In these cases, the NHEJ probe overlaps with and is on the same strand as the HDR probe. The NHEJ probe thus competes with the HDR probe for WT binding and a dark probe is not necessary. Also, the HDR allele becomes FAM++ and HEX-. (b) Other droplet populations defined as in Fig. 1c

with WT allele and both WT and NHEJ alleles ( $N_{\text{WT+}}$ ), and other droplets with HDR allele ( $N_{\text{HDR+}}$ ) as black, blue, green, and orange populations, respectively (Figs. 3, 4, and 5).

- To obtain the numbers of droplets in each population, export the data as a CSV file by clicking “Export CSV” icon in the QuantaSoft software. In the CSV file, the Ch1+Ch2+, Ch1 +Ch2-, Ch1–Ch2+, and Ch1–Ch2– columns are  $N_{\text{HDR+}}$ ,  $N_{\text{NHEJ}}$ ,  $N_{\text{WT+}}$ , and  $N_{\text{Empty}}$ , respectively. Use formulas below to calculate the frequencies of WT, HDR, and NHEJ alleles from these numbers.
- The standard formula for ddPCR quantification is:

$$c = -\ln(N_{\text{neg}}/N_{\text{total}})/V_{\text{droplet}}$$

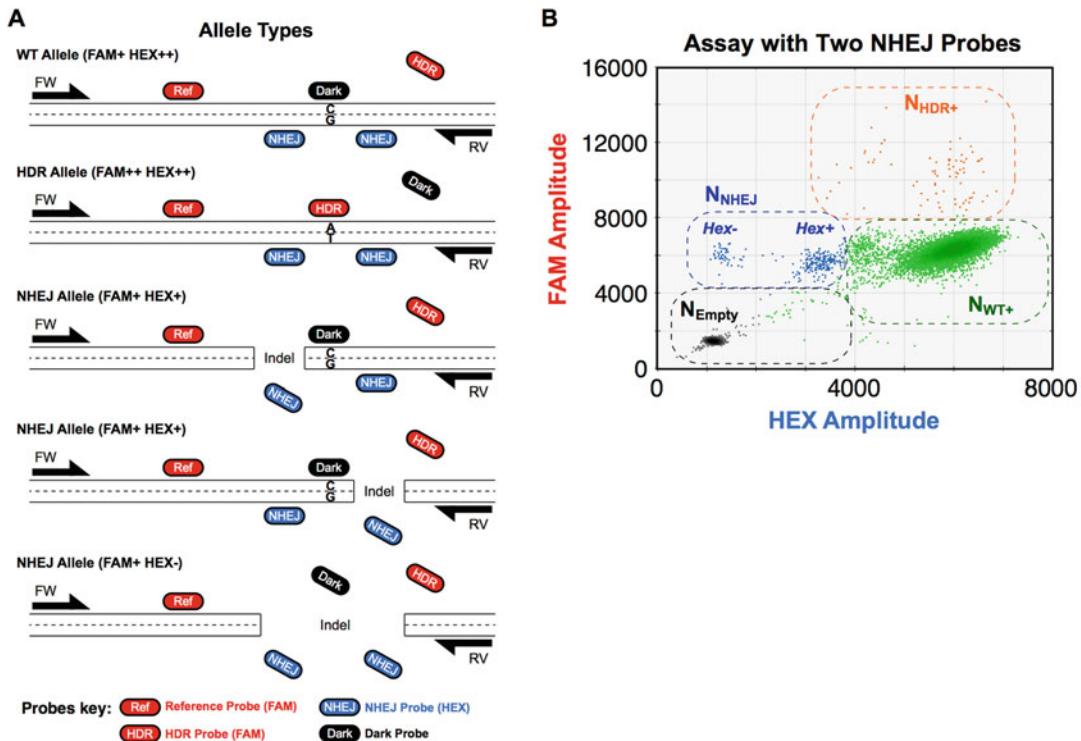
where

$c$  = copies per microliter of initial reaction  $N_{\text{neg}}$  = the number of droplets that do not contain the species of interest.

$N_{\text{total}}$  = the total number of droplets.

$V_{\text{droplet}}$  = the volume of an individual droplet (0.85 nl with current reaction conditions).

- In this ddPCR-based assay, some of the droplet populations cannot easily be separated, such as the droplet group containing WT and NHEJ + WT droplets ( $N_{\text{WT+}}$ ) or the droplet group containing HDR and HDR + WT droplets ( $N_{\text{HDR+}}$ ). To



**Fig. 5** Two cut sites are introduced: ddPCR assay design and 2D droplet plot. **(a)** Assay design when two cut sites are introduced. Because dual Cas9 systems introduce two cuts, two NHEJ probes are included in the assay to detect possible indels on either side of the mutation site. When neither NHEJ probe competes with the HDR probe—as shown here—a dark probe is also designed to avoid binding of the HDR probe to the WT allele. Because two NHEJ probes are included in this assay, the WT allele is detected as FAM+ and HEX++, and the NHEJ alleles are detected as FAM+ and HEX+, or FAM+ and HEX-. **(b)** Two-dimensional plot of an assay with two NHEJ probes. Definitions of  $N_{\text{Empty}}$  and  $N_{\text{HDR}+}$  populations are the same as in Fig. 1c. However, because two NHEJ probes are included in this assay, there are two droplet populations containing only NHEJ alleles—one that lost one of the two NHEJ probe binding sites of NHEJ alleles (FAM+ HEX+) and one that lost both (FAM+ HEX-). All other droplets were gated as  $N_{\text{WT}+}$  population (FAM+ HEX++). These definitions are used to calculate the HDR and NHEJ allelic frequencies (see Subheading 3.4)

quantify in this case, an appropriate subset of droplets is used to calculate  $N_{\text{neg}}$  and  $N_{\text{total}}$  (see Note 10).

#### Definitions

- $N_{\text{empty}} = \text{number of droplets in the double-negative cluster labeled "empty."}$
- $N_{\text{NHEJ}} = \text{number of droplets in the cluster labeled "NHEJ."}$
- $N_{\text{WT}+} = \text{number of droplets in the clusters labeled "WT+"}.$
- $N_{\text{HDR}+} = \text{number of droplets in the clusters labeled "HDR+"}.$

8. For NHEJ quantification

- $N_{\text{neg}} = N_{\text{empty}}$
- $N_{\text{total}} = N_{\text{empty}} + N_{\text{NHEJ}}$

9. For HDR quantification

- $N_{\text{neg}} = N_{\text{empty}} + N_{\text{NHEJ}} + N_{\text{WT+}}$
- $N_{\text{total}} = N_{\text{empty}} + N_{\text{NHEJ}} + N_{\text{WT+}} + N_{\text{HDR+}}$

10. For WT quantification

- $N_{\text{neg}} = N_{\text{empty}} + N_{\text{NHEJ}} + N_{\text{HDR+}}$
- $N_{\text{total}} = N_{\text{empty}} + N_{\text{NHEJ}} + N_{\text{WT+}} + N_{\text{HDR+}}$

11. Frequencies of WT, HDR, and NHEJ are calculated as:

$$F_{\text{WT}} = c_{\text{WT}} \times 100 / (c_{\text{WT}} + c_{\text{NHEJ}} + c_{\text{HDR}})$$

$$F_{\text{HDR}} = c_{\text{HDR}} \times 100 / (c_{\text{WT}} + c_{\text{NHEJ}} + c_{\text{HDR}})$$

$$F_{\text{NHEJ}} = c_{\text{NHEJ}} \times 100 / (c_{\text{WT}} + c_{\text{NHEJ}} + c_{\text{HDR}})$$

where

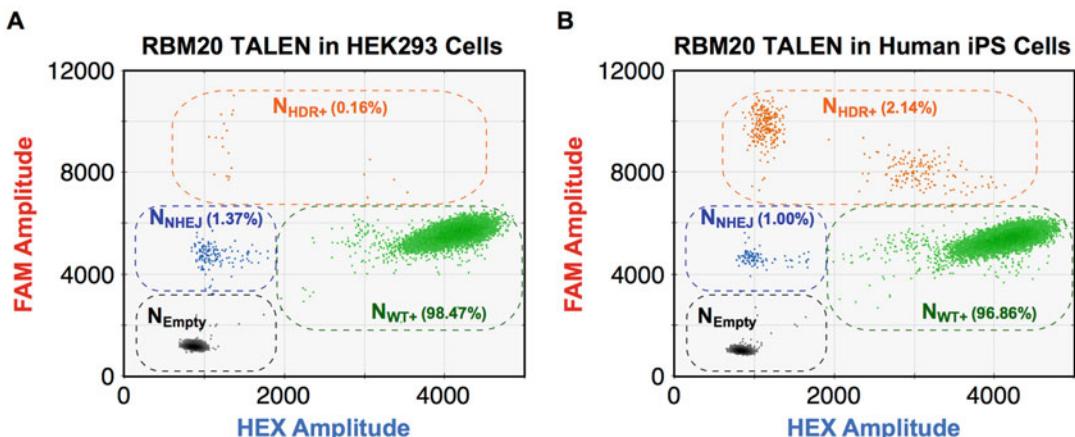
$F$  = allelic frequency (%).

$c_{\text{WT}}$  = the WT allelic concentration.

$c_{\text{HDR}}$  = the HDR allelic concentration.

$c_{\text{NHEJ}}$  = the NHEJ allelic concentration.

An example of analysis is shown in Fig. 6.



**Fig. 6** Different HDR and NHEJ frequencies induced by TALEN in RBM20 in HEK293 cells and human induced pluripotent stem cells. **(a and b)** HDR and NHEJ allelic frequencies induced by the same TALENs targeting RBM20 in HEK293 cells **(a)** and human induced pluripotent stem cells **(b)**. The NHEJ and HDR probes directly compete to each other, so  $N_{\text{HDR+}}$  population was FAM++ HEX– as in Fig. 4. The frequencies are shown in the two-dimensional plots. The two different cell types showed different HDR and NHEJ frequencies

---

## 4 Notes

1. Master high-copy gBlock stocks should be kept in a post-PCR environment to avoid contamination. The control gBlock stocks must be handled with utmost care to avoid contamination, only bringing highly dilute solutions into the assay setup area.
2. Filtered tips must be used to avoid contamination.
3. In order to slowly pipette droplets to avoid shearing of them, a 50  $\mu$ l pipette is recommended rather than a 200  $\mu$ l pipette.
4. Predicted cut sites are 3 bp upstream of PAM for CRISPR and equidistant between DNA binding domains for TALEN or FokI-dCas9.
5. Handle positive control mixtures carefully (e.g., careful opening of the control stock tubes and changing gloves after handling the control stock). If assay contamination is suspected, remake the 20 $\times$  assay.
6. Avoid creating bubbles in the DG8 sample wells. Bubbles floating on the surface of the sample may not affect droplet generation, but bubbles in the bottom of the well will disrupt droplet generation. Spinning prepared samples before droplet generation in tubes or plates removes bubbles. Take note of orientation of the cartridge with respect to sample order to ensure correct loading of the final Eppendorf twin.tec 96-well plate.
7. Samples should be loaded onto the chip prior to droplet generation oil.
8. Do not press the pipette tightly to the bottom of the cartridge or pipette too vigorously as this will shear the droplets. Cover the PCR plate with the foil sheet immediately after transfer to reduce the risk of contamination.
9. With too much input DNA, each droplet would have too many genomic DNA copies, which interferes with proper separation of the four distinct populations and thus accurate estimation of allelic concentrations. With too little input DNA, the number of genomic copies analyzed by the assay would not be enough to have high sensitivity. Therefore, the amount of genomic DNA input should not exceed 150 ng per ddPCR reaction.
10. For NHEJ quantification, only NHEJ single positives and the Empty (double-negative) droplets were used. For WT and for HDR quantification, all droplets are used.

## Acknowledgment

We thank Jennifer R. Berman, Samantha B. Cooper, Bin Zhang, and George A. Karlin-Neumann (Bio-Rad) for technical help and helpful discussions. This work was supported by the National Institutes of Health (U01-HL100406, U01-GM09614, R01-HL108677, U01-HL098179, U01-HL099997, P01-HL089707, and R01-HL060664 to B.R.C.); the UCSF Liver Center to B.R.C., the Bluefield Project to Cure Frontotemporal Dementia to B.R.C., the Uehara Memorial Foundation Research Fellowship to Y.M., and Gladstone-CIRM Fellowship to Y.M.

## References

1. Gaj T, Gersbach CA, Barbas CF III (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* 31(7):397–405. <https://doi.org/10.1016/j.tibtech.2013.04.004>
2. Certo MT, Ryu BY, Annis JE, Garibov M, Jarjour J, Rawlings DJ, Scharenberg AM (2011) Tracking genome engineering outcome at individual DNA breakpoints. *Nat Methods* 8 (8):671–676. <https://doi.org/10.1038/nmeth.1648>
3. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154 (6):1380–1389. <https://doi.org/10.1016/j.cell.2013.08.021>
4. Hendel A, Kildebeck EJ, Fine EJ, Clark JT, Punjya N, Sebastian V, Bao G, Porteus MH (2014) Quantifying genome-editing outcomes at endogenous loci with SMRT sequencing. *Cell Rep* 7(1):293–305. <https://doi.org/10.1016/j.celrep.2014.02.040>
5. Lin S, Staahl BT, Alla RK, Doudna JA (2014) Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *elife* 3:e04766. <https://doi.org/10.7554/elife.04766>
6. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC, Kitano TK, Hodel MR, Petersen JF, Wyatt PW, Steenblock ER, Shah PH, Bousse LJ, Troup CB, Mellen JC, Wittmann DK, Erndt NG, Cauley TH, Kochler RT, So AP, Dube S, Rose KA, Montesclaros L, Wang S, Stumbo DP, Hodges SP, Romine S, Milanovich FP, White HE, Regan JF, Karlin-Neumann GA, Hindson CM, Saxonov S, Colston BW (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem* 83(22):8604–8610. <https://doi.org/10.1021/ac202028g>
7. Miyaoka Y, Chan AH, Judge LM, Yoo J, Huang M, Nguyen TD, Lizarraga PP, So PL, Conklin BR (2014) Isolation of single-base genome-edited human iPS cells without antibiotic selection. *Nat Methods* 11(3):291–293. <https://doi.org/10.1038/nmeth.2840>
8. Miyaoka Y, Berman JR, Cooper SB, Mayerl SJ, Chan AH, Zhang B, Karlin-Neumann GA, Conklin BR (2016) Systematic quantification of HDR and NHEJ reveals effects of locus, nucleic acid, and cell type on genome-editing. *Sci Rep* Mar 31(6):23549. PMID: 27030102, PMC4814844
9. Bio-Rad Bulletin 6407, Droplet Digital Applications guide



# Chapter 21

## DNA Methylation Analysis Using Droplet Digital PCR

Ming Yu, Tai J. Heinzerling, and William M. Grady

### Abstract

Droplet digital (ddPCR) is a recent advance in PCR technology that enables the precise detection and absolute quantification of nucleic acid target sequences and that has a range of applications for both research and clinical diagnostic studies. Here, we discuss the parameters important in the design and performance of ddPCR for the detection and quantification of methylated DNA. We provide explicit instructions for conducting methylation specific ddPCR (aka MethylLight ddPCR). We also present an example that demonstrates the sensitivity and precision of the method for detecting methylated DNA in the promoter region of *mir342/EVL*, a potential DNA methylation biomarker for colon cancer risk. Common technical problems and troubleshooting for conducting successful MethylLight ddPCR assays are also discussed.

**Key words** Droplet digital PCR, MethylLight, DNA methylation, Colorectal cancer, Risk biomarkers, MethylLight ddPCR

---

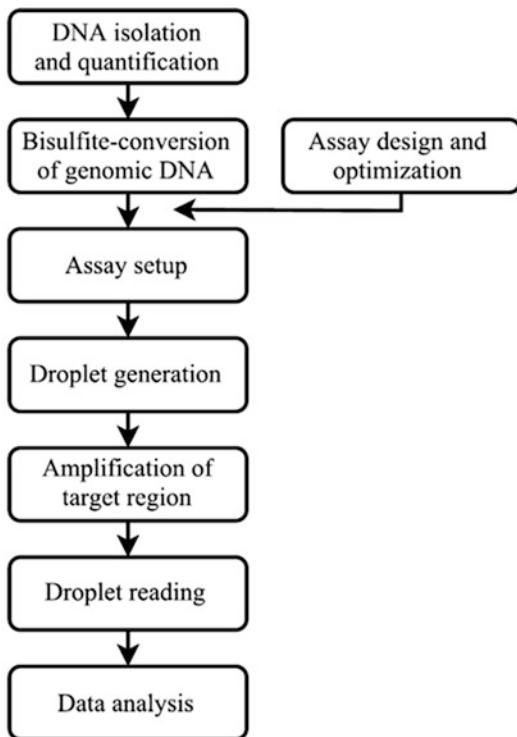
### 1 Introduction

DNA methylation is a commonly found epigenetic modification of mammalian DNA that results from the addition of a methyl group to the 5' carbon of a cytosine in a CpG dinucleotide sequence. The majority of the mammalian genome is heavily methylated, and DNA methylation has been shown to play an important role in development and disease by regulating gene expression in cells. Aberrant DNA methylation, which entails global hypomethylation throughout the genome and focal hypermethylation of CpG islands located in gene promoter regions and intergenic DNA, is a common epigenetic alteration in many human malignancies, including colorectal cancer (CRC). The study of DNA methylation alterations in CRC has provided the basis for many principles that have guided our understanding of epigenetic alterations in cancer in general. Intriguingly, DNA methylation alterations have been observed in normal colon tissue that is predisposed to develop polyps or cancer, a phenomenon called “field cancerization” [1–3], albeit at a much lower frequency than in diseased tissue

itself. Thus, these infrequently methylated DNA genes, for example *EVL* [4], have the potential to serve as biomarkers to identify individuals at high risk for developing CRC. It is likely that the field cancerization process is relevant for other cancers. MethylLight ddPCR is an ideal method for the study of field cancerization given its precision for detecting low levels of infrequently methylated alleles.

PCR-based methods have been applied extensively for the assessment of DNA methylation. The first generation of PCR for detecting methylated alleles was based on PCR amplification of DNA after digestion with methylation-sensitive restriction enzymes [5]. This was followed by a more flexible method called methylation-specific PCR (MSP), for which primers are designed specifically against DNA that has been treated with sodium bisulfite, which converts unmethylated cytosine to uracil through deamination [6]. In its initial iteration, the bisulfite-converted sequences of DNA in MSP are subjected to end-point analysis by UV transillumination after horizontal gel electrophoresis generating qualitative results. However, due to its qualitative nature, this original version of the MSP method proved suboptimal for applications that require quantitative assessment of the amount of methylated DNA present, which is often needed in epigenetic studies as well as in biomarker development. The development of MethylLight, a fluorescence-based quantitative PCR (TaqMan®) assay to measure DNA methylation, has considerably enhanced our ability to detect and quantify methylated alleles [7]. In MethylLight, the primers and/or fluorescent probes are designed against bisulfite-converted DNA sequence and quantitative information is obtained in real-time. MethylLight PCR typically is designed to detect methylated alleles and can quantitate the amount of the methylated allele in a relative fashion following correction for DNA loading and normalization to a methylated DNA standard. Despite its advantages over endpoint MSP, MethylLight is still susceptible to the presence of PCR inhibitors, has limited sensitivity for detecting rare methylated alleles in a background of unmethylated alleles, and can be skewed through the normalization process, which can lead to inconsistent results, especially in the detection of low levels of DNA methylation. Thus, conventional quantitative methylation specific PCR is often unable to reach the degree of precision and sensitivity needed to accurately detect low levels of methylated DNA and thus fully utilize DNA methylation as a field cancerization effect marker or cancer risk biomarker.

Here, we describe the development of a novel and highly sensitive assay for detecting methylated DNA based on droplet digital PCR, called the MethylLight ddPCR assay. We put emphasis on how to detect and perform absolute quantification of rare methylation events in low-concentration DNA samples with MethylLight ddPCR. An overview of the steps involved in a typical



**Fig. 1** Workflow diagram of the procedure for the MethyLight ddPCR, which is used for DNA methylation analysis

MethyLight ddPCR assay is shown in Fig. 1 and described in more detail in the text in this chapter. Our group and others have shown that this method allows the precise detection of infrequently methylated alleles that cannot be detected by conventional MethyLight, [8, 9].

## 2 Materials

The quality of DNA preparation from the biological samples of interest can impact DNA methylation analysis using ddPCR MethyLight. Acceptable sample sources include those that are acceptable for conventional PCR applications and include, but are not limited to, fresh frozen tissue, formalin-fixed, paraffin-embedded (FFPE) tissue, cell lines, and cell-free DNA in the plasma/serum. It is also important to be aware that genomic DNA will be subjected to bisulfite conversion under harsh chemical conditions, leading to fragmented DNA. Therefore, there is no need to do restriction digestion of DNA samples, as recommended for intact DNA. Although ddPCR is less sensitive to PCR inhibitors than

other technologies, we recommend thoroughly removing PCR inhibitors during DNA purification. Consider reducing the impact of PCR inhibitors on reactions by diluting samples 1:10. The recommended range of input DNA of the QX100/200 system is from 3 to 106060.6 copies/20 µL reaction. Assuming 3.3 pg/human haploid genome based on the Celera Genomics estimate), the amount of input DNA is estimated to range from 10 pg to 350 ng of human DNA/20 µL reaction. To estimate the number of copies/ng of DNA of a different organism, one needs to know the mass and the number of base pairs in the genome of interest. If the target copy number/genome is unknown, the optimal starting amount needs to be determined empirically using a dilution series of each sample at the relevant range for the anticipated experiments.

### **2.1 Genomic DNA (gDNA) Isolation**

1. We have routinely used DNA isolated from fresh frozen tissue using the DNeasy® Blood and Tissue Kit (QIAGEN, catalog #69504). If starting with FFPE tissues, we recommend using the QIAamp® DNA FFPE Tissue Kit (QIAGEN, catalog #56404) instead (*see Note 1*).
2. Water bath.
3. Xylene for deparaffinization of FFPE samples.
4. 100% ethanol.
5. Microcentrifuge tubes, 1.5 mL (RNase and DNase free).
6. Microcentrifuge.

### **2.2 gDNA Quantification (See Note 2)**

1. QuantiT™ PicoGreen® dsDNA Assay Kit (Life Technologies, catalog #p7589).
2. Falcon 96-well Black/Clear Flat-bottom Imaging Plate with Lid (Fisher Scientific, catalog #353219).
3. Molecular biology grade water.
4. Microcentrifuge tubes, 1.5 mL (RNase and DNase free).
5. Fluorescence microplate reader, e.g., Fluoroskan Ascent from ThermoLabsystems with Ascent software version 2.4.2.

### **2.3 Bisulfite Conversion of gDNA (See Note 3)**

1. Molecular biology grade water.
2. EZ DNA Methylation Kit (Zymo Research, catalog #D-5001).
3. Water bath or other device that can be used to incubate the samples at a constant temperature overnight.
4. Microcentrifuge tubes, 1.5 mL (DNase, RNase free).
5. Microcentrifuge.

### **2.4 Assay Components**

1. Control DNA: EpiTect Control DNA, 100% methylated (QIAGEN, catalog #59655) and EpiTect Control DNA, 100% unmethylated (QIAGEN, catalog #59665, *see Note 4*).

2. Software for converting DNA sequence to bisulfite-converted sequence, e.g., PyroMark Assay Design v2.0.1.15 (QIAGEN, catalog #9019077).
3. Software for designing primers and probes, e.g., Primer Express® Software v3.0.1 (Life Technologies, catalog #4363991).
4. Probes and forward and reverse primers for both the reference gene and gene of interest. To control for the total input of methylated and unmethylated DNA, we use a reference gene assay that is insensitive to CpG methylation status (referred to as the C-less assay) [10].
5. Droplet Digital PCR Supermix for Probes (Bio-Rad, catalog #1863010, *see Note 5*).

## **2.5 Bio-Rad ddPCR System Consumables**

1. Droplet Generation Oil for Probes (Bio-Rad, catalog #1863005).
2. DG8™ Cartridges for QX200™/QX100™ Droplet Generator (Bio-Rad, catalog #1864008). One is needed for every eight reactions.
3. DG8™ Gaskets for QX200™/QX100™ Droplet Generator (Bio-Rad, catalog #1863009). One is needed for every eight reactions.
4. ddPCR™ Droplet Reader Oil (Bio-Rad, catalog #1863004).
5. Pierceable Foil Heat Seals (Bio-Rad, catalog #1814040).
6. Microcentrifuge tubes, 1.5 mL (DNase, RNase free).
7. Semiskirted 96-well PCR plate with notches in at least two corners (e.g., Eppendorf, catalog #951020362).

## **2.6 Bio-Rad ddPCR System Equipment**

1. T100™ PCR Thermal Cycler (Bio-Rad, catalog #1861096).
2. QX200™ Droplet Generator (Bio-Rad, catalog #1864002).
3. PX1™ PCR Plate Sealer (Bio-Rad, catalog #1814000).
4. QX200™ Droplet Reader (Bio-Rad, catalog #1864003).
5. QuantaSoft™ Software, Regulatory Edition (Bio-Rad, catalog #1864011).

---

## **3 Methods**

Carry out all procedures at room temperature, unless otherwise noted, and in a room guaranteed free of potentially contaminating DNA, such as plasmids and amplicons. All steps after DNA isolation and during setup of the PCR (prior to placing the samples in the thermocycler) should be carried out in an environment kept completely free of potentially contaminating DNA, such as in a

PCR workstation (*see Note 6*). All instruments used in the PCR set-up steps should be decontaminated by exposure to UV light (254 nm) for at least 15 min prior to starting this phase of Methyl-Light ddPCR.

### **3.1 Genomic DNA (gDNA) Isolation and Quantification**

1. Isolate gDNA from tissue using a kit such as the DNeasy® Blood and Tissue Kit or the QIAamp® DNA FFPE Tissue Kit for FFPE tissues (*see Note 1*). Perform gDNA isolation according to the manufacturer's instructions for special considerations when using FFPE tissues, *see Note 7*.
2. Measure gDNA concentration using the Quant-iT™ PicoGreen® dsDNA Assay Kit according to manufacturer's instructions scaled down to a 96-well format: Use a 96-well imaging plate with black walls and a clear flat bottom. Add 100 µL of DNA solution and 100 µL of Quant-iT PicoGreen reagent working solution to each well for measurement (*see Note 2*).

### **3.2 Bisulfite Conversion of gDNA (See Note 3)**

1. Perform bisulfite conversion of sample DNA using a kit such as the EZ DNA Methylation Kit from Zymo Research (D5002) following the manufacturer's instructions with the following optimizations.
2. Consistent with the manufacturer's instructions in the EZ DNA Methylation Kit, we recommend using less than 1 µg DNA per column since too much DNA may result in incomplete conversion.
3. The overnight incubation (**step 4** of the kit) may be done in a water bath, thermocycler, or heat block. Any of these are acceptable as long as the samples are in a DNA clean environment and protected from light. It is also important to avoid excessive shaking of the samples during this process.
4. After the final wash step (**step 11**), empty the collection tube and spin-dry the column for an additional 1 min in order to ensure that there is no carryover of wash buffer into the elute.
5. To obtain the highest yield, repeat the elution step with both rounds of eluate going into the same tube (10 µL eluate per elution step, total volume = 20 µL).
6. Bisulfite converted-DNA should be used as soon as possible after conversion, or can be stored at -70 °C for later use. Freeze the samples immediately after conversion if they are to be stored. Repeated freezing and thawing is strongly discouraged as greater than three freeze-thaw cycles can lead to DNA fragmentation that impairs PCR DNA amplification. Therefore we recommend creating aliquots of samples before freezing.

### 3.3 Primer and Probe Design (See Note 8)

The primer/probe design follows the principles of MethyLight assays in general, which has been discussed in detail previously, with modifications noted below [11].

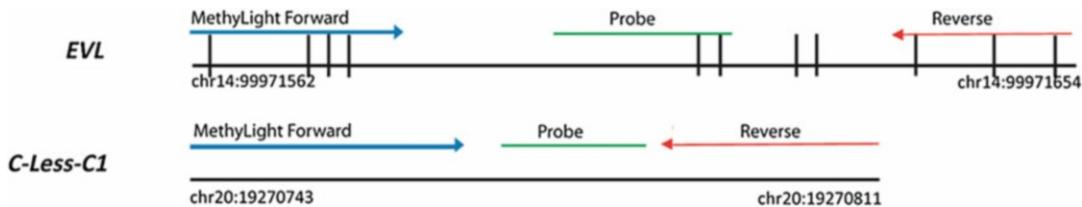
1. First virtually convert both the forward and reverse strand of a 500 bp sequence of DNA surrounding the CpG of interest to its bisulfite-converted sequence using PyroMark Assay Design v2.0 (QIAGEN).
2. Next, use ABI Primer Express software Version 3.0.1 (Applied Biosystems, Life Technologies) to design the primers and probe used in a MethyLight ddPCR assay (An example of the use of ABI Primer Express designed primers and probes for a methylated *EVL* assay (NP\_057421) is described later and in Yu et al. [9].) After virtual bisulfite conversion of the sequence, input either the forward or reverse strand into Primer Express set to the TaqMan® MGB Quantification setting.
3. Once the sequence is in Primer Express, follow the instructions for use of the software to select primers and probes for the locus of interest.
4. If an optimal set of primers and probes cannot be generated by the software, then modify the sequence of interest repeatedly by shortening or extending the sequence until a primer or probe covers the CpG of interest.
5. In addition, if an acceptable primer and probe set is not found despite modifying the target sequence, try using more lenient parameters (e.g., alter melting temperature parameters) or designing the assay for the complementary strand.
6. The selected probes are synthesized with a 5' fluorescent emitter and a 3' nonfluorescent quencher with a minor groove binding (MGB) domain. We design probes with 5' fluorescent emitters that can be used in a duplex assay. So we typically design the probes for the gene of interest with FAM and the reference gene with VIC or HEX, which allows us to simultaneously detect and quantify the methylated target gene (e.g., *EVL* labeled with FAM) and the reference gene (e.g., *C-LESS-C1* [10] labeled with VIC).

The reference gene selected will be amplified regardless of its methylation status and therefore is used to determine the total amount of amplifiable DNA. Other investigators have used beta-actin (ACTB) to normalize for input DNA, which can also work as a reference gene but is less well studied compared to *C-LESS-C1* (Belshaw et al. [3]). A complete list of the MethyLight primer and probe sequences for the *EVL* and *C-Less-C1* MethyLight ddPCR assays are provided in Table 1. The location of CpGs in the MethyLight primers and probes and the amplicons for methylated *EVL* and *C-LESS-C1* assays are provided in Fig. 2.

**Table 1**  
**Primer and probe sequences used in MethyLight ddPCR assays**

Assay	Primer/probe	Primer/probe sequence (5' to 3')
EVL	Forward	AACGACTCCGAATCCTCGAA
	Reverse	GCGAATAGTAACGCGGTATT
	Probe	FAM-CGCGAACTAATCTAACCA-MGBNFQ <sup>a</sup>
C-less-C1	Forward	TTGTATGTATGTGAGTGTGGGAGAGAGA
	Reverse	TTTCTCCACCCCTCTCTCC
	Probe	VIC-CCTCCCCCTCTAACTCTAT-MGBNFQ <sup>a</sup>

<sup>a</sup>MGBNFQ refers to a Minor Groove Binding nonfluorescent quencher in the 3' terminus of the probe



**Fig. 2** Relative location of the CpG dinucleotides in the MethyLight primers and probes for the assay to detect methylated *EVL* and for the C-LESS-C1 assays. The forward primers, probe and reverse primers are indicated by the blue arrows, green lines and red arrows, respectively. The small vertical lines indicate the CpG sites. The DNA strand shown is the top strand

### 3.4 Plate Lay-Out Design

1. Design the experimental plate layout using a 96-well plate in columns (see Note 9).
2. Always include a NTC, a positive control using 100% methylated Epitect DNA (Pos.), and a negative control using 100% unmethylated EpiTect DNA (Neg.)

### 3.5 Droplet Generation and Transfer

1. Thaw all reagents and DNA on ice, and equilibrate to room temperature for 3 min before generating droplets. Protect the probe from exposure to light.
2. The “Master Mix” (MM) is a mixture of Supermix, primers, probes, and water, in correct proportions, as described in Table 2 below, which will be added to each sample. Calculate the volumes of MM required based on 25 µL total volume per sample (20 µL MM plus 5 µL DNA). The primer and probes are used at final concentrations of 900 nmol/L and ~250 nmol/L, respectively according to manufacturer’s recommendations. See Table 2 for an example of Master Mix calculations. Reagents should be well mixed by pipetting or gentle vortexing.

**Table 2**  
**Example master mix calculations**

Component	Starting	Final	1×
Water			5
2× ddPCR supermix for probes	2×	1×	12.5
20× FAM <sup>a</sup>	20×	1×	1.25
20× VIC <sup>a</sup>	20×	1×	1.25
Bisulfite-converted DNA	2	ng/ $\mu$ L	5
Final volume in $\mu$ L <sup>b</sup>			25
Final ng/ $\mu$ L in ddPCR			0.4

<sup>a</sup>Recommend setting up 25  $\mu$ L/reaction and loading 20  $\mu$ L droplets in DG cartridge

<sup>b</sup>20× primer-probe mix: 18 uM PCR primers (each), 5 uM probe

3. For each sample and control, prepare one “reaction tube”: a 1.5 mL microcentrifuge tube containing all of the Master Mix and DNA for the replicates, with 20  $\mu$ L MM and 5  $\mu$ L DNA for each replicate. Multiply MM and DNA volumes by the number of replicates (e.g., for four replicates, mix 80  $\mu$ L Master Mix and 20  $\mu$ L DNA in a tube). Store reaction tubes in the dark at room temperature and start loading them one at a time onto the generator cartridge immediately after the reagents are combined.
4. Open the droplet generator cartridge holder by pressing latches on each side. Slide the disposable cartridge into the right side of the holder with the notch on the upper left corner of the cartridge. Make sure it is in place before pressing the two halves of the holder together.
5. Vortex the reaction tube at a moderate speed for at least 10 s and transfer 20  $\mu$ L of the reaction tube mixture into each well in the middle row of the cartridge (*see Note 10*).
6. Pour droplet generator oil into a pipetting reservoir and use a multichannel pipette to load 70  $\mu$ L of oil into each well on the bottom row of the cartridge. All eight wells must be filled or the droplets will not be generated properly.
7. Hook a DG8™ Gasket for QX200™/QX100™ Droplet Generator over the cartridge holder.
8. Open the QX200 Droplet Generator by pressing the green button on the top of the Generator and place the cartridge holder in the instrument. Verify that the two left indicator lights on the generator are green.

9. The run will begin automatically. When all three indicator lights are green, the run is finished and the door may be opened. Remove the cartridge holder from the instrument and remove the gasket from the holder. Discard the gasket.
10. Using an electronic multichannel pipette set to a volume of 40 µL and speed of one (lowest), gently transfer the droplets from the upper wells of the cartridge to one column of a 96-well plate (*see Note 11*).
11. Remove the cartridge from holder and discard.

### **3.6 Sealing the Plate**

1. Configure the PX1 PCR plate sealer to seal for 5 s at 180 °C.
2. Open the PX1 PCR plate sealer by pressing the arrow on the instrument's screen. Position your plate on the tray.
3. Position a pierceable foil seal over the wells of the plate with the red stripe of the seal facing upward. Place the metal frame included with the instrument over the PCR plate.
4. Close the door of the sealer and once the machine is up to temperature, press "Seal."
5. The PX1 will open automatically when it is finished sealing. Turn the PCR plate horizontally 180° and repeat the steps described above.
6. Thermal cycle the droplets within 30 min of pipetting the reaction mixtures into the plate.

### **3.7 PCR**

1. Run the plate in a thermocycler with the cycling conditions recommended for the Supermix used. For cycling conditions to use with ddPCR Supermix for Probes (Bio-Rad, catalog #1863010), *see Table 3*.
2. After thermocycling, the plate should be read using the QX200 droplet reader immediately or kept at 4 °C for no more than 1 day.

**Table 3**  
**Thermalcycling conditions for ddPCR Supermix for Probes<sup>a</sup>**

Cycling step	Temperature °C	Time	Ramp rate	Number of cycles
Enzyme activation	95	10 min		1
Denaturation	94	30 s	2 °C/s	45
Annealing/extension	60 <sup>b</sup>	1 min	2 °C/s	
Enzyme deactivation	98	10 min		1
Hold	4	Infinite		1

<sup>a</sup>Use a heated lid set to 105 °C and sample volume set to 40 µL

<sup>b</sup>Annealing temperatures vary for each primer/probe set. Test for optimal temperature

### 3.8 Droplet Reading

1. Turn on the QX200 droplet reader and allow it to warm up for 30 min.
2. Verify that the left two lights are solid green. If the left light is not solid, there is an issue with the power. If the second light is not on then either the oil bottle is too low or the waste bottle is too full. Correct these issues before proceeding with the plate reading.
3. Open the plate holder by flipping the two black release tabs into the up position and removing the cover. Place the plate on the holder with well A1 in the upper left, replace the cover, and flip the black tabs back down.
4. Open the reader by pressing the green button on top of the door, place the plate holder in the instrument, and close the door again using the same button. The plate indicator light should now be solid green.
5. Connect the reader to your computer via a USB cable and open the QuantaSoft software. Select Setup and define the experiment design. Then select Run (see Note 13).
6. While the run is in progress, the right-most indicator light flashes green. When the run is finished, all indicator lights should be solid green. At this point, open the door and remove the plate holder. Remove the PCR plate from the holder and discard.

### 3.9 Interpreting 2D Plot Results (See Table 4)

After the run is complete, we use QuantaSoft software to analyze the data in each well. If the plate was set up for absolute quantification (ABS) analysis. QuantaSoft sets the threshold automatically and determines concentration in the data tables in the analysis mode of the software. However, one should inspect each well and manually adjust the threshold if necessary to ensure correct designation of the droplets in each well.

Figure 3 shows an example of a duplex MethyLight ddPCR experiment in which the target gene *EVL* and the reference gene *C-LESS-C1* have been PCR amplified and viewed in a 2D plot of droplet fluorescence. For detecting infrequently methylated alleles in clinical samples, each sample is run in eight replicate wells and the droplet counts (positive and negative) from all replicated wells are combined to yield a “merged” well. The concentration and Poisson confidence intervals for each “merged” well are computed using the QuantaSoft software version 1.4.0.99 (Bio-Rad, Hercules, CA). Troubleshooting advice can be found in Table 4.

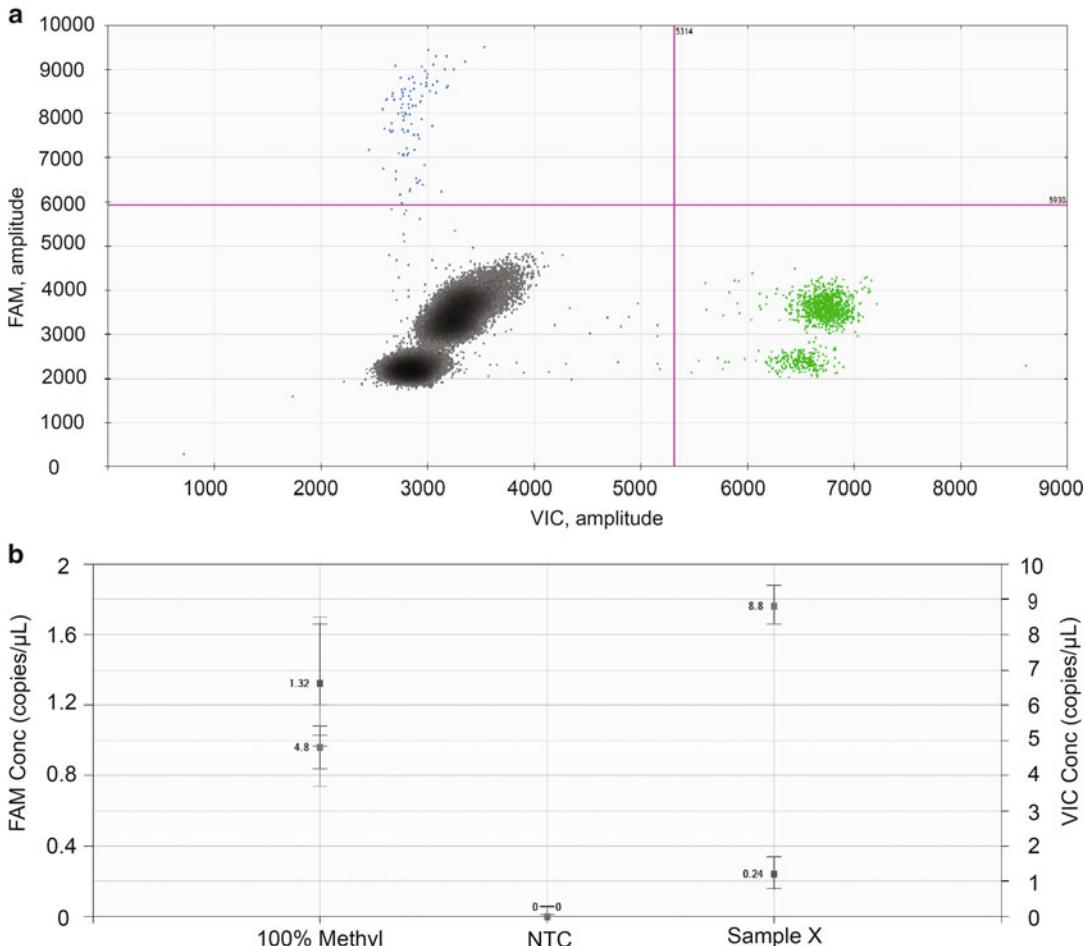
**Table 4**  
**Troubleshooting table**

Problem	Possible reason	Possible solution
No positive droplet counts from positive control samples for the reference gene	The ddPCR reaction mix has not been assembled correctly	Check the volumes and concentrations of each component in the PCR reaction mix
No positive droplet counts of the reference gene from the samples, but the positive control works properly	The amount of input DNA is too low	Increase the amount of the input DNA
The DNA concentration measurements by ddPCR are considerably lower than spectroscopic measurements	Bisulfite-converted DNA is partially degraded as ddPCR gives a concentration estimate of intact DNA targets Suboptimal assay design/efficiency	Use recommended methods to prepare, handle and store bisulfite-converted DNA for methylation studies Use the recommended primer and probe concentration. Optimize ddPCR MethyLight assay thermocycler conditions by running a temperature gradient experiment to determine an optimal annealing temperature
Only the reference gene gives positive droplet counts. No signal from a newly designed assay for the gene of interest	Suboptimal assay design for the gene of interest The assay conditions are not optimal  The ddPCR reaction mix is not assembled correctly  The probe/primers sequences are not correct. This can happen if an incorrect sequence is used for synthesizing the primers or probe	Redesign the assay  Optimize assay conditions by performing an annealing/extension temperature gradient Recheck the volumes and concentrations of the components in the PCR reaction mix  Set up a conventional MethyLight assay to check if the probe/primers will amplify the DNA. Reorder probe/primers if no PCR product generated
Duplexing two assays together does not give any positive droplet counts	Duplex assays are interfering with each other	Test the two assays separately in singleplex assays
Positive droplet counts in NTC	Contamination	Decontaminate the PCR workstation, PCR preparation area, and equipment. Practice good sample handling to avoid contamination. Repeat the experiment with new PCR grade water and reagents.

(continued)

**Table 4**  
**(continued)**

Problem	Possible reason	Possible solution
		Separate the post-PCR area from the PCR preparation room
Significant number of mid-level amplitude droplets (“Heavy rains” shown on the fluorescence amplitude plot) in DNA samples. Poor separation of fluorescence intensity between negative and positive droplets	The presence of PCR inhibitors in DNA samples	Dilute DNA samples further with PCR grade water
The number of total droplet counts per well is too low (<10,000)	The ddPCR reaction mix is not assembled correctly	Use the recommended concentration of primers (900 nM), probe (250 nM), and 1× ddPCR Supermix for probe
	Suboptimal droplet generation	Use a P-20 pipette and slowly load the DG8 cartridge with 20 µL of samples into the bottom of the sample well. Then load 70 µL of oil to the oil well. Begin droplet generation immediately after loading
	Poor droplet handling	Use the recommended electronic multichannel P-50 pipette and settings to ensure full-volume transfer of droplets into 96-well plate. Make sure the angle of the pipette tip near the bottom of the well to avoid shearing the droplet when dispensing. Handle droplets in a gentle and consistent way
	Undersealed plates result in oil evaporation and affect droplet quality	Properly seal the 96-well plate
Wide variation in technical replicates	Poorly mixed reaction mix	When creating technical replicates, make sure the reaction mixture is thoroughly mixed
	Poor well-to-well temperature uniformity of the PCR thermal cycler	Verify the thermal uniformity and accuracy of the PCR thermal cycler



**Fig. 3** Quantification of methylated *EVL* in colon mucosa biopsy samples by a duplex MethylLight ddPCR assay. Genomic DNA samples were bisulfite converted and analyzed by MethylLight ddPCR. Each sample (8 ng or 2424.2 haploid genome equivalents of total input DNA) was partitioned into an average of 15,000 droplets per well and replicated in 8 wells. The droplet counts (positive and negative) from all replicated wells were combined to yield a “merged” well. The concentration and Poisson confidence intervals for each “merged” well were computed using the QuantaSoft software version 1.4.0.99 (Bio-Rad, Hercules, CA). (a) 2D amplitude plot of the duplex assay. For each fluorophore, a threshold was manually set to designate positive events. Blue: methylated *EVL*-positive droplets (FAM); Green: C-Less-C1 positive droplets (VIC); Black: negative droplets. (b) Measured concentration (copies/μL) of the methylated *EVL* (blue) in the positive control (EpiTect 100% methylated control DNA), NTC and a representative clinical sample. The total input DNA was estimated based on a reference gene C-LESS-C1 assay (green). Error bars indicate the Poisson 95% confidence intervals for each measurement. NTC: no template control

---

## 4 Special Considerations for Detection and Absolute Quantification of Rare Methylated Alleles with ddPCR MethyLight

Rare methylation events occur in the context of field cancerization effect, when one methylated allele of a candidate biomarker gene exists in a background of (over 10,000) unmethylated alleles. DDPCR MethyLight enables detection and absolute quantification of rare methylation events at a level of sensitivity and precision beyond the capabilities of conventional MethyLight [9]. This section will discuss proper experimental setup to achieve a desirable level of quantification (LOQ) and accurate quantification of rare methylated alleles.

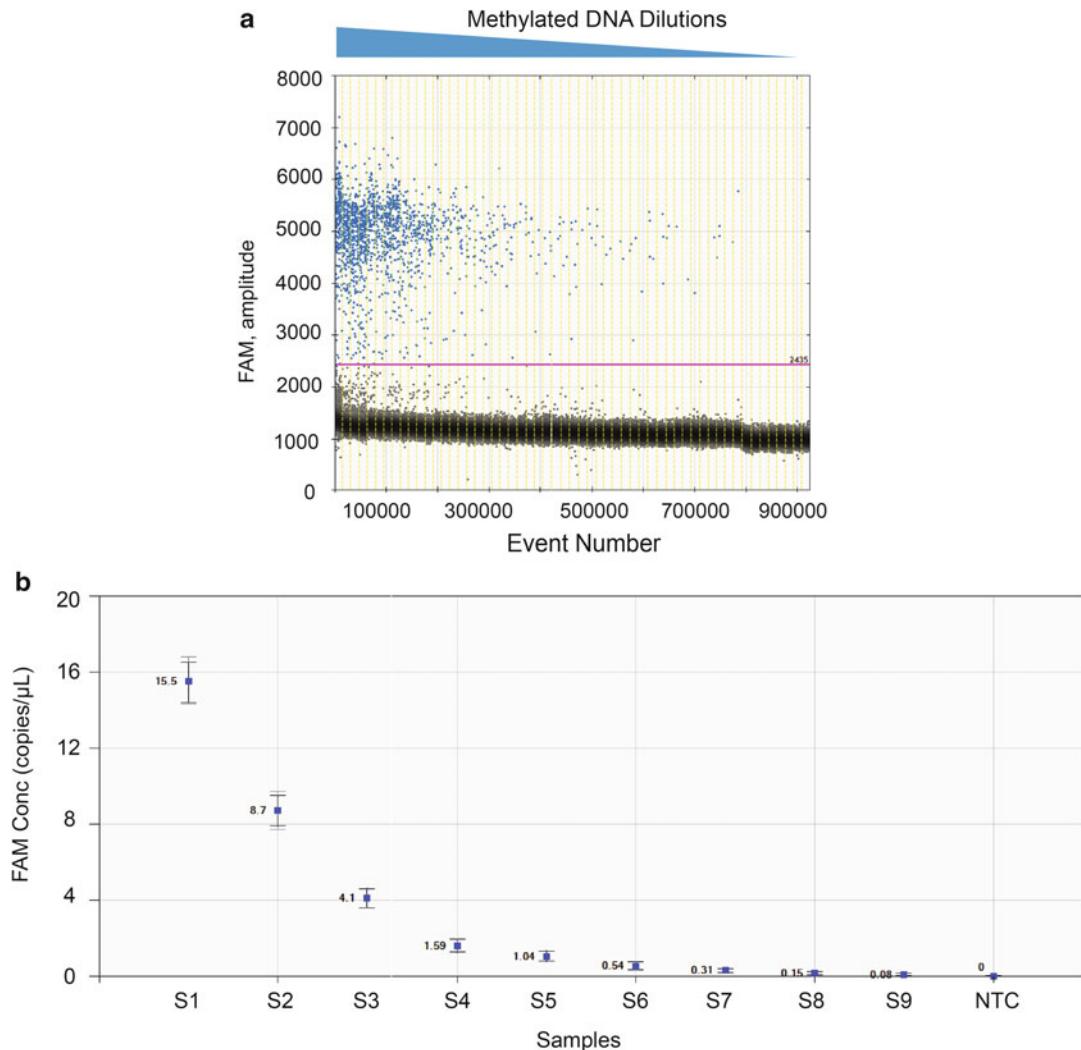
### 4.1 Sample Preparation

DNA extracted from a biological sample will have variable integrity and concentration, depending on the sample source. For example, FFPE tissues, although used routinely in cancer research, provide DNA that is highly fragmented and chemically cross-linked, which may negatively impact the performance of a MethyLight ddPCR assay. This should be kept in mind if poor quality results are generated from the test samples despite having good results from the positive control samples.

Most methods in DNA methylation studies use bisulfite conversion, which results in fragmented genomic DNA. One needs to take into account the actual amount of amplifiable DNA in these samples when calculating the amount of starting material needed. Furthermore, bisulfite converted DNA may produce a “rain” profile and a shift of baseline in a 1D amplitude plot (Fig. 4a). It is possible to account for this artifact by simply setting the threshold manually to obtain a good separation of positive droplets from negative background and achieve accurate absolute quantification (Figs. 3a and 4a). In addition, incomplete bisulfite-converted DNA will result in misleading data and should be considered when an unexpected amount of methylated DNA is detected or low PCR amplification occurs.

### 4.2 Testing an Assay

For development of methylation assays for assaying for a field cancerization effect, one important consideration is the limit of quantification (LOQ), which is the minimum concentration of the methylated allele in a sample that can be reliably distinguished from a background of a much higher proportion of unmethylated alleles, within an acceptable (predetermined) coefficient of variation (CV, measurement error). We use a multiplex assay, that includes an assay for a reference gene (e.g., C-LESS-C1) to quantify the amount of input DNA (over 99.9% of which are unmethylated), and an assay for the methylated gene of interest (e.g., *EVL*) to quantify the amount of methylated *EVL* allele. LOQ is typically quoted as a ratio or a percentage: for example, 1 methylated allele in



**Fig. 4** Analysis of limit of quantification (LOQ) for methylated *EVL* by MethyLight ddPCR. We determine the LOQ by serial dilution of the 100% methylated EpiTect Methyl control DNA into a solution containing 100% unmethylated EpiTect Unmethyl control DNA (8 ng of total input DNA in each well). Each sample is partitioned into an average of 15,000 droplets per well and replicated in four wells. The droplet counts (see Subheading 4 with regard to protocol for selection of positive and negative droplets) from all four replicates are combined for the final analysis. The concentration and Poisson confidence intervals for each “merged” well are computed using the QuantaSoft software version 1.4.0.99 (Bio-Rad, Hercules, CA). (a) 1D plot shows the separation between positive droplet (blue) and negative droplets (black). An event with fluorescence amplitude value  $>2435$  was considered a methylation-positive event (red line). Note the “rain” profile and a shift in baseline in the 1D amplitude plot, possibly due to the bisulfite conversion treatment. (b) Representative example of results of ddPCR of serial DNA dilution series. The concentration (copies/ $\mu$ L) of methylated *EVL* in serially diluted samples (8 ng of total input DNA per well) are shown: S1 = 80%, S2 = 40%, S3 = 20%, S4 = 10%, S5 = 5%, S6 = 2.5%, S7 = 1.25%, S8 = 0.625%, and S9 = 0.313%. The error bars represent Poisson 95% confidence intervals

1000 of unmethylated alleles, or 0.1%. In ddPCR, the LOQ is primarily scored based on the number of unmethylated alleles that are detected in the sample using the reference gene assay. For a given assay, we determine the LOQ by serially diluting the 100% methylated EpiTect Methyl control DNA with 100% unmethylated EpiTect Unmethyl control DNA in 8 ng of total DNA in each well. Figure 4 shows analysis of LOQ for methylated EVL by MethyLight ddPCR.

#### **4.3 The Amount of Input DNA**

One consideration for increasing the detection capability is to increase the amount of input DNA. For example, to reach a LOQ of 0.001%, statistically at least 300,000 human haploid genome in a well, or 1 µg of human DNA, must be screened within 20% CV. If enough sample DNA is available, one can use about 1 µg of bisulfite-converted DNA in the 20 µL reaction to effectively detect more methylated alleles at a LOQ of 0.001%. It is important to keep in my mind that the source of the DNA, such as a tissue biopsy may provide very limited amount of DNA. Thus, a practical LOQ of 0.1% is expected if the total input DNA is less than 10 ng per well.

#### **4.4 Experimental Replications**

Random sample partition is the fundamental feature of ddPCR technology. It also has an important impact on improving an assay's LOQ by effectively diluting away the background DNA that can interfere with detection of methylated alleles, resulting in greater relative abundance of a rare methylated template in a given droplet. Running the same samples in multiple wells will further increase the number of positive events that can be detected by ddPCR, thus improving the LOQ for low input DNA samples. To achieve a lower LOQ than 1 in 25,000 or 0.04% using human DNA, we recommend replicates of four or higher and creating a merged well for analysis for detecting rare methylated alleles.

#### **4.5 Proper Controls in Experimental Setup**

A proper experimental setup for detecting rare methylated alleles includes positive control wells (100% methylated control DNA), negative control wells (100% unmethylated control DNA), and non-template-control (NTC) wells. The number of positive droplets in the negative control wells should be zero, which indicates optimal specificity of the assay. In addition, NTC wells should also produce zero positive droplets, which reflects optimal laboratory practices. Refer to Table 4 if positive droplets are detected in NTC. The false-positive rate must be considered when designing assays to achieve accurate quantification.

#### **4.6 Absolute Quantification Data Analysis**

We use the absolute quantification (ABS) experiment design to quantify the concentration of the methylated alleles in a sample. QuantaSoft software counts the number of positive and negative droplets for each fluorophore in each sample, and fits the fraction of

positive droplets to a Poisson statistics to determine the concentration in copies/ $\mu$ L of the final ddPCR reaction. We then back-calculate the original concentration in the starting material. This type of ABS experiment does not require a standard curve as needed in conventional MethylLight assays, thus simplifying the experimental planning and reducing the cost.

To analyze replications, select the merged button in QuantaSoft software to combine counts from all the wells with the same sample name and analyze these data together in multiple wells.

#### 4.7 Types of Errors

For experiments with replicates, QuantaSoft software calculates two types of errors: a theoretical replicate error (technical errors, or Poisson error) and a standard error of the means (the total error). It is recommended that total errors are reported for each measurement. For samples with DNA concentration within the range of dynamic range of ddPCR (3–106060.6 copies of target molecules/well), the total error and the Poisson error should be nearly identical.

However, for quantifying rare methylated alleles that are in the extreme low end of the concentration range, both the Poisson error and the total errors will tend to be larger. In addition, errors introduced by pipetting during dilution or aliquoting are not negligible. Improving pipetting accuracy is helpful to reduce the total error.

---

## 5 Notes

1. The methods described for isolating and quantifying genomic DNA are based on protocols we have established in our laboratory and have yielded excellent results over many years. However, genomic DNA can be isolated from FFPE archived tissue samples, fresh frozen tissues, blood, or stool samples by a variety of methods, and other established methods that have worked well for a laboratory should work well for this protocol [12]. It is essential to include a proteinase K digestion step in the DNA extraction step to eliminate chromatin structures that might interfere with DNA dissociation, which is required for the bisulfite treatment to convert the unmethylated cytosines to uracil. For DNA extraction from larger FFPE samples or fresh frozen samples, we prefer the DNeasy® Blood and Tissue Kit (QIAGEN, catalog #69504). For isolating DNA from FFPE tissues, we have used QIAamp® DNA FFPE Tissue Kit (QIAGEN, catalog #56404), which is recommended for up to eight FFPE tissue sections, each with a thickness of up to 10  $\mu$ m and a surface area of up to 250 mm<sup>2</sup> in a single preparation. For DNA extraction from larger FFPE samples or fresh frozen

samples, we use the DNeasy® Blood and Tissue Kit (QIAGEN, catalog #69504).

2. Other established methods for DNA quantification that have worked well for a laboratory should work well for this protocol. However, we recommend using fluorometry based methods to measure the DNA concentration (e.g., Quant-iT PicoGreen dsDNA Assay Kit) rather than by absorbance using spectrophotometry to avoid interference from other biomolecules and dust particles and to obtain the most accurate estimates of DNA concentration from FFPE tissue sources, which often yields highly fragmented DNA.
3. For the majority of DNA methylation studies currently being performed, the genomic DNA is subjected to a chemical modification induced by sodium bisulfite, which deaminates unmethylated cytosine, followed by conversion to uracil and ultimately to thymine after PCR amplification. Methylated cytosine remains intact during this process. After bisulfite conversion, the two strands of DNA are no longer complementary to each other, which should be kept in mind during the design of primers and probes as either strand of DNA can be used for primer and probe selection. The resulting differential sequence information between methylated and unmethylated alleles is employed to design methylation-specific primers and probes. Therefore, bisulfite conversion is a critical step for DNA methylation studies and bisulfite conversion efficiency of >95% is necessary. It is essential that a robust sodium bisulfite conversion method be used which is dependent on the starting materials. There is published literature on this topic that can be used to guide the selection of a suitable method [13].
4. Control DNA: The EpiTect 100% methylated Control DNA from QIAGEN are provided already bisulfite-converted and at a concentration of 10 ng/mL. Store aliquots at –80 °C until use. Repeated freezing and thawing is strongly discouraged as bisulfite-converted DNA is single stranded and less stable than double stranded DNA. Other commercial sources such as CpGenome Human Methylated & Unmethylated DNA standard set (Millipore, catalog #S8001) should work just as well.
5. ddPCR Supermix for Probes: Store at –20 °C until use. Once thawed, the Supermix can be stored at 4 °C for up to 2 weeks. Repeated freezing and thawing is not recommended, but we have not observed a decrement in performance of the assays with up to two freeze-thaw cycles.
6. For an environment completely free of potentially contaminating DNA, we have had good results using the dead air Air-Clean600 Workstation from Air Clean Systems.

7. DNA can be isolated from FFPE tissues in the form of tissue cores or sections. In either case, we use the QIAamp® DNA FFPE Tissue Kit (QIAGEN, catalog #56404) with only slight modifications: Cores of tissue will require longer incubations in xylene, as well as a longer digestion period with a greater volume of Proteinase K solution (up to 4 days). For sections of tissues on slides, we add a rehydration step after deparaffinization by incubating the slides in a series of baths containing decreasing concentrations of ethanol. The tissue is then removed from the slide using a clean (i.e. DNA and RNA free, RNase free) razor blade and placed into a microcentrifuge tube where it is then digested by Proteinase K as normal. We prefer to perform the final elution in two steps in order to obtain the greatest yield possible, as FFPE samples are often quite small.
8. Primer and probe design: Primers and probes that work on conventional MethyLight should always work for MethyLight ddPCR, but may require further optimization. A minor-groove binding (MGB) domain on the quencher increases the melting temperature of the probe, thus allowing for shorter, more specific probes. The 3' emitters must be FAM and VIC/HEX because these are the only emitters detected by the QX200 system droplet reader.
9. A 96-well plate is mandatory for use with the droplet reader as opposed to PCR strip tubes or a 384-well plate. Designing your plate in columns allows for easier transfer of samples from the DG8 cartridge to the plate.
10. Droplet generation: One of the most crucial steps is the droplet generation and transfer, during which a 20 µL PCR reaction is randomly partitioned into on average 15,000 uniformly sized droplets, which enable sensitive detection and precise quantification. Handling the droplets in a gentle and consistent way is the key to an optimal droplet count. Reaction mixes should be well mixed prior to dispensing and the DG8 cartridge should be preloaded in the DG8 cartridge holder. When loading samples into the cartridge, the pipette tip should be held at a ~15° angle (from the vertical) and should rest on the ridge near the bottom of the well while slowly dispensing to avoid introduction of bubbles (which can disrupt droplet formation). Be sure to fill all eight wells with reaction mixes; it is helpful to have a second cartridge holder so that one can start preparing the next set while the first set is running in the generator.
11. Transfer of droplets: After generating droplets in the DG8 cartridge, the droplet solution should look cloudy. We strongly recommend transferring the droplets to a 96-well PCR plate using an electronic multichannel pipette set at a very slow speed. Gently dispense the droplets on the wall of the plate

wells near the bottom of the well. We also recommend designing the experiment plate layout on a 96-well plate in columns for ease of transfer from the eight-well droplet generator cartridges.

## Acknowledgments

This work was supported by National Cancer Institute (NCI) RO1CA194663, P30CA15704, UO1CA152756, U54CA143862, and P01CA077852 (WMG); Burroughs Wellcome Fund Translational Research Award for Clinician Scientist (WMG); NIH 2T32DK007742-16 (MY); the Lattner Foundation (WMG), R.A.C.E. Charities (WMG).

## References

1. Slaughter DP, Southwick HW, Smejkal W (1953) Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer* 6:963–968
2. Shen L, Kondo Y, Rosner GL, Xiao L, Hernandez NS, Vilaythong J, Houlihan PS, Krouse RS, Prasad AR, Einspahr JG et al (2005) MGMT promoter methylation and field defect in sporadic colorectal cancer. *J Natl Cancer Inst* 97:1330–1338
3. Belshaw NJ, Pal N, Tapp HS, Dainty JR, Lewis MP, Williams MR, Lund EK, Johnson IT (2010) Patterns of DNA methylation in individual colonic crypts reveal aging and cancer-related field defects in the morphologically normal mucosa. *Carcinogenesis* 31:1158–1163
4. Grady WM, Parkin RK, Mitchell PS, Lee JH, Kim YH, Tsuchiya KD, Washington MK, Paraskeva C, Willson JK, Kaz AM et al (2008) Epigenetic silencing of the intronic microRNA hsa-miR-342 and its host gene EVL in colorectal cancer. *Oncogene* 27:3880–3888
5. Zuccotti M, Grant M, Monk M (1993) Polymerase chain reaction for the detection of methylation of a specific CpG site in the G6pd gene of mouse embryos. *Methods Enzymol* 225:557–567
6. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB (1996) Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc Natl Acad Sci U S A* 93:9821–9826
7. Eads CA, Lord RV, Wickramasinghe K, Long TI, Kurumboor SK, Bernstein L, Peters JH, DeMeester SR, DeMeester TR, Skinner KA et al (2001) Epigenetic patterns in the progression of esophageal adenocarcinoma. *Cancer Res* 61:3410–3418
8. Hayashi M, Guerrero-Preston R, Sidransky D, Koch WM (2015) Paired box 5 methylation detection by droplet digital PCR for ultra-sensitive deep surgical margins analysis of head and neck squamous cell carcinoma. *Cancer Prev Res* 8:1017–1026
9. Yu M, Carter KT, Makar KW, Vickers K, Ulrich CM, Schoen RE, Brenner D, Markowitz SD, Grady WM (2015) MethyLight droplet digital PCR for detection and absolute quantification of infrequently methylated alleles. *Epigenetics* 10:803–809
10. Weisenberger DJ, Trinh BN, Campan M, Sharma S, Long TI, Ananthnarayan S, Liang G, Esteva FJ, Hortobagyi GN, McCormick F et al (2008) DNA methylation analysis by digital bisulfite genomic sequencing and digital MethyLight. *Nucleic Acids Res* 36:4689–4698
11. Trinh BN, Long TI, Laird PW (2001) DNA methylation analysis by MethyLight technology. *Methods* 25:456–462
12. Potluri K, Mahas A, Kent MN, Naik S, Markey M (2015) Genomic DNA extraction methods using formalin-fixed paraffin-embedded tissue. *Anal Biochem* 486:17–23
13. Holmes EE, Jung M, Meller S, Leisse A, Sailer V, Zech J, Mengdehl M, Garbe LA, Uhl B, Kristiansen G et al (2014) Performance evaluation of kits for bisulfite-conversion of DNA from tissues, cell lines, FFPE tissues, aspirates, lavages, effusions, plasma, serum, and urine. *PLoS One* 9:e93933

# **Part V**

## **Gene Expression and RNA Quantification**



# Chapter 22

## Simultaneous Quantification of Multiple Alternatively Spliced mRNA Transcripts Using Droplet Digital PCR

Bing Sun and Yun-Ling Zheng

### Abstract

Currently there is no sensitive, precise, and reproducible method to quantitate alternative splicing of mRNA transcripts. Droplet digital™ PCR (ddPCR™) analysis allows for accurate digital counting for quantification of gene expression. Human telomerase reverse transcriptase (hTERT) is one of the essential components required for telomerase activity and for the maintenance of telomeres. Several alternatively spliced forms of hTERT mRNA in human primary and tumor cells have been reported in the literature. Using one pair of primers and two probes for hTERT, four alternatively spliced forms of hTERT ( $\alpha-/ \beta+$ ,  $\alpha+/ \beta-$  single deletions,  $\alpha-/ \beta-$  double deletion, and nondelletion  $\alpha+/ \beta+$ ) were accurately quantified through a novel analysis method via data collected from a single ddPCR reaction. In this chapter, we describe this ddPCR method that enables direct quantitative comparison of four alternatively spliced forms of the hTERT messenger RNA without the need for internal standards or multiple pairs of primers specific for each variant, eliminating the technical variation due to differential PCR amplification efficiency for different amplicons and the challenges of quantification using standard curves. This simple and straightforward method should have general utility for quantifying alternatively spliced gene transcripts.

**Key words** mRNA alternative splicing, mRNA quantification, hTERT, qPCR, Droplet digital PCR

---

### 1 Introduction

Humans produce around 150,000 different proteins from their 25,000–30,000 genes, by alternative splicing. More than 70% of human protein-coding genes are alternatively spliced [1], allowing the production of different proteins with diverse and even antagonistic functions from a single DNA sequence. As alternative splicing affects numerous genes, it is not surprising that changes in alternative splicing are frequently associated with human diseases. A large proportion of human genetic disorders result from splicing variants. Tauopathies is an example of a disease caused by a change in the ratio of protein isoforms generated by alternative splicing [2]. Spinal muscular atrophy (SMA) is another example of a recessive disease caused by a point mutation in an exonic regulatory

element [3]. Abnormal splicing variants also contribute to the development of cancer [4, 5]. Given the important role of alternative splicing in regulating cellular functions, accurate quantification of multiple alternatively spliced transcripts will facilitate the discovery of new biomarkers for clinical application and enhance our understanding of the role of alternatively spliced transcripts in health and disease.

Around 20 alternatively spliced variants of human telomere reverse transcriptase (hTERT) have been reported; only four variants (i.e.,  $\alpha-/ \beta+$  [ $\alpha$  deletion],  $\alpha+/ \beta-$  [ $\beta$  deletion],  $\alpha-/ \beta-$  [ $\alpha \& \beta$  double deletion], or  $\alpha+/ \beta+$  [nondeletion]) are commonly present in most tumor tissues and may serve as specific markers for cancer diagnosis, prediction of clinical outcome, and drug targeting [6–12]. hTERT is involved in the control of telomerase activity and plays significant roles in cell replication and carcinogenesis. Only the full-length hTERT (or  $\alpha+/ \beta+$ ) has been shown to be associated with telomerase activity. The  $\alpha$ -deletion transcript has a 36-bp deletion within the conserved reverse transcriptase motif and is a dominant negative inhibitor of telomerase activity [5]. The  $\beta$ -deletion transcript has a 182-bp deletion leading to a truncated protein before the conserved reverse transcriptase motifs, resulting in a catalytically inactive telomerase [8, 9]. Over 90% of tumor tissues have increased telomerase activity when compared to adjacent normal tissues [10, 11]. However, the correlation between the level of telomerase activity and the level of hTERT transcripts in the tumor tissues was low to moderate, i.e., high level expression of full-length hTERT (or  $\alpha+/ \beta+$ ) is not always associated with high telomerase activity [12], which could be the result of post-transcriptional processes leading to the formation of several alternatively spliced variants that serve as an additional mechanisms of telomerase regulation. The relative quantities of each spliced variant may determine the overall telomerase activity. For example, the co-presentation of  $\alpha$ -deletion hTERT will result in the reduced telomerase activity depending on its expression level relative to the expression level of full length hTERT (the ratio between  $\alpha-/ \beta+$  and  $\alpha+/ \beta+$ ).

Next generation sequencing is considered the gold standard for profiling the expression pattern of alternatively spliced variants. However its application is limited due to high cost and the large amount of RNA required, and it does not typically render absolute transcript copy numbers. Microarray assays (e.g., the Affymetrix exon microarray, ExonHit) have been developed to detect alternatively spliced variants for specific exons; however, these methods often cannot distinguish between the different variants (i.e., containing exon A only or exon B only from variants with both exons A and B). Several real-time PCR (qPCR)-based assays for the enumeration of the common hTERT spliced variants have been developed to determine hTERT splicing patterns, but the correlation

between the levels of telomerase activity and hTERT transcript levels is not consistent [12–16]. One caveat is that the qPCR-based methods have limited sensitivity to precisely quantify some of the alternatively spliced transcripts that are expressed at very low levels. In addition, a separate qPCR reaction is required to measure each alternatively spliced variant with different primer pairs. Variations in PCR amplification efficiency could compromise the quantification accuracy, making comparisons less reliable. In the present chapter, using hTERT expression as an example, we describe the use of droplet digital PCR (ddPCR) for the quantification of the four major alternatively spliced transcripts of hTERT, by using one pair of primers in a single PCR reaction, and without the need for internal standards. This simple method reliably quantitates four distinguishable clusters of droplets and allows the direct comparison of expression levels among the four spliced variants in different cell types.

---

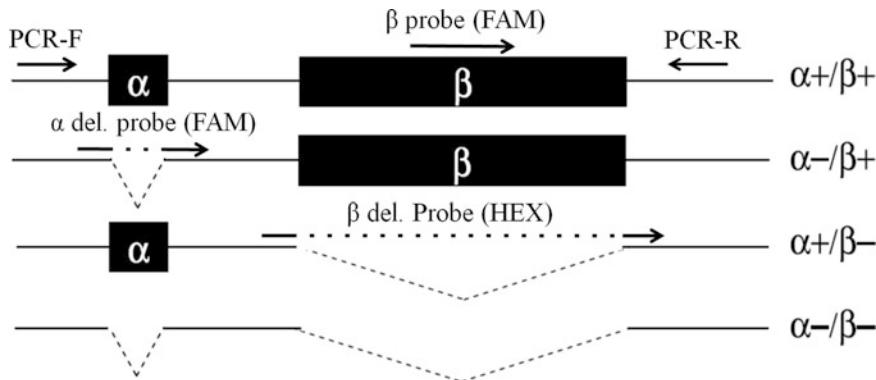
## 2 Materials

### 2.1 Cell Lines Evaluated

All the cell lines (ovarian cancer OVCar-3, osteosarcoma SAOS-2 and U2OS, breast cancer MCF-7, lung fibroblast WI-38, and endometrium adenocarcinoma HEC-1-A) used in the study were obtained from ATCC (Manassas, VA) or from Cell Culture Core Facility of Lombardi Cancer Center, Georgetown University Medical Center, and cultured with the media as recommended by the providers. RNA can also be extracted from fresh tissues or frozen fresh tissues, and used for transcript quantification with this method.

### 2.2 Primer Design and Sequences

1. All primers and probes were synthesized by Integrated DNA Technologies, Inc. (Coralville, IW).
2. The forward-primer for hTERT cDNA amplification (PCR-F in Fig. 1) is located at upstream of  $\alpha$ -deletion splicing site (5'-GCCTGAGCTGTACTTGTCA-3'). The reverse-primer (PCR-R in Fig. 1) is at downstream of  $\beta$ -deletion splicing site with sequence of 5'-CAGCGTGGAGAGGATGGAG-3'. The amplicon size are 376, 230, 194, and 412 bp for  $\alpha-$ / $\beta+$  ( $\alpha$  deletion),  $\alpha+/-\beta-$  ( $\beta$  deletion),  $\alpha-/$  $\beta-$  ( $\alpha\&\beta$  double deletion), or  $\alpha+/\beta+$  (full length), respectively.
3. Dual-labeled  $\alpha$ -deletion probe was designed with FAM as the fluorescent indicator and ZEN-Iowa Black as the quencher—positioned as shown in Fig. 1. Sequence [5'-/FAM/TACTTGTC/ZEN/AAGGACAGGCTCACG/IABk/-3')] partially overlaps the transcript both before and after the deleted region, and specifically recognizes the  $\alpha$ -deletion transcripts only.



**Fig. 1** Location of primers and probes for the detection of hTERT mRNA  $\alpha$ -deletion ( $\alpha-/ \beta+$ ),  $\beta$ -deletion ( $\alpha+/ \beta-$ ), and  $\alpha-/ \beta-$  double deletion. Deleted regions are designated by dashed lines

4. The dual-labeled  $\beta$ -deletion probe was designed in a similar fashion [5'-/HEX/CTTCAAGAG/ZEN/CCACGTCC-TACGTC/IABk/-3'], using HEX as the fluorescent indicator.
5. The dual-labeled  $\beta$  probe was designed within the deleted  $\beta$  region [5'-/6-FAM/CTCTACCTT/ZEN/GACA-GACCTCCAGC/IABkFQ/-3'], using FAM as the fluorescent indicator. This probe was used to identify the copresence of single  $\alpha$ -deletion and single  $\beta$ -deletion in the same droplet using cloned hTERT variants.
6. GAPDH primers and probes: forward primer, 5'-ATTCCACCCATGGCAAATTC-3'; reverse primer 5'-TGGGATTTCATTGATGACAAG-3' with an amplicon of 72 bp; and probe, 5'-/HEX/CAAGCTTCCCCTTCT-CAGCC/IABkFQ/-3' was used as a normalization control for the comparison between samples, assuming that GAPDH expression is stable across all samples.

### 2.3 Positive Control for hTERT Transcripts

To serve as positive controls for identification and quantification of each variant, cloned hTERT DNA fragments of the four alternatively spliced variants were amplified by PCR using the hTERT PCR-F primer and another reverse primer (5'-CAAACAGCTTCTCCATGT-3') downstream of hTERT PCR-R. These amplification products (419, 273, 237, and 455 bp for  $\alpha-/ \beta+$ ,  $\alpha+/ \beta-$ ,  $\alpha-/ \beta-$ , or  $\alpha+/ \beta+$ , respectively) were previously cloned into pCR4-TOPO TA vector (Life Technologies, Grand Island, NY), and their sequences were verified by Sanger DNA sequencing (Genewiz, Germantown, MD). These controls are available upon request.

## 2.4 ddPCR Related Materials and Reagents

1. 2× Droplet Digital™ PCR Supermix for probes (without dUTP; Bio-Rad, cat #186-3025).
2. Droplet generation oil for probes (Bio-Rad, cat #186-3005).
3. Eight-channel DG8™ droplet generation cartridges (Bio-Rad).
4. Pierceable foil plate heat seal (Bio-Rad);
5. Heat-sealable semiskirted 96-well PCR plates (Eppendorf North America, Inc., Hauppauge, NY).
6. Droplet Reader oil (Bio-Rad).
7. Pipet-Lite XLS (2–20 µL) and RT-L10F pipette tips (filtered, Mettler Toledo), both from Rainin (Columbus, OH) for loading ddPCR mix into droplet generation cartridge.
8. Pipet-Lite XLS (8 channels, 5–50 µL) and RT-L200F pipette tips (filtered, Mettler Toledo), both from Rainin, for transferring droplets from droplet generation cartridge to Eppendorf PCR plate.
9. Plate Sealer (Eppendorf).
10. Thermocycler T100 (Bio-Rad).

## 3 Methods

### 3.1 RNA Extraction and cDNA Synthesis

1. Total RNA from cultured tumor cell lines and fibroblasts was extracted with TRIzol total RNA extraction kit (Life Technologies).
2. RNA measurement: 1 µL of extracted total RNA was analyzed by spectrometry on a NanoDrop-1000 to determine concentration, with 260/280 ratio ≥ 2.0.
3. One µg total RNA was used for cDNA synthesis in a 10 µL reaction, using the iScript Select cDNA synthesis kit (Bio-Rad) and 19 nt oligo dT + A/G/C to ensure 1:1 reverse transcription (*see* Subheading 4, Note 1).

### 3.2 Setting Up ddPCR Reaction

1. hTERT reaction mixtures were assembled by combining 11.5 µL of 2× ddPCR Supermix (*see* Subheading 4, Note 2), forward and reverse primers (250 nM final concentration), probes (125 nM final concentration each), 1 µL cDNA (equivalent to 50 ng of initial total RNA, depending on expression levels and RT efficiency) and nuclease free water in a final volume of 22.5 µL.
2. For GAPDH, the cDNA was diluted 1:2000 and mixed with same reagents as above. Quantification was performed in a separate ddPCR reaction.

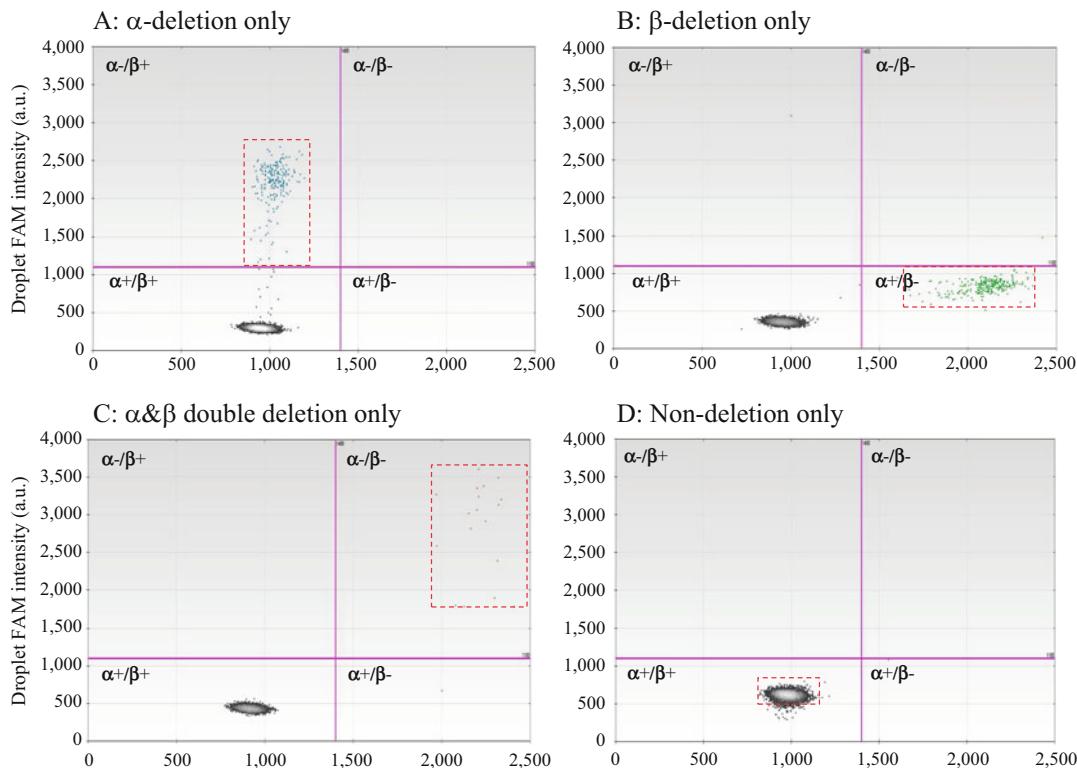
### 3.3 ddPCR Reaction Partitioning and Amplification

PCR reaction and detection were carried out according to manufacturer's instructions [17].

1. Each assembled ddPCR reaction mixture (20 µL) was loaded into the sample well of an eight-channel droplet generator cartridge.
2. Seventy microliters of droplet generation oil was loaded into the oil well for each channel.
3. The cartridge was placed into the droplet generator after securing a gasket over the loaded chip, for droplet generation (up to 20,000 droplets per reaction).
4. The droplets for each sample which collected in the cartridge droplet wells were then manually transferred with a multichannel pipette to a 96-well PCR Eppendorf plate.
5. The plate was heat-sealed with a foil seal and then placed on a conventional thermal cycler (T100, Bio-Rad).
6. Thermal cycling conditions were 95 °C for 10 min (1 cycle), followed by 45 cycles of 95 °C for 30 s and 60 °C for 1 min, then by 98 °C for 10 min (1 cycle), and hold at 12 °C until detection. Longer extension time may be required for the detection of long amplicon.
7. After PCR, the 96-well PCR plate was transferred to the droplet reader (QX100, Bio-Rad), which automatically reads the droplets from each well of the plate using appropriate settings predefined by the manufacturer (Bio-Rad).

### 3.4 Analysis of Positive Controls

To establish the transcript detection specificity of the multiplexed hTERT assay, each of the four cloned hTERT variants was amplified by ddPCR in a separate reaction to determine the respective cluster locations. Representative results are shown in a 2D plot in Fig. 2. Droplets for each variant are clearly shown: α deletion (Fig. 2a), β deletion (Fig. 2b), α&β double deletion (Fig. 2c, see Subheading 4, Note 3), and nondeletion (Fig. 2d, see Subheading 4, Note 4). The droplets with nondeletion hTERT variant appeared at slightly higher fluorescence amplitude relative to the double negative droplets (lacking any of the hTERT transcript isoforms) in the lower left quadrant in Fig. 2a–c (see explanation in the Subheading 4, Note 4). The ability of this multiplex assay to distinguish multiple isoforms in the same well is shown in Fig. 3, in which 2, 3, or 4 cloned hTERT variants are combined in a single ddPCR reaction to simulate samples from human cells or tissues. To identify the location of copresence of α-deletion and β-deletion in the same droplets, cloned α-deletion and β-deletion templates were titrated and used at high concentrations with 25–50% droplets showing positive signals (Fig. 3a, c, and d).



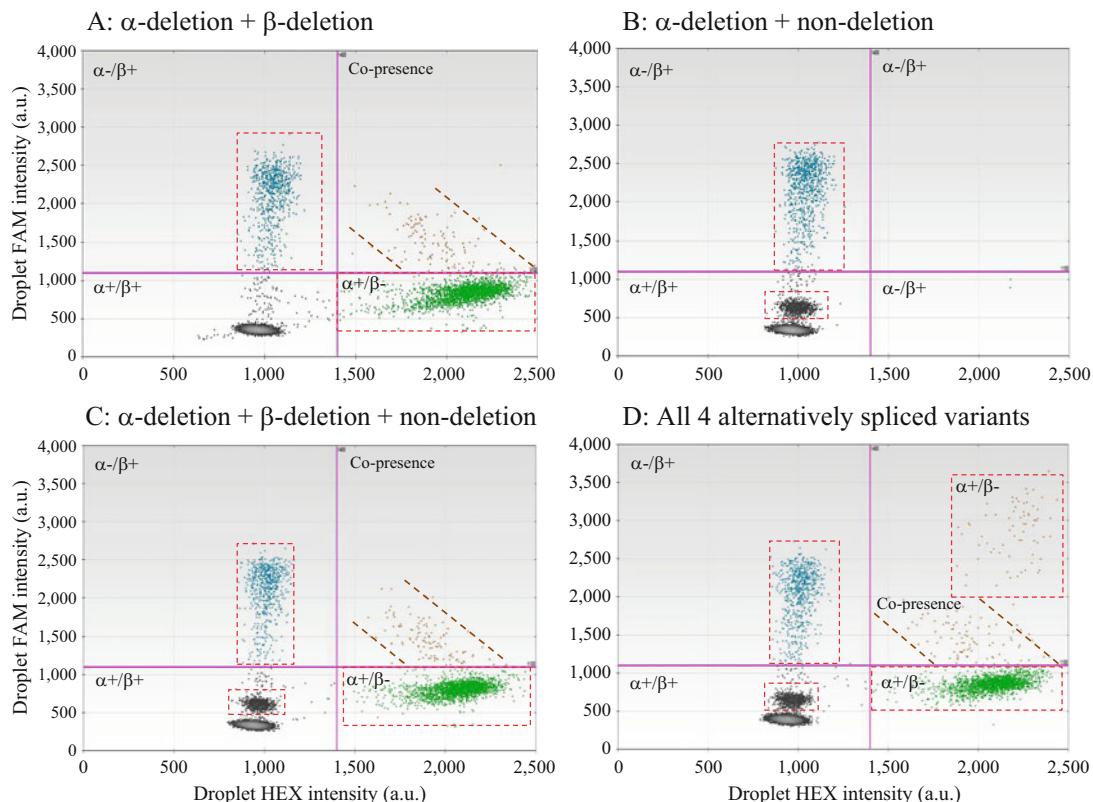
**Fig. 2** 2D display of droplet distribution after ddPCR amplification of cloned hTERT alternatively spliced transcripts (**a**— $\alpha$  deletion only, **b**— $\beta$  deletion only, **c**— $\alpha\&\beta$  double deletion only, and **d**—Nondeletion only) in a single droplet digital PCR reaction containing both  $\alpha$ -deletion (FAM) and  $\beta$ -deletion (HEX) probes. FAM-positive droplets (i.e.,  $\alpha-/ \beta+$ ) are in the upper left quadrant. HEX-positive droplets (i.e.,  $\alpha+/ \beta-$ ) are in the lower right quadrant. Double-positive droplets ( $\alpha-/ \beta-$  in dashed rectangle) are in the upper right quadrant. Weak FAM-positive ( $\alpha+/ \beta+$  in dashed rectangle) or negative droplets are in the lower left quadrant

### 3.5 Verifying the Detection of Colocalized hTERT Variants

A separate ddPCR reaction with the same primer pair but including only the nondeletion  $\beta$  probe (FAM-labeled, Fig. 1) and the  $\beta$ -deletion probe (HEX) was performed to assess the ability to detect the copresence of single  $\alpha$ -deletion and single  $\beta$ -deletion in the same droplet. In the case of hTERT expression, the transcript levels of each variant in all the cells evaluated are low, even in cancer cells. Thus the likelihood of having colocalization of two or more hTERT variants in one droplet is extremely low. It is generally advised to titrate cDNA sample input for ddPCR to have less than 5000 positive droplets in a reaction in order to avoid the presence of new clusters due to the colocalization of two or more variants in one droplet (see Subheading 4, Note 5).

### 3.6 Data Analysis

Instead of using autoanalysis after data acquisition with QuantaSoft analysis software, manually select “+/-”, “-/+”, “-/-” counts by using the “*lasso*” function in the 2D plots, relative to the positive counts for each alternatively spliced transcripts, allowing the



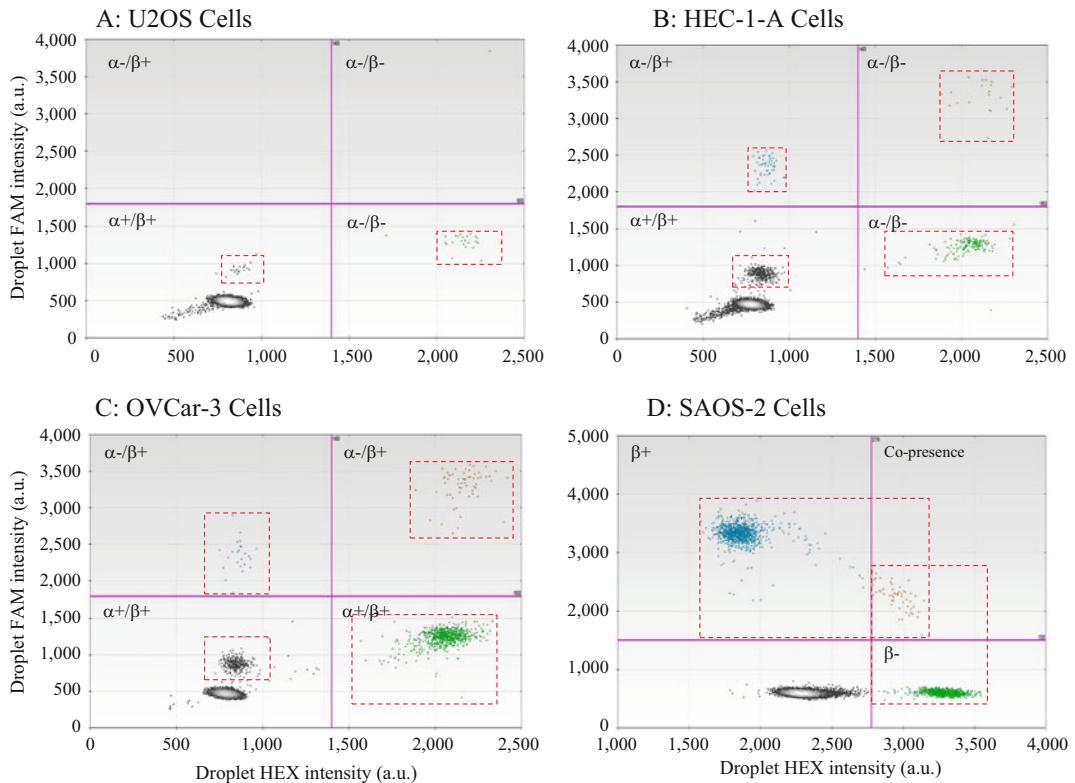
**Fig. 3** 2D display of droplet distribution after ddPCR amplification of cloned hTERT variants in different combination (**a**— $\alpha$ -deletion +  $\beta$ -deletion, **b**— $\alpha$ -deletion + nondeletion, **c**— $\alpha$ -deletion + -  $\beta$ -deletion + nondeletion, and **d**—all four alternatively spliced variants) in a single droplet digital PCR reaction containing both  $\alpha$ -deletion (FAM) and  $\beta$ -deletion (HEX) probes. FAM-positive droplets (i.e.,  $\alpha-/ \beta+$ ) are in the upper left quadrant. HEX-positive droplets (i.e.,  $\alpha+/ \beta-$ ) are in the lower right quadrant. Double-positive droplets ( $\alpha-/ \beta-$  in dashed rectangle, and the copresence of  $\alpha$ - and  $\beta$ -deletion transcripts between dashed lines) are in the upper right quadrant. Weak FAM-positive ( $\alpha+/ \beta+$ ) or negative droplets are in the lower left quadrant

calculation of the copy numbers of  $\alpha$ -deletion (FAM-positive),  $\beta$ -deletion (HEX-positive),  $\alpha\&\beta$  double deletion, or copresence of single  $\alpha$ - and  $\beta$ -deletions (FAM- and HEX-double positive) and nondeletion forms (weak FAM-positive) in the four quadrants. More tedious but accurate way to quantitate the four variants in a given sample can be done by greying out all the droplets and selecting FAM or HEX positive droplets that belong to a specific group (for example, Fig. 3d cluster just above negative droplets  $\alpha+/ \beta+$  are selected as positive. For  $\alpha-/ \beta+$ , high FAM cluster and Copresence cluster would be selected, and all the others would be grey).

---

## 4 Notes

1. It is important to make an effort to achieve 1:1 cDNA synthesis from mRNA by using specific oligonucleotide for your targeted gene transcripts. The oligo dT we used for reverse transcription, not random primers, in combination with reverse transcriptase with RNase H activity is preferred for cDNA synthesis, to ensure transcripts were not used more than once to generate cDNA.
2. To minimize “rain drops” (droplets with relatively much lower fluorescent counts) when using cloned cDNA in a vector as control, linearized template should be used. In addition, primer design and PCR conditions need to be modified to achieve high PCR amplification and specificity.
3. FAM- and HEX-double positive droplets were detected, as shown in the upper right quadrant in Fig. 4a, c (dashed rectangles). There are two possibilities for droplets shown in the upper right quadrant: the presence of a transcript with  $\alpha\&\beta$  double deletion or co-presence of  $\alpha$ -deletion and  $\beta$ -deletion transcripts in the same droplet. In the latter case, two independent PCR amplifications using either  $\alpha$ -deletion transcript or  $\beta$ -deletion transcript as template, happened in the same droplet. In theory, the fluorescence intensity of double-positive droplets would be stronger in droplets containing a  $\alpha\&\beta$  double deletion variant than in droplets containing both  $\alpha$ -deletion and  $\beta$ -deletion variants, possibly due to the difference in length of their amplicon, with shorter amplicon ( $\alpha\&\beta$  double deletion variants) having higher PCR amplification efficiency. To demonstrate this, we conducted experiments using cloned hTERT variants with following combination:  $\alpha$ -deletion and  $\beta$ -deletion,  $\alpha\&\beta$  double deletion, and all four cloned fragments together. As shown in Fig. 3a, c, the fluorescent intensity of the droplets with copresence of single  $\alpha$ -deletion and single  $\beta$ -deletion in the upper-right quadrant (between the dashed parallel lines) were noticeably weaker comparing to that with  $\alpha\&\beta$  double deletion seen in the same quadrants of Figs. 2c and 3d (dashed rectangle). The large differences in signal intensity between the copresence of  $\alpha$ - and  $\beta$ -deletions and the  $\alpha\&\beta$  double deletion allows the  $\alpha\&\beta$  double deletion droplets to be counted. Accordingly, all the four cloned control templates in the Fig. 3 were calculated and presented in Table 1. It was shown that the counts for each variant are consistent among the different combination measurements. As for the counts of  $\alpha\&\beta$  double deletion and of co-presence, these two populations of double-positive droplets can be well separated and



**Fig. 4** 2D display of droplet distribution after ddPCR amplification of mRNA from human cancer cell lines. Each of the three cell line (**a**—U2OS, **b**—HEC-1-A, and **c**—OVCar-3) was amplified in a single droplet digital PCR reaction containing both  $\alpha$ -deletion (FAM) and  $\beta$ -deletion (HEX) probes. FAM-positive droplets (i.e.,  $\alpha-/ \beta+$ ) are in the upper left quadrant. HEX-positive droplets (i.e.,  $\alpha+/ \beta-$ ) are in the lower right quadrant. Double-positive droplets (i.e.,  $\alpha-/ \beta-$  or the copresence of single  $\alpha$ - and  $\beta$ -deletions) are in the upper right quadrant. Weak FAM-positive ( $\alpha+/ \beta+$ ) or negative droplets are in the lower left quadrant. **d**—cell line SAOS-2 was amplified in a single droplet digital PCR reaction with nondeletion  $\beta$  (FAM) and  $\beta$ -deletion (HEX) probes. FAM-positive droplets (i.e.,  $\beta+$ ) are in the dashed rectangle in upper half. HEX-positive droplets (i.e.,  $\beta-$ ) are in the dashed rectangle on the right. Double-positive droplets, i.e., copresence of non- $\beta$  deletion ( $\beta+$ ) and  $\beta$ -deletion ( $\beta-$ ), are in the upper right quadrant

counted separately. The hTERT expression levels are generally low in normal tissue or tumor samples, and thus co-preservation of  $\alpha$ - and  $\beta$ -deletion variants in the same droplet is unlikely. Figure 4d shows that the co-preservation of  $\beta$ -deletion and non- $\beta$  deletion transcripts (the two most abundant spliced variants in all the cells) in the same droplet is very rare (<0.1%). This observation is true even in SAOS2 cells that have the highest hTERT expression level among all the cells tested. Therefore, the double-positive droplets in the upright corners of the upper right quadrants in Fig. 4b, c represent the detection of  $\alpha\&\beta$  double deletion transcripts.

**Table 1**

**Calculated copies of each cloned hTERT constructs for the co-presence of  $\alpha-$  and  $\beta-$  deletion constructs as shown in Fig. 3**

Mix of cloned variants	Double-positive droplets				
	$\alpha$ -deletion	$\beta$ -deletion	Non-deletion	Double-deletion	Co-presence
$\alpha-/ \beta+$ and $\alpha+/ \beta-$	1392	4580	36	2	202
$\alpha-/ \beta+$ and $\alpha+/ \beta+$	1392	3	2098	0	0
$\alpha-/ \beta+$ and $\alpha+/ \beta-$ and $\alpha+/ \beta+$	1258	4460	1774	0	192
$\alpha-/ \beta+$ and $\alpha+/ \beta-$ and $\alpha+/ \beta+$ and $\alpha-/ \beta-$	1354	4140	1874	108	182

Values were from single measurement

4. During our initial testing, it was observed that, in addition to the quantification of  $\alpha$ -deletion ( $\alpha-/ \beta+$ , as in Fig. 4b, c) and  $\beta$ -deletion ( $\alpha+/ \beta-$  in the right lower quadrant, such as in Fig. 4a–c), a group of weak FAM-positive droplets were detected just above the negative droplets in the lower left quadrant (as shown in dashed rectangles of Fig. 4a–c). We suspected that these droplets contain nondeletion hTERT transcripts ( $\alpha+/ \beta+$ ) which were detected with weak FAM fluorescent signals due to inefficient hydrolysis of  $\alpha$ -deletion probe, as  $\alpha$ -deletion probe probably partially annealed to the nondeletion form at the annealing temperature during PCR. To confirm this, we carried out an experiment with cloned  $\alpha$ -deletion plasmids plus the cloned nondeletion plasmid as templates for ddPCR. Compared to the results obtained with cloned  $\alpha$ -deletion plasmid only reaction as shown in Fig. 2a, a new group of weak FAM-positive droplets in the lower left quadrant (dashed rectangle) appeared after adding cloned nondeletion hTERT fragments to the cloned  $\alpha$ -deletion (Fig. 3b–d). The quantification accuracy for the nondeletion form was also confirmed by the subtraction of  $\alpha$ -deletion counts from non- $\beta$  deletion counts obtained from the experiments with nondeletion  $\beta$  probe and  $\beta$ -deletion probe. Although the detection is coincident and may not apply to all situations, similar approach with proper probe design plus optimization testing may work for other gene transcript detection as well.
5. When the expression levels are high, precaution should be taken to appropriately dilute template/sample if ddPCR will be used. On the other hand, especially when the gene

expression level is very low, ddPCR amplification conditions including the primer design should be optimized to avoid any nonspecific amplification in order to obtain accurate counts for each variant.

6. For normalization with a reference gene, titration of the input template of the reference gene is required for accurate quantification because most of the reference genes such as GAPDH or  $\beta$ -actin are expressed at much higher levels compared to the targeted gene. After normalization by GAPDH expression, this method permits the comparison of  $\alpha$ -deletion,  $\beta$ -deletion,  $\alpha\&\beta$  double deletions, nondeletion, and total hTERT expression among different samples. Table 2 shows that SAOS-2 cells expressed the highest level of total hTERT, double deletion and  $\beta$ -deletion forms, whereas HEC-1-A cells expressed the highest percentage of  $\alpha$ -deletion among all the cells tested. The U2OS is a cell line using an alternative lengthening of telomeres (ALT) mechanism to maintain its telomere length and have been shown lack of telomerase activity [18]. We demonstrated the expression of extremely low levels of  $\beta$ -deletion and nondeletion hTERT transcripts in this cell line, indicating the high sensitivity of the method. The lack of telomerase activity in the U2OS cells is likely due to the absence of hTERC expression, the template component of telomerase complex (unpublished observation with ddPCR). Lack of hTERC expression is also observed in several other cell lines depending on the ALT to maintain telomere length [19].
7. Most importantly, this method allows direct comparison of the expression levels among the four alternatively spliced hTERT variants. As shown in Fig. 4b, the  $\alpha$ -deletion variant HEC-1-A cells are expressed relatively higher in comparison to  $\beta$ -deletion variant than that in OVCar-3 cells (Fig. 4c). The ratio of double deletion variant to  $\alpha$ -deletion variant in the OVCar-3 cells is higher than that in HEC-1-A cells. The expression profiling of MCF-7 cells obtained in the present study with  $\alpha$ -deletion at 1.87%,  $\beta$ -deletion at 63.50%, double deletion at 2.79%, and nondeletion at 31.85% (Table 2) are very similar to the data reported previously with qPCR (estimated at 3.5%, 62.5%, 3.0%, and 31.5%, respectively) [15]. To the best of our knowledge, the method presented here is the first to provide a simple, precise and absolute quantification of four alternatively spliced hTERT transcripts, in particular for the detection of single as well as double deletions simultaneously for direct comparison.
8. The new method not only minimizes experimental errors that may be introduced by pipetting, but also eliminates variations from PCR amplification efficiency.

**Table 2**  
The expression of 4 alternatively spliced hTERT transcripts in cell lines

Cell line	Expression relative to GAPDH (x 1000)				Percentage of total hTERT expression				
	<i>α-deletion</i>	<i>β-deletion</i>	<i>Double-deletion</i>	<i>Non-deletion</i>	Total hTERT	<i>α-deletion</i>	<i>β-deletion</i>	<i>Double-deletion</i>	<i>Non-deletion</i>
SAOS-2	47.09	3668.28	253.03	1049.76	5018.16	0.94	73.10	5.04	20.92
U2OS	0.00	13.93	0.72	8.93	23.58	0.00	59.08	3.07	37.85
WI-38	0.00	1.44	0.00	0.47	1.91	0.00	75.44	0.00	24.56
HEC-1-A	37.61	162.44	14.35	622.07	836.46	4.50	19.42	1.72	74.37
OVCar-3	23.10	779.54	45.89	435.76	1284.28	1.80	60.70	3.57	33.93
MCF7	20.83	708.88	31.12	355.50	1116.33	1.87	63.50	2.79	31.85

Values were the average of duplicate measurements

## Acknowledgments

Research in YLZ's laboratory is supported by grants from the National Cancer Institute of the National Institutes of Health (R01CA132996) and Susan G. Komen for the Cure (KG100283).

## References

1. Matlin AJ, Clark F, Smith CWJ (2005) Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol* 6:386–398
2. Hong M, Zhukareva V, Vogelsberg-Ragaglia V et al (1998) Mutation-specific functional impairments in distinct tau isoforms of hereditary FTDP-17. *Science* 282:1914–1917
3. Lefebvre S, Burglen L, Reboullet S et al (1995) Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 80:155–165
4. Faustino NA, Cooper TA (2003) Pre-mRNA splicing and human disease. *Genes Dev* 17:419–437
5. Cáceres JF, Kornblith AR (2002) Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet* 18:186–193
6. Ulaner GA, JF H, TH V et al (1998) Telomerase activity in human development is regulated by human telomerase reverse transcriptase (hTERT) transcription and by alternate splicing of hTERT transcripts. *Cancer Res* 58:4168–4172
7. Colgin LM, Wilkinson C, Englezou A et al (2000) The hTERTalpha splice variant is a dominant negative inhibitor of telomerase activity. *Neoplasia* 2:426–432
8. Listerman I, Sun J, Gazzaniga FS et al (2013) The major reverse transcriptase-incompetent splice variant of the human telomerase protein inhibits telomerase activity but protects from apoptosis. *Cancer Res* 73:2817–2828
9. Lincz LF, Mudge LM, Scorgie FE et al (2008) Quantification of hTERT splice variants in melanoma by SYBR green real-time polymerase chain reaction indicates a negative regulatory role for the beta deletion variant. *Neoplasia* 10:1131–1137
10. Liu Y, BQ W, Zhong HH et al (2012) Quantification of alternative splicing variants of human telomerase reverse transcriptase and correlations with telomerase activity in lung cancer. *PLoS One* 7:e38868
11. Fujiwara-Akita H, Maesawa C, Honda T et al (2005) Expression of human telomerase reverse transcriptase splice variants is well correlated with low telomerase activity in osteosarcoma cell lines. *Int J Oncol* 26:1009–1016
12. Zaffaroni N, Villa R, Pastorino U et al (2005) Lack of telomerase activity in lung carcinoids is dependent on human telomerase reverse transcriptase transcription and alternative splicing and is associated with long telomeres. *Clin Cancer Res* 11:2832–2839
13. Yi X, Shay JW, Wright WE (2001) Quantitation of telomerase components and hTERT mRNA splicing patterns in immortal human cells. *Nucleic Acids Res* 29:4818–4825
14. Ohyashiki JH, Hisatomi H, Nagao K et al (2005) Quantitative relationship between functionally active telomerase and major telomerase components (hTERT and hTR) in acute leukaemia cells. *Br J Cancer* 92:1942–1947
15. Mavrogiannou E, Strati A, Stathopoulou A et al (2007) Real-time RT-PCR quantification of human telomerase reverse transcriptase splice variants in tumor cell lines and non-small cell lung cancer. *Clin Chem* 53:53–61
16. Hara T, Noma T, Yamashiro Y et al (2001) Quantitative analysis of telomerase activity and telomerase reverse transcriptase expression in renal cell carcinoma. *Urol Res* 29:1–6
17. Hindson BJ, Ness KD, Masquelier DA et al (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem* 83:8604–8610
18. Ulaner GA, JF H, TH V et al (2000) Regulation of telomerase by alternate splicing of human telomerase reverse transcriptase (hTERT) in normal and neoplastic ovary, endometrium and myometrium. *Int J Cancer* 85:330–335
19. Hoare SF, Bryce LA, Wisman GBA et al (2001) Lack of telomerase RNA gene hTERC expression in alternative lengthening of telomere cells is associated with methylation of the hTERC promoter. *Cancer Res* 61:27–32



# Chapter 23

## Using Droplet Digital PCR to Analyze Allele-Specific RNA Expression

Nolan Kamitaki, Christina L. Usher, and Steven A. McCarroll

### Abstract

Genome-wide association studies have discovered thousands of common alleles that associate with human phenotypes and disease. Many of these variants are in non-protein-coding (regulatory) regions and are believed to affect phenotypes by modifying gene expression. In any organism with a diploid genome, such as humans, measuring the expression of each allele of a gene provides a well-controlled way to identify allelic influences on that gene's expression. Here, we describe a protocol for precisely measuring the allele-specific expression of individual genes. This method targets the nucleotide differences between the two alleles of a gene within an individual and measures the “allelic skew,” the extent to which one allele is expressed more than the other. We cover the design of effective assays, the optimization of reactions, and the interpretation of the resulting data.

**Key words** Allele-specific expression, Allelic skew, Allelic imbalance, Droplet digital PCR, Digital PCR, mRNA expression, Assay design

---

### 1 Introduction

Genome-wide association studies (GWAS) have mapped hundreds of human traits to thousands of common variants [1]. Less than 30% of these variants are nonsynonymous or are in linkage disequilibrium with a variant that is [2], suggesting that much of the variation driving phenotypic differences does so not by altering protein function, but by altering gene expression. Furthermore, systematic substitution of DNA bases in known enhancers has shown that the vast majority of these variants modestly affect expression (less than twofold change) and very few abolish or greatly increase expression [3]. Measuring these modest effects on gene expression requires techniques that can both ascertain modest differences in expression and can filter through the noise created by environmental influences, genomic background, and other *trans*-acting effects.

Once a candidate gene has been selected, usually by hypotheses about the mechanisms and genetic architecture of a phenotype, we can compare the expression of the two alleles within the same biological sample, thus avoiding many sources of noise. In general, a person's two alleles are subject to the same environmental and *trans*-acting genetic influences, which are potentially strong enough to obscure modest genetic effects and can be difficult to control for in study designs that assay total expression in genetically diverse humans living in variable environments [4].

In the absence of allele-specific regulatory effects, the two alleles will be expressed at a ratio of 1:1, but local *cis*-acting genetic and epigenetic variation can cause one allele to be expressed more than the other. Measuring this difference requires a technology capable of measuring the two alleles precisely and with equal sensitivity. Techniques that take this measurement after PCR amplification—including RNA-seq and pyrosequencing—run the risk of confusing amplification effects with real genetic effects. Even a subtle difference of 1% in amplification efficiency will, after 30 PCR cycles, result in a 1.35-fold relative difference in the abundance of the two alleles in the amplified material, an effect size that would exceed that of most enhancer SNPs. In addition, digital measurement, which counts the number of transcripts with each allele in a sample, is preferable to analog measurement, which estimates ratios, since digital measurement allows the statistical significance of allelic skews to be estimated for each person.

The ability to digitally count the abundance of alleles in an unamplified sample is therefore critical. 1-Step RT-droplet digital PCR (RT-ddPCR) accomplishes this by partitioning the individual RNA transcripts of a gene into separate reaction compartments (droplets) before amplification/detection [5]. Within these ~20,000 droplets, the RNA is reverse-transcribed into cDNA and amplified using a molecular assay that is composed of: a pair of primers that will amplify both alleles, a fluorescence probe targeting the reference allele, and a probe (with a different fluorophore) targeting the alternate allele. The number of RNA molecules originating from each allele is quantified by counting the number of droplets that fluoresce with the corresponding probe colors. Provided that the positive and negative droplets are clearly and consistently distinguished, ddPCR is robust to differences in the amplification kinetics of the alleles (a clear advantage over pyrosequencing and real-time PCR). Droplet digital PCR is also more resistant to the effects of nonspecific binding of the probes to the opposite allele, allowing the determination of the correct allele in each droplet as long as each probe has greater affinity for its respective allelic target.

Here, we share our protocol for measuring the allelic skew of a gene using ddPCR. We cover how to design a successful allele-specific assay, how to screen and optimize that assay, and finally,

how to use it in ddPCR. We focus on assay design in humans, but the protocol is readily adaptable to any species with a diploid genome and substantial heterozygosity. It is also readily adapted to measure the expression levels of paralogous genes in a paralog-specific manner.

However, this method has limitations in that it can only assay those individuals heterozygous for the transcribed reporter SNP, and historic recombination between the reporter SNP and the functional variant may introduce uncertainty about the direction of effect. Thus, the clearest overall picture may emerge when this strategy is combined with another that assays total expression across a large cohort [6].

---

## 2 Materials

### 2.1 Locus-Specific Analysis Reagents

1. 18  $\mu\text{M}$  forward primer; 18  $\mu\text{M}$  reverse primer; 5  $\mu\text{M}$  5' FAM-labeled, 3' Iowa Black ZEN-quenched probe (*see Note 1*) that targets the reference allele; and 5  $\mu\text{M}$  5' HEX-labeled, 3' Iowa Black ZEN-quenched probe that targets the alternate allele. Store at  $-20^\circ\text{C}$  away from light (*see Note 2*). Vortex and centrifuge before use.

### 2.2 Biological Sample

1. RNA samples (100 fg–100 ng of total RNA) from individuals heterozygous for the reporter SNP (*see Note 3*). They need not be heterozygous for the candidate functional variant. Store at  $-80^\circ\text{C}$  in RNase-free ddH<sub>2</sub>O. Vortex and centrifuge before use.

### 2.3 ddPCR Components and Equipment

1. 20 $\times$  assay mix from Subheading 2.1.
2. One-Step RT-ddPCR Advanced Kit for Probes (Bio-Rad). Store at  $-20^\circ\text{C}$ . Invert to mix before use.
3. Droplet Generation Oil for Probes (Bio-Rad). Store at room temperature.
4. Droplet Reader Oil (Bio-Rad). Store at room temperature.
5. DG8™ Cartridges for droplet generation (Bio-Rad).
6. DG8™ Gaskets for droplet generation (Bio-Rad).
7. QX200™ Droplet Digital PCR System: droplet generator and cartridge holders, droplet reader, and QuantaSoft reader software (Bio-Rad).
8. Rainin multichannel pipettors and corresponding tips for pipetting 20  $\mu\text{L}$  and 40  $\mu\text{L}$  volumes (*see Note 4*).
9. Half-skirted Eppendorf 96-well PCR plates for droplet thermal cycling and reading.

10. Pierceable, heat-sealable foil seals (e.g., Bio-Rad Pierceable Foil Heat Seal).
11. Plate sealer capable of sealing for 5 s at 180 °C (e.g., Bio-Rad PXI™ Plate Sealer).
12. Thermal cycler, preferably one capable of a 4–12 °C hold and a heated lid.

## 2.4 Web Resources

1. UCSC genome browser (hg19): <http://genome.ucsc.edu/cgi-bin/hgGateway>.
2. 1000 Genomes browser: <http://browser.1000genomes.org/index.html>.
3. SNAP (SNP annotation and Proxy Search): <https://www.broadinstitute.org/mpg/snap/>.
4. Genotype-Tissue Expression (GTEx) Project: <http://www.gtexportal.org/home/>.
5. IDT OligoAnalyzer: <http://www.idtdna.com/calc/analyzer>.
6. Primer3Plus primer design tool: <http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi> [7].

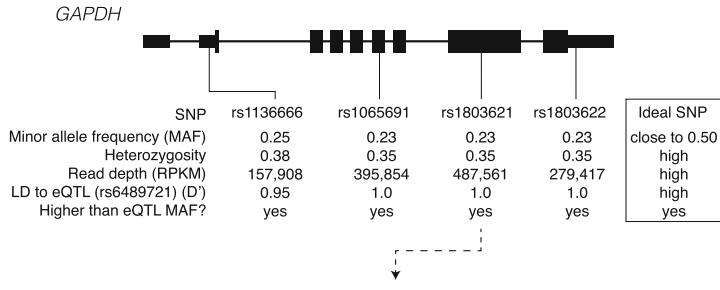
## 3 Methods

### 3.1 Selection of Reporter SNPs for Assays

The first step (*see* Fig. 1, Subheading 3.1) is to identify one or more reporter variants (transcribed heterozygous sites) in the gene of interest, which is generally chosen through hypotheses about the mechanisms of disease. These reporter variants are generally single nucleotide polymorphisms (SNPs) within the target gene's RNA transcript. The chosen SNP will be used just as a reporter of a transcript's chromosome-of-origin; it need not be the variant causing the allelic skew, nor necessarily correlated via linkage disequilibrium with that variant, though it may be useful in downstream analysis to know the relationship between the reporter and causal SNP.

Using the UCSC genome browser, enter the gene name of interest and select the assembly hg19 to focus on that locus. Under the “Variation” header at the bottom of the page, change the “1000G Ph1 Vars” track display setting to “full” and click “refresh.” Any exonic variant appearing in this track, as well as those in the 5'UTR and 3'UTR, can be used (*see* Fig. 2). In addition, intronic SNPs can sometimes be utilized, especially if RNA-seq data indicates that the corresponding sequence is detected at a meaningful level in total RNA (*see* Note 5), reflecting that sequence's presence in cells for an appreciable period after transcription and before splicing.

### 3.1 Marker SNP selection



### 3.2 Assay design

#### 1. Obtain the DNA sequence.

```
TGAAGCAGGCCTGGAGGGCCCCCTCAAGGGCATCTGGGTACACTGAGCACAGGTGG
TCTCTCTGACTTCAACAGCGACACCCACTCTCACCTTTGACGCTGGGCTGGCATTTG
CCCTAACGACCACTTTGCAAGCTCATTTCTGTAIGTGGCTGGGCAAGAGACTGGC
TCTTAAAAGTGAGGCTGGGCCCTCTGGCTGGCTGGCTAGAAAAGGGCCCTGACA
ACTCTTTTTCATTTCTAGGTATGACAACGAATTGGTACAGCAACAGGGTGGTGGACCT
```

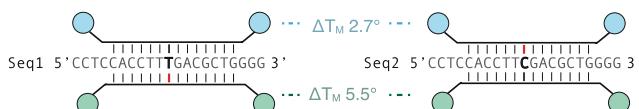
#### 2. Modify the DNA sequence – mask introns and SNPs.

```
TGAAGCAGGCCTGGAGGGCCCCCTCAAGGGCATCTGGGTACACTGAGCACAGGTGG
TCTCTCTGACTTCAACAGCGACACCCACTCTCACCTTTGACGCTGGGCTGGCATTTG
CCCTAACGACCACTTTGCAAGCTCATTTCTGTAIGTGGTACAGCAACGAATTGGCTAC
AGCAACAGGGTGGTGGACCT
```

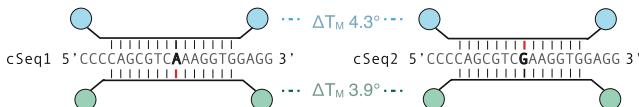
#### 3. Determine which strand the probes should bind.

Reference allele	Alternate allele
Seq1 5' CCTTACACCTT <del>T</del> GACGCTGGGG 3'	Seq2 5' CCTTACACCTT <del>C</del> GACGCTGGGG 3'
cSeq1 3' GGAGGTGGAA <del>A</del> CTCGGACCCC 5'	cSeq2 3' GGAGGTGGAA <del>G</del> CTCGGACCCC 5'

#### Probe Configuration 1



#### Probe Configuration 2

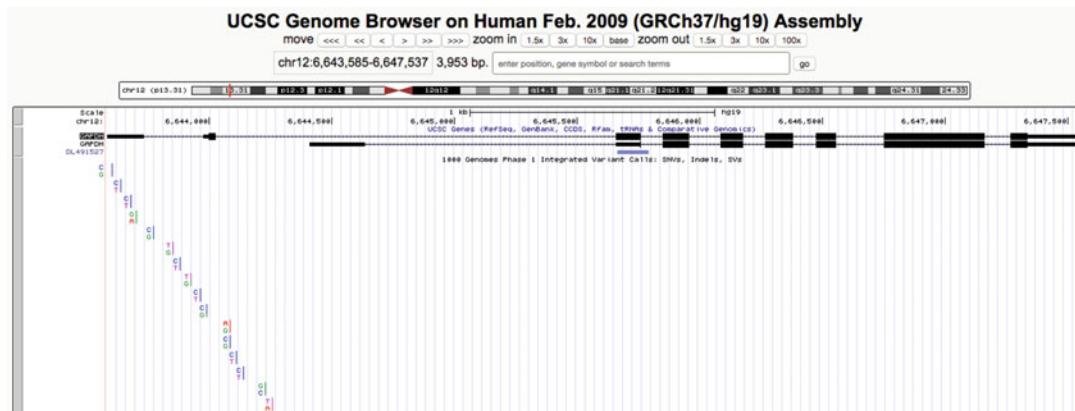


#### 4. Design Assay.

Probe Configuration 2 has more balanced ΔT<sub>M</sub>'s. Design assay using reverse complement of obtained DNA sequence.

```
AGGTCCACCACTCTGTTGCTGAGCCAATTCTGTTGATACATACCGGAAATGAGCTT
GACAAAGTGGTGTGAGGGCACTGCGAGCCCCAGCGTCAAGGTTGGAGGAGTGGGTGTC
GCTGTTGAAGTCAGAGGAGACCACTGTTGCTCAGTGAGCCAGGATGCCCTGAGGGG
GCCCTCGACGCGCTGCTTCA
```

**Fig. 1** A schematic following the steps of assay design for the *GAPDH* gene, as referenced sequentially in the chapter text



**Fig. 2** A screenshot of the UCSC genome browser

1. Find the parameters needed for evaluating potential reporter SNPs. There are two main components to consider: the number of individuals that are heterozygous for the reporter SNP and the abundance of transcription at that SNP. Note that due to the evolutionary pressure on coding sequence variation, many genes may have only one or two common, high read-depth SNPs that are amenable for evaluating allelic skew.
  - (a) Because homozygotes cannot be evaluated, determine or estimate the number of heterozygotes among the available RNA-sample donors. If genotype data are already available, this can be measured directly. If not, heterozygosity can be estimated from population-level data from a 1000 Genomes Project population sample whose ancestry matches your study population. Obtain the minor allele frequency (MAF) of the SNP using a source, such as the 1000 Genomes browser, and calculate the heterozygosity— $2 \times (\text{MAF}) \times (1-\text{MAF})$ , which is derived from the Hardy–Weinberg equilibrium value  $2pq$ . The greater the heterozygosity, the more samples will be usable.
  - (b) Determine or estimate the level of expression of the potential reporter SNPs; this will be useful for determining the amount of RNA input and selecting a SNP that is substantially expressed, despite alternative splicing or potentially being in an intron (*see Note 5*). The average number of sequencing reads, termed read depth and often reported as RPKM or FPKM (reads or fragments per kilobase per million reads), that contain the exon/intron in which the reporter SNP resides can be used as a rough proxy for expression. The read depth can be found for any exon across a range of tissues at the GTEx website by searching for a gene and positioning the mouse cursor over the exon in the row corresponding to the desired

tissue type. The higher the read depth, the more the exon containing the reporter SNP is expressed, and the more informative that SNP will be given less input RNA.

- (c) **(Optional)** If there is a prior hypothesis that a particular functional variant is generating allele-specific expression differences, then the linkage disequilibrium between this variant and any potential reporter SNPs should be calculated using the 1000 Genomes browser or SNAP in the population most genetically similar to that of the RNA-sample donors (e.g., samples from donors of Caucasian ancestry should be evaluated using European-ancestry populations such as CEU). A reporter SNP with a high D' (a measure of recombination) and an MAF greater than the candidate functional variant's MAF will be most helpful for validating that candidate and determining its direction of effect (*see Subheading 3.5, step 2 and Note 6*).

2. Determine which potential reporter SNP(s) will be the most informative. A general guideline for ranking these SNPs is (heterozygosity)×(expression\_level), which assigns roughly equal weighting to the number of heterozygotes, which influences sample size, and expression level, which influences measurement precision. With ample RNA (>20 ng) and moderate to high expression of the gene (>10 RPKM), the number of heterozygous samples available will be more important than the level of expression. Conversely, when a gene is expressed at low levels or a smaller quantity of usable RNA is available for analysis, it may be more important to select variants with higher expression levels. Select the top reporter SNP(s) based on these criteria. If there are multiple suitable reporter SNPs, it is useful to select two or more with low D' to each other, to increase the likelihood of having a reporter SNP with high D' to the functional variant (*see Subheading 3.1, step 1c*).

### **3.2 Assay Design**

1. Obtain the DNA sequence flanking the reporter SNPs for assay design (*see Fig. 1, Subheading 3.2, step 1*). Enter the SNP coordinates into the UCSC genome browser and navigate to that genomic region. The browser should be zoomed in on that variant.
  - (a) Obtain the surrounding DNA sequence by selecting View > DNA from the tool bar.
  - (b) Under the options, add 100 extra bases upstream and 100 extra bases downstream.
  - (c) Check the “Mask Repeats” option and select “Mask to N.” This blocks the primer-design software from using

sequences that are repeated many times in the human genome.

- (d) Hit the “extended case/color options” button and check both the “Underline/Human mRNAs” option and the “Bold/Common SNPs” option.
- (e) Hit submit.

The resulting sequence will be 201 bp long, with the target SNP in the middle, all SNPs in bold, and the exonic mRNA sequence underlined. If not all of the sequence is underlined, increase the “extra bases” parameter (possibly up to 1000s of bases) until a total of 100 bases on either side of the SNP are underlined. Repeat-masked sequence does not count in this total.

2. Change the DNA sequence to exclude introns and SNPs (*see Fig. 1, Subheading 3.2, step 2*). Unless designing a reporter SNP assay to an intronic SNP, remove the intronic sequence (the sequence that is not underlined), leaving the mRNA transcript. To prevent the primer-design software from designing primers that bind to polymorphic bases, change all the bolded SNPs (except the target SNP) to N (*see Note 7*).
3. Determine which strand the probes should bind (*see Fig. 1, Subheading 3.2, step 3*). Selection of the best strand (plus or minus strand) for an assay increases the extent to which the correct allele-specific probe will outcompete the other probe when amplifying the target allele and is done by balancing each probe’s  $\Delta T_M$  (delta melting temperature = correct match  $T_M$ —mismatch  $T_M$ ). Thus, the  $\Delta T_M$ ’s for a pair of probes binding to one strand of both allelic sequences will be compared to a pair of probes binding to the other strand. To determine all possible  $\Delta T_M$ ’s, select the sequence from 10-bp upstream to 10-bp downstream of the reporter SNP (Seq1). Copy/paste the sequence and change the reporter SNP to its alternate allele (Seq2). Obtain the reverse complement sequence for both (cSeq1 and cSeq2). Use IDT OligoAnalyzer with these settings for all following tests:
  - (a) Target type: DNA.
  - (b) Oligo Conc: 0.05  $\mu$ M.
  - (c)  $\text{Na}^+$  Conc: 50 mM.
  - (d)  $\text{Mg}^{++}$  Conc: 3 mM.
  - (e) dNTPs Conc: 0 mM.

Enter Seq1 and select the Tm Mismatch tool. Change the generated complementary sequence to cSeq2. Record the DELTA  $T_M$ . Repeat for Seq2, cSeq1, and cSeq2 with editing the generated complementary sequence to cSeq1, Seq2, and Seq1 respectively. If the DELTA  $T_M$  values when entering Seq1

and Seq2 are more equal than those for cSeq1 and cSeq2, proceed with the sequence already obtained. If cSeq1 and cSeq2 are more equal, then proceed with the reverse complement of the 201-bp mRNA segment from the previous step. If there were multiple potential reporter SNPs, prioritize on those with a higher sum of DELTA  $T_M$  for Seq1 and Seq2 or cSeq1 and cSeq2.

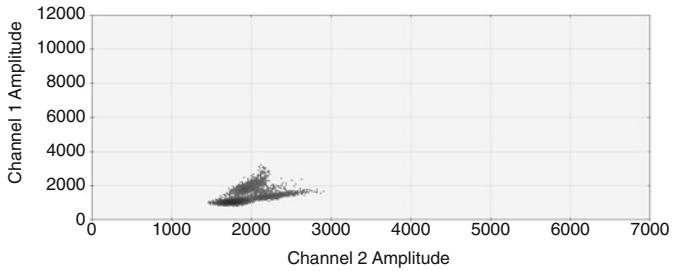
4. Design the primers and probe to the reference allele using the Primer3Plus primer design tool. Enter the 201-bp mRNA sequence obtained in Steps 1–3 into the box at the top of the webpage. Check “Pick hybridization probe (internal oligo)” under the input sequence. Under “General Settings,” click the “Mispriming/repeat library” pull-down menu above the sequence box and choose “HUMAN.” The following conditions are for an assay that will be run with an annealing temperature around 60 °C, but depending on local sequence surrounding the SNP (for example, when this sequence is AT-rich), it may be necessary to use a different annealing temperature (*see Note 8*).
  - (a) Under “General Settings” set:
    - The Primer Size as Min: 18, Opt: 20, Max: 27.
    - The Primer Tm as Min: 57, Opt: 60, Max: 63.
    - The Primer GC% as Min: 20, Opt: 50, Max: 80.
    - The Concentration of divalent cations to 3.8.
    - The Concentration of dNTPs to 0.8.
  - (b) Under “Advanced Settings” set:
    - The Table of thermodynamic parameters to “SantaLucia 1998”.
    - The Salt correction formula to “SantaLucia 1998”.
  - (c) Under “Internal Oligo” set:
    - The Hyb Oligo size as Min: 15, Opt: 20, Max: 27.
    - The Hyb Oligo Tm as Min: 64, Opt: 65, Max: 70.
    - The Hyb Oligo GC% as Min: 20, Opt: 50, Max: 80.
    - The Hyb Oligo Mishyb Library to “HUMAN”.
  - (d) Force Primer3Plus to design the probe over the target SNP by setting the Hyb Oligo Excluded Region to “1,80 121,80”.
  - (e) Click “Pick Primers” and select a probe and primer set in which the probe does not start with a G (5' G) and the probe overlaps the mutation of interest with at least four bases on either side.

- (f) If no assay designs pass these design restrictions, relaxing the parameters can be attempted. However, designs requiring significant changes may require trying a different reporter SNP if one is available (*see Note 9*).
- 5. Design the probe to the alternate allele. Within the mRNA sequence, change the target SNP to the alternate (nonreference) allele. Copy and paste the left primer from Step 4 into the box below “Pick left primer, or use left primer below:” and copy and paste the right primer from Step 4 into the box below “Pick right primer, or use right primer below.” Using the same primer design settings as Step 4, select a probe that does not start with a 5' G and overlaps the mutation of interest with at least four bases on either side. Note, succeeding in the design of the first allele does not guarantee success with the other allele.
- 6. Ensure that the assay targets the correct region and is not predicted to have off-target amplification. Use the UCSC BLAT tool (under the “Tools” menu at the top of the page) to check that the two primers and reference-allele probe match only the region of interest perfectly (*see Note 10*). View the region with the “Common SNPs” track displayed to ensure the primers and probes do not bind over other SNPs. In addition, check whether the primers and probes are likely to bind each other by using the “Hetero-Dimer” option in the right menu bar of IDT’s OligoAnalyzer; ΔGs lower than –7 should be avoided.
- 7. Order the primers and probe from your usual oligonucleotide supplier. For example, the 5' FAM and 5' HEX probe fluorophores, both with the Iowa Black 3' ZEN quencher from Integrated DNA Technologies. When making the 20× assay mix, please note that the proportion of primers to probes is different than in qPCR.

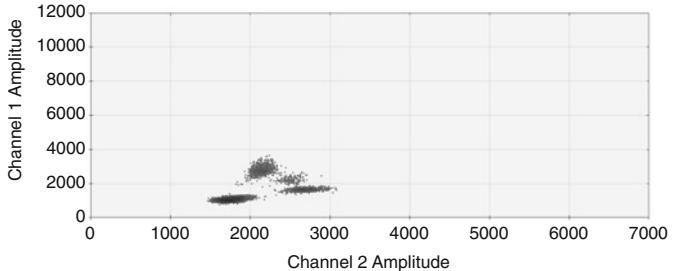
### **3.3 RT-ddPCR for Determining Allelic Imbalance**

1. **(Optional)** Before running the assay on samples of interest, and particularly if the DELTA TM values from Subheading 3.2, step 3 were low, it may be useful to run trial reactions with varying annealing temperatures using heterozygous genomic DNA (*see Fig. 3*) and with varying RNA inputs to optimize cluster separation (*see Notes 16 and 17*).
2. Thaw the RNA and the reagents of the One-Step RT-ddPCR Advanced Kit for Probes on ice for 30 min. Vortex and centrifuge briefly.
3. For each reaction, assemble:
  - 6.25 μL Supermix (One-Step RT-ddPCR Advanced Kit for Probes).

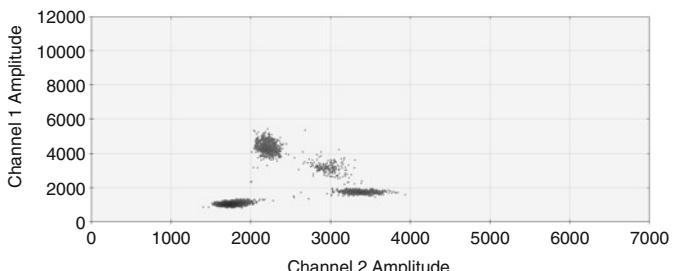
a. 60°C Annealing Temperature



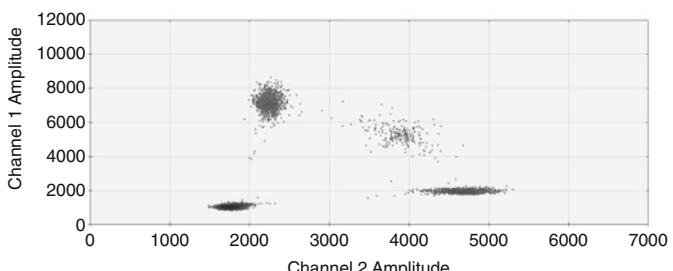
b. 58°C Annealing Temperature



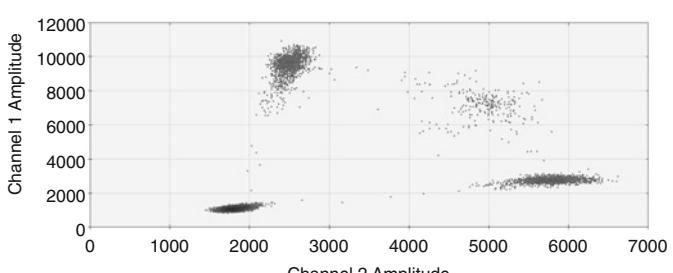
c. 56°C Annealing Temperature



d. 53°C Annealing Temperature



e. 50°C Melting Temperature



2.5 µL Reverse Transcriptase (One-Step RT-ddPCR Advanced Kit for Probes).

1.25 µL 300 mM DTT (One-Step RT-ddPCR Advanced Kit for Probes).

µL of 20× assay targeting the SNP.

variable input of RNA (100 fg–100 ng per reaction) (*see Note 3*).

<13.75 µL of DEPC-treated water.

25.0 µL total volume (*see Note 11*).

\*Some may find it helpful to run a control sample with heterozygous DNA to confirm where the droplets cluster. In addition, because the allelic imbalance statistics can be run using a single reaction, duplicates of a sample are not necessary unless the RNA is rare, necessitating combining several reactions to have enough positive droplets to make a definitive call.

4. Vortex and spin the plate to collect the liquid at the bottom of wells. Keep the plate protected from light until droplet generation, and allow the reactions to equilibrate to room temperature for 3 min prior to droplet generation (*see Note 12*).
5. Place a DG8 cartridge into the QX200 droplet generation cartridge holder (Bio-Rad).
  - (a) Pipette 20 µL of the PCR mix into the middle row of the cartridge. Only push down to the first stop when ejecting liquid to ensure there are no air bubbles in the sample (*see Note 13*). Using a Rainin multichannel pipettor with Rainin tips is preferred at this stage (*see Note 4*).
  - (b) Pipette 70 µL of oil into the bottom row of the cartridge using a new set of tips. Always be sure to pipette the oil after the samples. The top row (where droplets are outputted) is left empty.
  - (c) Place a DG8 rubber gasket over the cartridge.
6. Place the cartridge holder with cartridge and gasket into the QX200 droplet generator. Close the generator and droplets will be formed.
7. When the triangles on the button on the droplet generator lid return to being lit solid green, remove the cartridge. Carefully remove its gasket and transfer the droplets in the top row to a clean, half-skirted Eppendorf plate. It is important that the pipetting at this stage is slow and careful with the pipette

---

**Fig. 3** A temperature gradient run using the same sample and assay. Note how the clusters are intermingled in (a), yet increase in separation with increasing annealing temperature in (b), (c), and (d) until an optimal temperature is reached in (e). If this gradient were run with values below 50 °C, we would see the clusters start to intermingle again as we move away from the optimal annealing temperature

oriented at 45 degrees, otherwise the droplets may shear. Afterward, discard the gasket and cartridge (*see Note 14*).

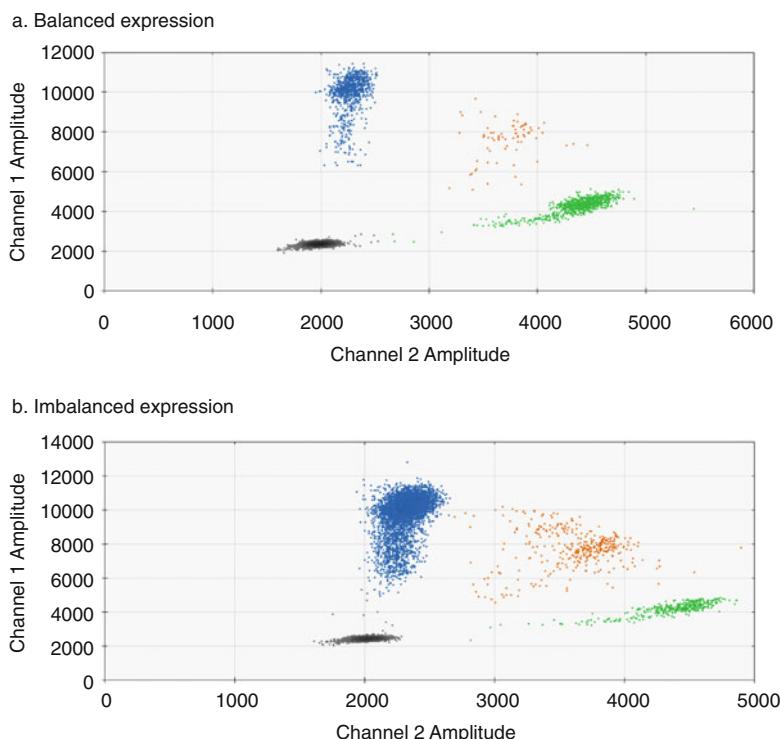
8. After all droplets are made, seal the droplet plate with a foil seal by heating the top to 180 °C for 5 s. Check to make sure that the outlines of the wells are visible through the foil, indicating a good seal.
9. Thermal cycle as follows:  
42–50 °C for 60 min (*see Note 15*).  
95 °C for 10 min.  
40 cycles of 95 °C for 30 s followed by 60 °C (or identified optimal annealing temperature; *see Note 16*) for 1 min.  
98 °C for 10 min.  
4 °C hold.  
Use a 2.5 °C per cycle ramp rate for all steps and set to 40 µL volume. Droplets can be stored protected from light at 4 °C after cycling for up to 24 h before reading.
10. Set up a template on the QX200 droplet reader computer. Open QuantaSoft. Under “Template” in the top left corner, select “New” in order to fill in a new plate map. Double click on the first well. In the “Sample” box, under “Experiment,” select any of the “ABS” experiments, and then, under “Supermix,” select “One-Step RT-ddPCR Kit for Probes.” In the “Target 1” box, enter the name of the FAM assay in the “Name” field and select “Ch1 Unknown” from the “Type” menu. In the “Target 2” box, enter the name of the HEX or VIC assay in the “Name” field and select “Ch2 Target” from the “Type” menu. Select all wells of the plate that will contain samples that are using the same assays. Click the blue “Apply” button in the top window. The name of each sample may also be entered. Once finished, click “OK” and save the template.
11. Read the droplets on the QX200 droplet reader. Put the plate into the plate holder of the QX200, then place the black plate holder on top and click the silver tabs on either side down into place, making sure the A1 well is in the top left corner. Close the lid of the QX200. In QuantaSoft on the QX200 computer, make sure the template created in Step 9 is loaded, then click “Run.” On the popup menu that appears, select “FAM/HEX” or “FAM/VIC,” depending on the pair of fluorophores used, then click “OK.”

### **3.4 Data Finalization and Quality Control**

1. After the ddPCR is complete, perform a well-by-well visual inspection of droplet clusters in QuantaSoft by clicking “Analyze” on the leftmost menu followed by “2D Amplitude.” Verify that there is clear separation between the positive and negative clusters for both the reference and alternate allele

fluorophores. Some droplets between the positive and negative clusters are acceptable, but if the clusters themselves overlap, it will undermine the accuracy of the measurements. If all the wells have bad cluster separation, first try optimizing the annealing temperature and RNA concentration (*see Notes 16* and *17*, Fig. 4). If the assay is still not generating clean data, redesign the assay to the same or a different reporter SNP (*see Subheading 3.1*).

- Evaluate and adjust the software's automatic calling for each droplet cluster in each well. Droplets in the top left corner of the 2D amplitude plot are FAM+, droplets in the bottom right corner are VIC+ (or HEX+), droplets in the top right corner are double positive, and droplets in the bottom left corner are double negative. Though ddPCR is usually highly accurate for assays targeting independent loci (as in CNV analysis, in which the two probes interrogate different nucleic acid sequences), for allele-specific expression, where the two sequences of interest are highly similar, the signal in the two fluorescence channels is often not independent, due to the probes sometimes binding the opposite allele. The lack of rectangularity in the cluster locations (in two-dimensional fluorescence space) may



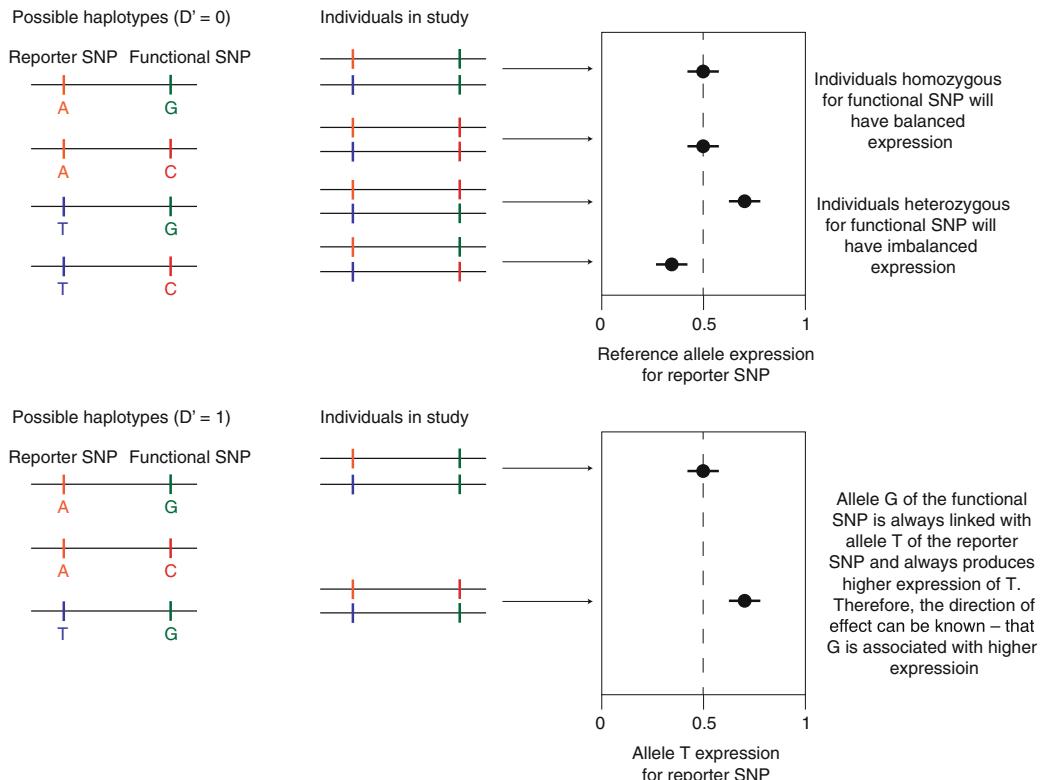
**Fig. 4** (a) A sample where the alleles (blue and green) appear to have equal expression. (b) A sample where the alleles have 90%:10% expression (blue:green)

complicate automated clustering and require interactive visual clustering by the user. To adjust cluster assignment in an unbiased fashion, highlight all the wells containing the same assay on the plate view in upper left and demarcate cluster boundaries identically across all samples on the 2D amplitude plot by using the “Threshold” or “Lasso” tools. If the “Status” column is set to “Check,” click anywhere in the amplitude plot to get the software to recognize the data (*see Note 18*).

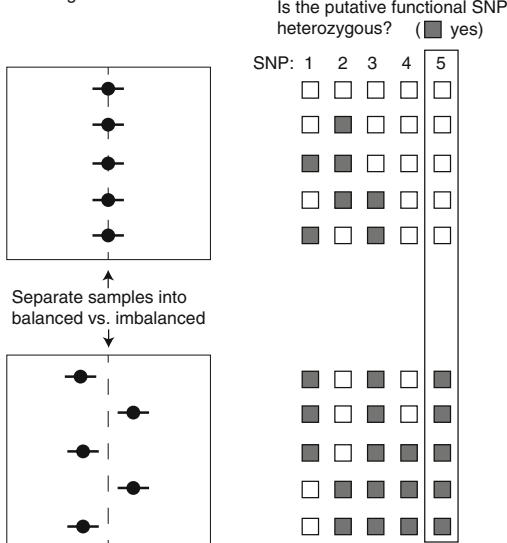
3. Export the data by selecting the wells and clicking “Export CSV.” The columns “Ratio” and “FractionalAbundance” correspond to different measures of allelic skew, (Allele A/Allele B) and Allele A/(Allele A + Allele B), respectively. The columns following these with labels such as “PoissonFractionalAbundanceMax” and “PoissonFractionalAbundanceMin” represent the upper and lower 95% confidence intervals for Allele A/(Allele A + Allele B), respectively (*see Note 19*).
1. Using the “PoissonFractionalAbundanceMax” and “PoissonFractionalAbundanceMin” values, we can determine which samples have significant allelic imbalance. Specifically, these values define the 95% confidence interval around the estimate of allelic imbalance recorded under “FractionalAbundance” and thus identify any samples that do not overlap 0.5 (which represents equal relative expression from each allele). The limit of detection in terms of effect size is, of course, dependent on the level of expression and RNA input, but for a moderately expressed gene with 10–20 ng of RNA (roughly 3000 or more positive droplets total), we can confidently determine effects down to a 10% difference in expression. Ideally, all imbalanced samples will display ratios with consistent absolute distances from 0.5, suggesting a single common variant affecting expression across the population. A wide variance in the magnitude of distance from 0.5 suggests multiple common influences acting on this gene, making each effect difficult to identify in isolation.
2. Map the functional SNP by correlating the presence of imbalance with each putative functional SNP’s heterozygosity. In short, samples that have balanced expression of the reporter SNP are homozygous for the functional variant(s), while samples that are imbalanced are heterozygous (*see Fig. 5a*). This can be used to map the functional variant by finding the regional SNPs that are always homozygous when there is balanced expression and heterozygous when there is imbalanced expression (*see Fig. 5b*) (*see Note 20*). If all the imbalanced samples are imbalanced in the same direction, this provides not only increased power to find the functional variant, but also allows the easy determination of which allele of the functional SNP is associated with higher expression (*see Note 21*). However, this

### **3.5 Using the Observed Allelic Imbalance to Map Functional Variants**

## a. Theory



## b. Finding the functional SNP



**Fig. 5 (a)** The possible haplotypes for a reporter and functional variant with a  $D'$  of 0 or 1, and the pattern of imbalance those configurations would yield. **(b)** The technique used to map the functional variant by correlating the presence of imbalance with heterozygosity

knowledge can also be obtained by determining the chromosomal phase molecularly in the DNA samples using methods we have previously described [8]. A chi square test should be used to determine the significance of the putative functional SNP.

---

#### 4 Notes

1. We have found that standard TaqMan probes generally perform well, but MGB or LNA probes may also be used. FAM and HEX probes can be ordered through Integrated DNA Technologies, and VIC probes can be ordered through Life Technologies. HEX and VIC fluorescence is read in the same channel in ddPCR, so either an HEX or VIC probe should be used with a FAM probe.
2. For ease of use, 20× mixes can be stored at 4 °C for short periods (~1 month).
3. Because the level of expression across different genes varies by orders of magnitude, there is not a single RNA input concentration that will work for all genes. We generally start with an input of 10 ng and adjust the concentration up or down until we obtain 10–40% positive droplets in ddPCR, which limits the likelihood of the clusters overlapping by limiting the size of those clusters, and provides enough double-negative droplets to confidently call clusters. This may require large amounts of RNA for genes with low expression. Note that this sample concentration is lower than most ddPCR assays.
4. Low-quality pipette tips can shred droplets during droplet generation, likely because they shed tiny pieces of plastic into the reaction. Though not strictly required, Rainin pipettors and tips are preferred when working with droplets.
5. At minimum, the SNP has to be within the transcript of the gene and preferably in an exon that is expressed in all isoforms (unless only one isoform is of interest). We have found that intronic SNPs can also be used, especially when they reside in long introns and are far from the 3' ends of such introns, ensuring slower degradation rates [9]. Assays on intronic SNPs always require a total-RNA sample. Such assays also tend to require more RNA input and are most feasible for highly expressed genes ( $\text{RPKM} > 25$ ). The GTEx website has expression information for introns as a downloadable file under the “Datasets” tab on “Download” page after registering for an account. The current release is “GTEx\_Analysis\_V4\_RNA-seq\_Flux1.6\_intron\_reads.txt.gz” at the time of writing. However, the data is given as read counts rather than equivalent RPKM on each gene page and needs to be further manipulated

for direct comparison. Briefly, divide each row by the length of the intron given in the first column, divide each column by the number of unique mapped reads in that library (found in “GTEx\_Data\_V4\_Annotations\_SampleAttributesDS.txt” under “SMMPPDUN” column label), and multiply by one billion (to correct for base/kilobase and read/(million read) conversion).

6. When the reporter SNP is more common than the proposed functional variant, particularly if  $D' = 1$ , it is possible to determine or validate the proposed functional variant by correlating the heterozygosity of the other variants to the allelic imbalance observed at the reporter. However, if the reporter SNP and actual functional variant have  $D' = 1$  and have similar allele frequencies (i.e., when  $r^2 = 1$ ), then there will be a limited set of possible genotypes and most samples heterozygous for the reporter SNP will have imbalance. Without any samples that are balanced and heterozygous, mapping is nearly impossible (because there is no balance/imbalance vs. heterozygosity to correlate). If prioritizing on finding a reporter SNP with  $D' = 1$  to the functional variant, then try to find a reporter SNP where the allele linked to the minor allele of the functional variant is twice as common as the functional variant’s minor allele frequency. For example, if the functional variant has an MAF = 0.05, then a reporter SNP with  $D' = 1$  and overlapping reporter SNP allele frequency of 0.5 will yield 50% usable reporter SNP heterozygotes, but with 45%:5% being balanced:imbalanced. A reporter SNP with  $D' = 1$  and overlapping reporter SNP allele frequency of 0.1 will yield fewer heterozygotes (18%), but more observations of imbalance (9%:9% being balanced:imbalanced), thus increasing the power of mapping.
7. SNPs in the primer or probe binding sites can prevent or hinder amplification. The sites of common SNPs must be excluded from the sequence used to design assays.
8. A temperature of 55 °C is another annealing temperature to consider for assay design at loci with lower nearby GC content. Use the following settings, particularly if design selection at 60 °C is difficult, while keeping the other settings constant:
  - (a) Under “General Primer Picking Conditions” set:
    - The Primer Size as Min: 18, Opt: 20, Max: 27.
    - The Primer Tm as Min: 55, Opt: 55, Max: 57.
    - The Primer GC% as Min: 20, Opt: 50, Max: 80.

- (b) Under “Internal Oligo (Hyb Oligo) General Conditions” set:
- The Hyb Oligo size as Min: 15, Opt: 20, Max: 27.
  - The Hyb Oligo Tm as Min: 56.5, Opt: 57.5, Max: 59.
  - The Hyb Oligo GC% as Min: 20, Opt: 50, Max: 80.
9. Generally the hybridization probe will be the source of design failure due to the constraint that the probe sequence must overlap the variant. To identify the specific requirement(s) that the surrounding sequence does not meet, enter the 21 bp sequence from Subheading 3.2, step 3 into the box below “Pick hybridization probe (internal oligo), or use oligo below” and click “Pick Primers.” The warning message at the top of the design will help troubleshoot the design. If problematic, parameter relaxing might first begin with decreasing “Hyb Oligo Max Mishyb” and increasing “Internal Oligo Max Poly-X,” both under the “Internal Oligo” tab. If other settings are leading to design failure, adjustments such as increasing annealing temperature range slightly can still lead to a usable assay, but consider designing assays for other reporter SNPs.
10. If the primers are too short for BLAT, use the in silico PCR function (“In-Silico PCR” under the UCSC genome browser “Tools menu”). If the region of interest is in a segmental duplication, one unique and perfect match may not be possible. Try choosing a different reporter SNP if the other duplications are also expressed.
11. Though the ddPCR instructions have 20 µL as the listed final reaction volume before droplet generation, scaling reaction volumes up by 25% ensures fewer air bubbles are introduced in Subheading 3.2, step 4a.
12. ddPCR droplets are made in sets of eight samples at a time and read in 96-well plate format, so it is easiest to set up the PCR in a 96-well plate.
13. Pushing the pipette down to the final stop introduces air bubbles, which compromise the number and quality of droplets. If air bubbles are introduced into the sample chamber, manually pop them with a clean pipette tip.
14. An automatic droplet generator (Bio-Rad QX200 AutoDG) and associated consumables can be used instead of manual droplet generation.
15. Reverse transcription temperature can be set to a higher temperature ( $\leq 50^{\circ}\text{C}$ ) to denature any secondary structure that may prevent conversion to cDNA, but using lower temperatures ( $\geq 42^{\circ}\text{C}$ ) can help reduce RNA degradation in solution.

16. Though assays are designed to work best at a particular annealing temperature ( $60^{\circ}\text{C}$ ), many SNP assays yield cleaner data at another temperature. Before running the assay on the samples of interest, run a set of reactions containing identical reagents and input sample on a gradient of melting temperatures (*see Fig. 4*). Select the melting temperature with the best cluster separation. An ideal plot will have tight clusters with ample room to draw separating thresholds. The clusters need not lie perfectly horizontal or vertical. In fact, it is expected that the probes will cross-react with the other allele, causing the trajectory of the single-positive droplets to form an acute angle, rather than the normal  $90^{\circ}$  right angle, but this will not affect the measurements as long as the four clusters are separate.
17. Because the correct identification of droplets containing one or both alleles depends on the competitive binding of the probes to sequences with which they are both similar, cluster separation can be more challenging. This is ameliorated by lower RNA inputs, which yield fewer double-positive droplets. If any two clusters overlap, run a range of reactions with decreasing sample input until there is clear separation. Allele-specific expression analysis benefits from using alleles as natural controls for each other, and will often require less sample input than total expression assays to obtain equivalent power in detecting the same regulatory effect. An RNA sample can also be assayed across multiple wells, with the resulting allele counts then combined across wells to increase the resolution/precision of the analysis.
18. The software only calls wells with 10,000 or more droplets and sets the “Status” column to “Check” for wells with fewer droplets.
19. “FractionalAbundance” and the related confidence interval defined by the min and max columns may better depict the magnitude of deviation as compared to the nonlinear distances between ratios. For example, if one allele generates 1 transcript to the other allele’s 2 transcripts, the “Ratio” column would read either 0.5 or 2, depending on the phase between the reporter and functional variants. Whereas “FractionalAbundance” would read either 33.3 or 66.6—depending on phase—and thus more intuitively depict the same effect size.
20. Some samples could have a chromosome containing a rare and/or recent recombination event that breaks the expected LD pattern seen in most individuals of similar ancestry. Because allele-specific imbalance is more resistant to variation arising from genetic and environmental *trans*-acting factors (*see 1*), a single sample can be confidently called as exhibiting allelic imbalance or not. Based on the allelic imbalance of that sample

resembling that of other samples with equivalent haplotypes on each side of the recombination event, this may suggest or eliminate sets of variants normally segregating on the same haplotype.

21. If all samples with significant allelic imbalance are imbalanced in the same direction, then we automatically know which allele is expressed more highly. This is because the reporter SNP and the functional SNP have a  $D' = 1$ . Effectively, of four possible combinations between two variants, a  $D' = 1$  implies that one or two of them are never seen. The reporter SNP heterozygotes have only two possible genotypes (see Fig. 5a), resulting in either balance or imbalance. This is different from the case where  $D'$  does not equal 1, where two different genotypes result in imbalance and its not clear from genotypes alone which allele is associated with higher expression.

## Acknowledgments

Our understanding of allelic skew has greatly benefited from the work of Tom Mullen and Jim Nemesh in our lab. This work was supported by a grant from the National Human Genome Research Institute (R01 HG006855, to SAM).

## References

1. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A 106:9362–9367. <https://doi.org/10.1073/pnas.0903103106>
2. Genomes Project C, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA (2010) A map of human genome variation from population-scale sequencing. Nature 467:1061–1073. <https://doi.org/10.1038/nature09534>
3. Patwardhan RP, Hiatt JB, Witten DM, Kim MJ, Smith RP, May D, Lee C, Andrie JM, Lee SI, Cooper GM, Ahituv N, Pennacchio LA, Shendure J (2012) Massively parallel functional dissection of mammalian enhancers *in vivo*. Nat Biotechnol 30:265–270. <https://doi.org/10.1038/nbt.2136>
4. Cowles CR, Hirschhorn JN, Altshuler D, Lander ES (2002) Detection of regulatory variation in mouse genes. Nat Genet 32:432–437. <https://doi.org/10.1038/ng992>
5. Chen R, Mias GI, Li-Pook-Than J, Jiang L, Lam HY, Chen R, Miriami E, Karczewski KJ, Hariharan M, Dewey FE, Cheng Y, Clark MJ, Im H, Habegger L, Balasubramanian S, O'Huallachain M, Dudley JT, Hillenmeyer S, Haraksingh R, Sharon D, Euskirchen G, Lacroute P, Bettinger K, Boyle AP, Kasowski M, Grubert F, Seki S, Garcia M, Whirl-Carrillo M, Gallardo M, Blasco MA, Greenberg PL, Snyder P, Klein TE, Altman RB, Butte AJ, Ashley EA, Gerstein M, Nadeau KC, Tang H, Snyder M (2012) Personal omics profiling reveals dynamic molecular and medical phenotypes. Cell 148:1293–1307. <https://doi.org/10.1016/j.cell.2012.02.009>
6. Battle A, Mostafavi S, Zhu X, Potash JB, Weissman MM, McCormick C, Haudenschild CD, Beckman KB, Shi J, Mei R, Urban AE, Montgomery SB, Levinson DF, Koller D (2014) Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. Genome Res 24:14–24. <https://doi.org/10.1101/gr.155192.113>
7. Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JA (2007) Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res 35:W71–W74. <https://doi.org/10.1093/nar/gkm306>

8. Regan JF, Kamitaki N, Legler T, Cooper S, Klitgord N, Karlin-Neumann G, Wong C, Hodges S, Koehler R, Tzonev S, McCarroll SA (2015) A rapid molecular approach for chromosomal phasing. PLoS One 10:e0118270. <https://doi.org/10.1371/journal.pone.0118270>
9. Gray JM, Harmin DA, Boswell SA, Cloonan N, Mullen TE, Ling JJ, Miller N, Kuersten S, Ma YC, McCarroll SA, Grimmond SM, Springer M (2014) SnapShot-Seq: a method for extracting genome-wide, *in vivo* mRNA dynamics from a single total RNA sample. PLoS One 9:e89673. <https://doi.org/10.1371/journal.pone.0089673>



# Chapter 24

## Very Low Abundance Single-Cell Transcript Quantification with 5-Plex ddPCR™ Assays

George Karlin-Neumann, Bin Zhang, and Claudia Litterst

### Abstract

Gene expression studies have provided one of the most accessible windows for understanding the molecular basis of cell and tissue phenotypes and how these change in response to stimuli. Current PCR-based and next generation sequencing methods offer great versatility in allowing the focused study of the roles of small numbers of genes or comprehensive profiling of the entire transcriptome of a sample at one time. Marrying of these approaches to various cell sorting technologies has recently enabled the profiling of expression in single cells, thereby increasing the resolution and sensitivity and strengthening the inferences from observed expression levels and changes. This chapter presents a quick and efficient 1-day workflow for sorting single cells with a small laboratory cell-sorter followed by an ultrahigh sensitivity, multiplexed digital PCR method for quantitative tracking of changes in 5–10 genes per single cell.

**Key words** Single-cell, Cell-sorting, Gene expression, Transcripts, RNA quantification, Digital PCR, dPCR, Droplet digital PCR, ddPCR, Assay multiplexing, Radial multiplexing

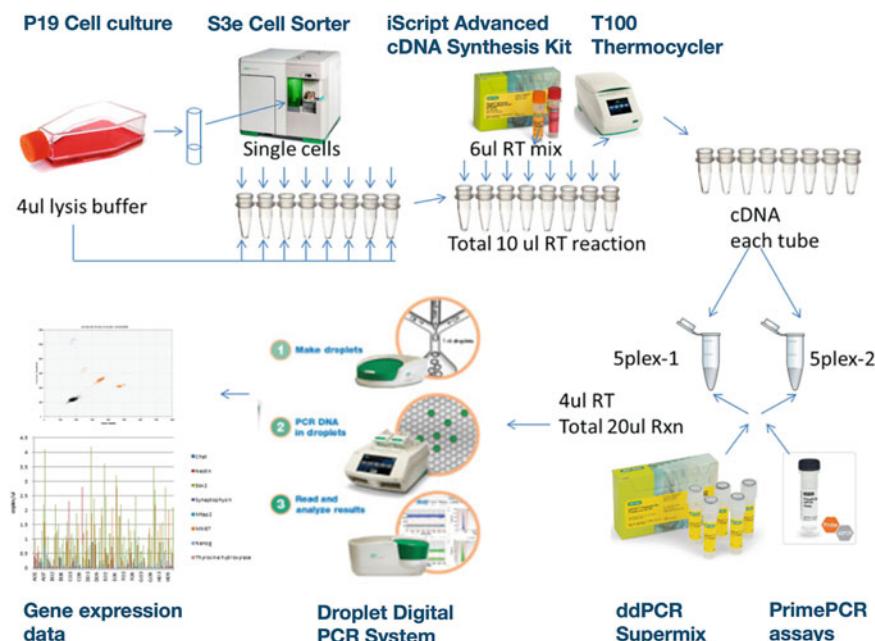
---

### 1 Introduction

Gene expression studies give valuable insight into the properties of cells and tissues and how these change in response to developmental signals and the environment. Methods to detect transcript levels of both known and unknown genes have changed dramatically over the past four decades becoming more efficient, more sensitive, more quantitative and enabling the profiling of larger numbers of gene transcripts at once. Historically, these studies have been done with northern blots [1, 2], nuclease-protection assays [3, 4], or PCR-based methods such as qRT-PCR [5, 6], targeting a limited number of gene transcripts, or DNA microarrays targeting much larger numbers [7–11]. Examples range across studies of bacteria, e.g., sporulation in *B. subtilis* [12], humans, e.g., discrimination of new breast cancer subtypes [13], and plants, e.g., light-induced development of the photosynthetic apparatus [14]. More recently, digital PCR (dPCR) and next-generation sequencing (NGS) have

been enlisted, in the former case to make absolute and highly sensitive measurements of transcript levels [15–17] and in the latter, to generate broad whole transcriptome or targeted RNA transcript profiles [18–20]. Although prominently expressed genes can be discerned in the complex mixture of cells comprising a tissue or even a whole organism, transcripts from minority cell populations such as stem cells are likely to be lost, especially for transcripts present at low abundance per cell.

Furthermore, since a major goal of gene expression studies is to identify genes cooperating to produce a phenotype of interest, the mixing together of different cell types in a single sample, or even transcriptional heterogeneity among apparently like cells, tends to confound the effort as it is unclear which transcripts are present in each individual cell (e.g., see Fig. 1, ref. 21). To address this challenge, various methods have been applied to isolate populations of single cells (e.g., by fluorescence-activated cell sorting [FACS], nanofluidic chambers, and most recently, droplets) and then to analyze either selected transcripts or the entire transcriptome of each cell by the methods mentioned above [22–25]. Each of the



**Fig. 1** A 1-day workflow for single-cell sorting followed by RT-ddPCR™. Cultured cells (shown) or dissociated cells are prepared as a suspension free of clumps and debris before using the S3e™ Cell Sorter to deliver a single-cell into each PCR tube well which contains lysis buffer. cDNA generated from the unpurified RNA lysate in each well is then divided into 2 ddPCR™ reactions per cell—one is assayed with 5-plex 1, the other with 5-plex 2. After converting all samples into droplets using the QX100™ or QX200™ Droplet Generator, followed by thermocycling, the plate(s) of endpoint ddPCR™ reactions are read in the Droplet Reader to quantify transcript levels for each of the 10 genes assayed

analysis methods has its strengths and drawbacks. Whereas NGS enables the most complete profiling of the transcriptome, it is not as sensitive as PCR-based methods—having capture efficiencies of only ~10% [24, 25]—and thus only detects transcripts present at >30–50 copies/cell. Unless unimolecular barcodes are used and sequencing is done to a much greater depth (>tenfold more than usual), it gives only relative quantification and is more costly and laborious. PCR-based methods while only detecting a smaller subset of genes per cell, can nonetheless detect transcripts present at only a few copies/cell in the present method; can flexibly and efficiently target any desired region in a transcript, no matter how far 5' it is from the 3' mRNA end; can utilize customized assay designs and thermal cycling conditions to more efficiently capture difficult transcript sequences (e.g., GC-rich regions) and transcript types (e.g., non-polyA-containing long noncoding RNAs); and is much quicker and less costly. Digital PCR, in particular, can provide the greatest sensitivity with absolute quantification of each gene transcript species being targeted, without a standard curve. Furthermore, because of the ability to multiplex and discriminate at least five different gene transcripts in a single droplet digital PCR reaction well (*see* ref. 26), it is possible to quantify 5–10 genes per cell at very high sensitivity without requiring preamplification of the single-cell cDNA as in all other methods. This both eliminates potential distortion of transcript quantification that can occur during preamplification and NGS library preparation, and provides a quick and efficient method for profiling key signature genes in populations of single cells.

Here we describe an ultrasensitive and rapid method for sorting single cells into PCR wells using the Bio-Rad S3e Cell Sorter, directly converting the unpurified RNA into cDNA, and utilizing a strategy of “radial multiplexing” to quantify the absolute number of transcripts per cell for ten genes via Bio-Rad’s QX200<sup>TM</sup> Droplet Digital<sup>TM</sup> PCR (ddPCR<sup>TM</sup>) System (*see* Figs. 1 and 2).

---

## 2 Materials

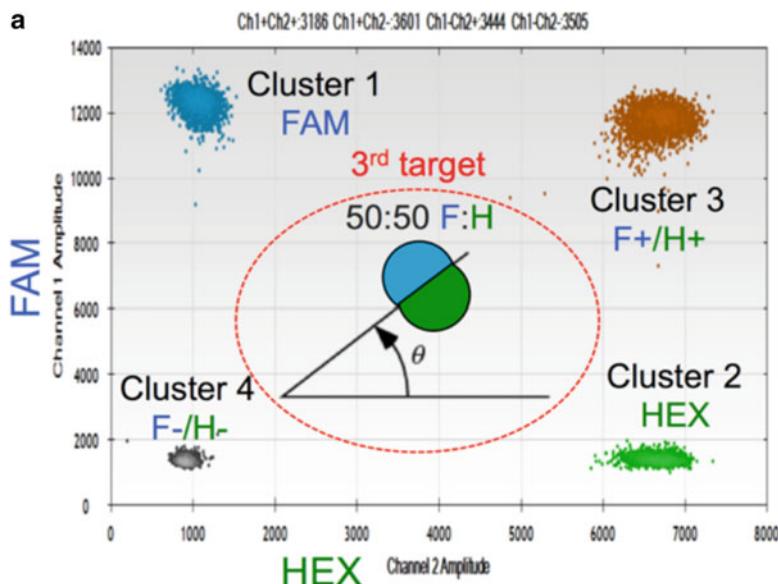
### 2.1 Single-Cell Preparation

1. For cultured cells, use appropriate growth medium, any necessary growth supplements and sterile plasticware, along with sterile hood and cell growth incubator. For P19 mouse embryonal carcinoma cells used here, normal-serum medium of Monzo et al. was used [27].
2. Pipettes and pipette tips.
3. 50 mL Falcon centrifuge tubes.
4. 0.25% Trypsin–EDTA (1×), phenol red solution for release of adherent cells from culture dish surface (e.g., ThermoFisher Scientific).

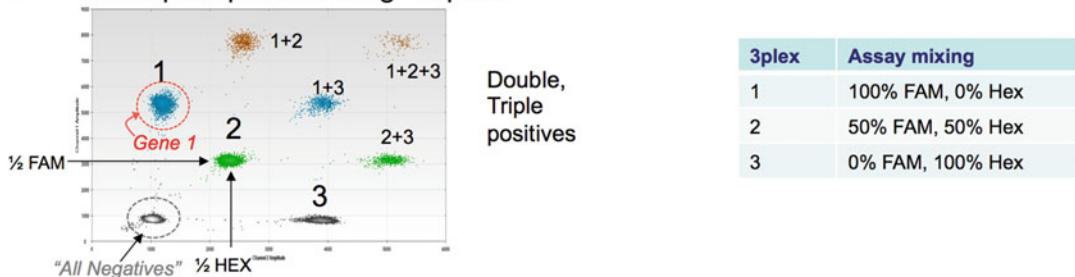
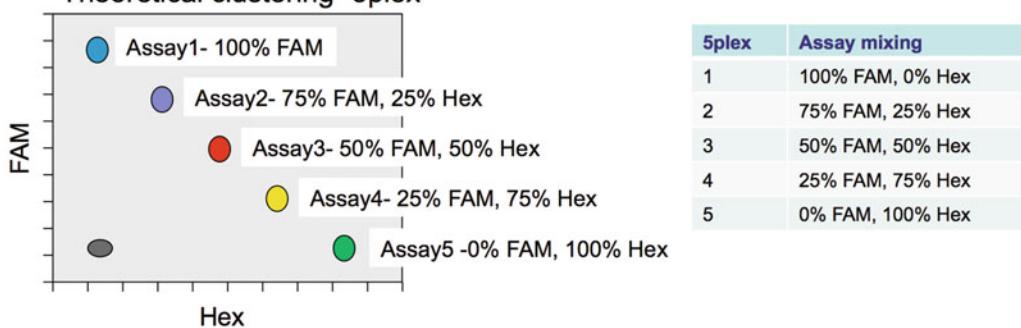
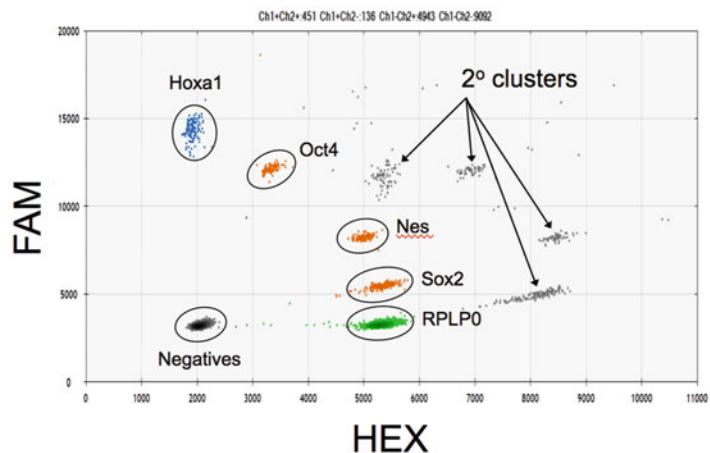
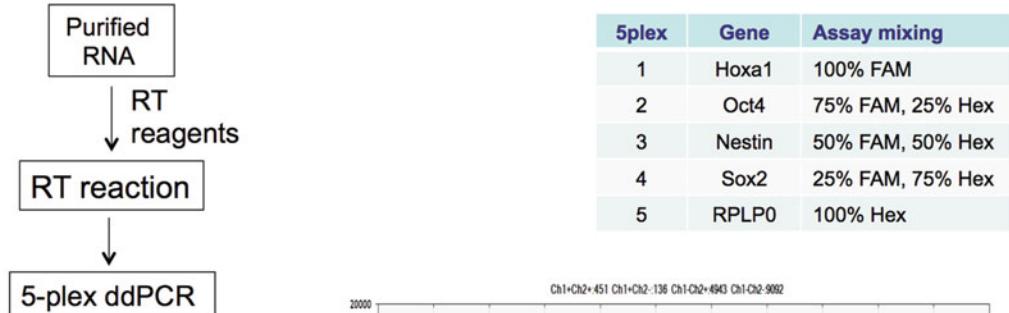
5. PBS (phosphate-buffered saline, pH 7.4).
6. BSA (20 mg/mL, Sigma).
7. 5 mL FACS tube with 40  $\mu$ m cell-strainer (Corning).
8. Clinical centrifuge (providing  $180 \times g$ -force).
9. Cell counter (e.g., Bio-Rad TC-20™ Automated Cell Counter)

## 2.2 Single-Cell Sorting

1. S3e™ Cell Sorter (Bio-Rad, Cat. No. 145-1001).
2. ProLine™ Calibration Beads (Bio-Rad, Cat. No. 145-1081).
3. ProLine™ Rainbow Beads (Bio-Rad, Cat. No. 145-1085).
4. 10% bleach solution.
5. Sterile distilled water.
6. PCR strip-tubes with caps (8 tubes/strip).



**Fig. 2** Radial multiplexing with ddPCR™ assays. (a) A 2D droplet plot of a standard duplex assay, w/ cartoon of a 3rd target along the 45° radius achieved by mixing together equal concentrations of FAM and HEX probes specific for that target. (b) Real Triplex assay with high concentrations of all targets. This results not only in primary clusters with droplets containing only a single type of transcript (1, 2 or 3), but also to secondary (1 + 2, 1 + 3, 2 + 3) and tertiary (1 + 2 + 3) clusters with two and three target types present, respectively. In practice, for single-cell and liquid biopsy applications, species concentrations are typically low and rarely occupy these higher order clusters. (See Fig. 4 for examples of individual single-cell 2D droplet plots). (c) Cartoon showing the concept of radial multiplexing extended to five quantifiable targets per ddPCR™ well using a continuum of probe mix ratios, from 100% FAM to 100% HEX. (d) 5plex-2 (assaying a different set of genes than 5plex-1) tested with purified bulk cell RNA from a mouse cell line, P19 embryonal carcinoma (EC) cells. Note that only with this more abundant amount of input RNA are there a noticeable number of droplets with more than a single transcript type. Compare with single cell transcript profiles in Fig. 4 which rarely have any 2° clusters

**b Concept of probe mixing - triplex****c Theoretical clustering - 5plex****d****Fig. 2** (continued)

7. DNA Suspension Buffer (10 mM Tris-HCl pH 8.0 with 0.1 mM EDTA);
8. 10% Triton X-100.
9. RNaseOUT (40×; Thermo Fisher Scientific).
10. Vortexer.

### **2.3 cDNA Synthesis**

1. Microcentrifuge (with 8-strip tube rotor).
2. iScript™ Advanced cDNA Synthesis Kit (Bio-Rad).
3. T100 (or C1000 Touch) Thermal cycler (Bio-Rad).

### **2.4 ddPCR™ Reactions**

1. QX100™ or QX200™ Droplet Digital™ PCR System, with Manual Droplet Generator and Automated Droplet Reader (Bio-Rad).
2. Rainin 8-channel pipettors and pipette tips (*see Note 1*).
3. ddPCR™ Supermix for Probes (no dUTP; Bio-Rad).
4. ddPCR™ Taqman™ 5' nuclease assays with FAM- and/or HEX-labeled probes.
5. Droplet Generation Oil for Probes (Bio-Rad).
6. DG8™ Cartridges and Gaskets (Bio-Rad).
7. Pierceable Foil Heat Seals (Bio-Rad).
8. 96-well semiskirted ddPCR™ plates (Bio-Rad, Cat. No. 12001925).
9. PX1™ PCR Plate Sealer (Bio-Rad).
10. QuantaSoft™ Analysis Software 1.7.4.0917 or earlier versions, or QuantaSoft™ Analysis Pro Software.

## **3 Methods**

The present method involving probe multiplexing has been developed and tested with Bio-Rad's QX100™ or QX200™ Droplet Digital™ PCR system though, in principle, the approach should also be deployable on other digital PCR platforms capable of detecting at least two distinct fluorescence emission wavelengths [28–32]. The Bio-Rad QX systems detect fluorescence in two spectral regions well suited for FAM- and HEX-(or VIC-)labeled Taqman probes, and in the simplest duplex assay format enables the discrimination of two distinct DNA or RNA targets per reaction well, where one probe is labeled in FAM and the other in HEX, each hybridizing to different target sequences. When both targets are abundantly present in a sample (as illustrated in the 2D droplet plot in Fig. 2a), this results in discrete clusters of droplets representing the presence of a first target (e.g., cluster 1, containing FAM signal only), a second target (e.g., cluster 2, containing

HEX signal only), both targets (cluster 3, containing both FAM and HEX signals and their corresponding target molecules), or neither target (cluster 4). By introducing a 50:50 mix of FAM and HEX probes, both directed against a third target sequence, it can be detected in a primary cluster located along the diagonal (or a 45° radius) as shown in Fig. 2a. An example of such a triplex assay is shown in Fig. 2b where a sufficiently high concentration of the 3 targets leads to primary clusters with a single target type in each (1, 2, or 3) closest to the origin, secondary clusters containing two targets each (1 + 2, 1 + 3, and 2 + 3) further from the origin, and a tertiary cluster containing all three targets (1 + 2 + 3) furthest out of all. Fig. 2c cartoons the application of this probe mixing strategy to five targets (a 5-plex) where ratios between 50:50 and 100% (75:25 and 25:75, FAM:HEX) are also used to position the fourth and fifth targets' primary clusters along radii (hence, the term “radial multiplexing”) between the vertical and diagonal (Assay 2) and the diagonal and the horizontal (Assay 4). And finally, an example of a 5-plex assay is shown in Fig. 2d.

### 3.1 Formulating and Optimizing a 5-Plex Droplet Digital PCR Assay

A successful 5-plex ddPCR™ assay requires identification of five individual assays, all with a similar annealing temperature, which do not interfere with one another when combined into a single ddPCR™ reaction. When formulating a 5-plex ddPCR™ assay, follow these steps.

1. Obtain singleplex ddPCR™ Taqman 5' nuclease probe assays with common annealing temperatures (e.g., 60 °C). These may either be from predesigned 20× assays optimized for ddPCR™ analysis (“<https://www.bio-rad.com/digital-assays/#/>”) or as five sets of primers and their corresponding probes designed against the transcript target sequences of interest according to best practices (*see* Bio-Rad Bulletin #6407 and Note 2) and synthesized by an oligo manufacturer (e.g., IDT, Coralville, IA or BioSearch, Petaluma, CA). We recommend ordering them first with a single probe fluorophore (e.g., four with FAM and one with HEX) and confirming their proper singleplex performance in ddPCR™ reactions, before ordering a second HEX-labeled probe for three of the component assays. Ultimately, eight probes will be necessary to detect the five transcripts of interests (*see* Fig. 2c), where three of them require a combination of both FAM- and HEX-labeled probes, one requires only a FAM probe, and one only a HEX probe. If formulating two 5-plex assays, design and test an additional five gene assays as for the first 5-plex.
2. Run a ddPCR™ thermocycler temperature gradient (as shown in Table 1) for each individual assay to test annealing/extension temperatures from ~5 °C below to ~5 °C above the calculated  $T_m$  of the primers (e.g., 55–65 °C) using cDNA from an appropriate bulk total RNA target (e.g., from cell

**Table 1**  
**ddPCR plate set up**

<b>ddPCR plate setup</b>												
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>
65 °C A	Gene 1 assay	Gene 2 assay	Gene 3 assay	Gene 4 assay	Gene 5 assay>	Gene 6 assay	Gene 7 assay	Gene 8 assay	Gene 9 assay	Gene 9 assay	Gene 9 assay	Gene 10 assay
63.4 °C B	Gene 1 assay	Gene 2 assay	Gene 3 assay	Gene 4 assay	Gene 5 assay	Gene 6 assay	Gene 7 assay	Gene 8 assay	Gene 9 assay	Gene 9 assay	Gene 10 assay	
62 °C C	Gene 1 assay	Gene 2 assay	Gene 3 assay	Gene 4 assay	Gene 5 assay	Gene 6 assay	Gene 7 assay	Gene 8 assay	Gene 9 assay	Gene 9 assay	Gene 10 assay	
60.6 °C D	Gene 1 assay	Gene 2 assay	Gene 3 assay	Gene 4 assay	Gene 5 assay	Gene 6 assay	Gene 7 assay	Gene 8 assay	Gene 9 assay	Gene 9 assay	Gene 10 assay	
59.2 °C E	Gene 1 assay	Gene 2 assay	Gene 3 assay	Gene 4 assay	Gene 5 assay	Gene 6 assay	Gene 7 assay	Gene 8 assay	Gene 9 assay	Gene 9 assay	Gene 10 assay	
57.8 °C F	Gene 1 assay	Gene 2 assay	Gene 3 assay	Gene 4 assay	Gene 5 assay	Gene 6 assay	Gene 7 assay	Gene 8 assay	Gene 9 assay	Gene 9 assay	Gene 10 assay	
56.4 °C G	Gene 1 assay	Gene 2 assay	Gene 3 assay	Gene 4 assay	Gene 5 assay	Gene 6 assay	Gene 7 assay	Gene 8 assay	Gene 9 assay	Gene 9 assay	Gene 10 assay	
55 °C H	Gene 1 assay	Gene 2 assay	Gene 3 assay	Gene 4 assay	Gene 5 assay	Gene 6 assay	Gene 7 assay	Gene 8 assay	Gene 9 assay	Gene 9 assay	Gene 10 assay	

lines or tissue of interest; Universal Mouse Reference RNA, Agilent Catalog #740100; or Universal Human Reference RNA, Agilent Catalog #740000). See Subheading 3.4 below for cDNA synthesis procedure, scaling the reaction shown there to 20 µL final volume, consisting of 12 µL of Reverse Transcription Master Mix (Table 3) and 8 µL of total RNA and water combined. See Subheading 3.5 below for set-up and running of ddPCR<sup>TM</sup> reactions, however using temperature gradient cycling conditions as illustrated in Table 1.

3. Select five assays that show optimal or near-optimal separation of positive and negative droplets at a similar annealing temperature for inclusion in the first 5-plex (see Note 3). If you are creating two 5-plexes, it is recommended to include one assay in common to the two multiplexes for more precise determination of transcript copies/cell from cDNA that is divided between two ddPCR<sup>TM</sup> reactions per cell. For each set of 5 component assays in a multiplex, you will need to order HEX probes with the same probe sequences as previously ordered with FAM, for three of the assays (see Note 4). Formulate a 5-plex ddPCR<sup>TM</sup> assay (as shown in Table 4) by adding assay 1 containing 100% FAM probe only, assay 2 with 75% FAM and 25% HEX probes, assay 3 with 50% of both FAM and HEX probes, assay 4 with 75% FAM and 25% HEX probes, and assay 5 with 100% HEX probe, each from a 20× stock with total probe concentration of 5 µM. If any of the assays did not perform well at the desired temperature (or at all), you may need to modestly redesign the assay for a higher or lower temperature or select an alternative design for the gene of interest.
4. Run ddPCR<sup>TM</sup> reactions for each 5-plex with cDNA from control bulk RNA—just as was done for the singleplex assays above—to check the separation of the clusters in 2D droplet plots. You can change the total probe concentration for an individual target to optimize the separation of the clusters (see Note 4).
5. For quantification of each individual transcript in each 5-plex, see Subheading 3.6 below. Compare the quantification of each transcript level in the 5-plex to its quantification in singleplex for the same control RNA target. Transcript copy numbers should be very similar. However, if there is some unexpected interference between any of the individual assays leading to significant disagreement between the singleplex and multiplex results (not commonly seen), it will be necessary to substitute at least one of the interacting assays. After achieving good cluster separation and accurate quantification in both multiplexes, the 5-plexes are now ready to use.

### 3.2 Preparation of Single Cells for FACS Sorting

The 5-plex single-cell gene expression protocol described in this chapter has been developed using adherent mammalian cells (P19 mouse embryonal carcinoma cells) grown according to Monzo et al. [27]. Alternatively, sources of single cells may be derived from suspension cell cultures; cells from blood, sputum or urine; or cells dissociated from tissue (*see Note 5*). However, although general guidance is given below for preparation of single cells ready for sorting, these other cell sources have not specifically been tested in this protocol.

1. First harvest cells by enzymatic release using 0.25% Trypsin-EDTA at 37 °C for 5 min, followed by quenching with media containing serum.
2. Centrifuge cells at  $180 \times g$  for 5 min at room temperature, discard supernatant, resuspend in 10 mL PBS and count cells with a cell counter (such as the Bio-Rad TC-20™ Automated Cell Counter).
3. Centrifuge cells at  $180 \times g$  for 5 min at room temperature, discard supernatant and gently resuspend pellet to a concentration of 1–10 million cells/mL in PBS with 0.1% BSA.
4. Filter cells into a FACS tube via a 40 µM cell strainer to obtain a single cell suspension (*see Note 6*). Keep on ice until sorting.

### 3.3 Flow-Sorting of Single-Cells into Lysis Buffer

Filtered single cells are sorted with the Bio-Rad S3e™ Cell Sorter into 8-well PCR strips containing Lysis buffer. Follow the manufacturer's instructions for cell sorter operations, referring to Bio-Rad Bulletin #10031105, if needed.

1. Prepare Lysis Buffer and Reverse Transcription Master Mix according to the recipes in Tables 2 and 3, respectively.
2. Power on the S3e™ Cell Sorter and perform a QC run using ProLine™ Calibration Beads, ensuring that it passes QC.
3. Perform a test sort with ProLine™ Rainbow Beads, making sure the beads are sorted into the center of the PCR tubes.
4. Clean the system with 10% Bleach at low pressure; clean the system again with sterile deionized water.

**Table 2**  
**Preparation of Lysis buffer**

Lysis buffer	1000 µL
DNA suspension buffer (10 mM Tris-HCl, pH 8, 0.1 mM EDTA)	965 µL
Triton X-100 (10%)	10 µL
RNAse OUT™ (40×)	25 µL

**Table 3**  
**Preparation of Reverse transcription master mix**

Reverse transcription master mix	1 ×	130 ×
5 × iScript advanced reaction mix	2 µL	260 µL
iScript advanced reverse transcriptase	0.5 µL	65 µL
Nuclease-free water	3.5 µL	455 µL
Total	6 µL	780 µL

5. Add 4 µL of prechilled Lysis Buffer into each PCR tube on 8-well PCR strips, and keep on ice before use.
6. Load the sample tube containing the filtered cells onto the sample station of the S3e<sup>TM</sup> sorter and collect data.
7. Set up gates for the target cells, making sure to exclude any cell debris and doublets.
8. In the ProSort software, click on “Sort Logic”: select “8-well strip”; choose “single sorting” mode, and set the sort limit to “1” to sort one single cell into each well (*see Note 7*).
9. Put 2 × 8-well PCR strips containing Lysis Buffer onto the two 8-strip tube adaptors in the sorting chamber.
10. Click on “Sort” to sort single cells into each well.
11. Remove both strips out of the sorter, cap the tubes and store them on ice.
12. Repeat steps 8–11, until single cells are sorted into as many as 12 strips (assuming 96 cells are to be collected).
13. Vortex the 8-well strips with sorted cells for 10 s and centrifuge briefly.
14. Leave the PCR tubes on ice for 5 min for complete lysis (*see Note 8*).

### 3.4 Preparation of cDNA

1. Add 6 µL of Reverse Transcription master mix (*see Table 3*) into each well; then place the caps back onto the tube strips.
2. Centrifuge the strips for a few seconds, vortex and centrifuge again.
3. Load PCR strips onto a thermal cycler\* and incubate under the following conditions:
  - (a) 42 °C for 30 min.
  - (b) 85 °C for 5 min.
  - (c) 4 °C (infinite).

\*Use heated lid set to 105 °C.

If you wish to stop here, the cDNA may be frozen at –20 °C at this step.

**Table 4**  
**Preparation of ddPCR™ master mix**

ddPCR™ master mix	1 well (1×)	1 plate (120×)	2 plates (240×)	per well (1.1×)
Water	1	120	240	1.1
2× Droplet Digital™ PCR Supermix for Probes (no dUTP)	10	1200	2400	11
20× assay 1 (100% FAM)	1	120	240	1.1
20× assay 2 (75% FAM:25% HEX)	1	120	240	1.1
20× assay 3 (50% FAM:50% HEX)	1	120	240	1.1
20× assay 4 (25% FAM:75% HEX)	1	120	240	1.1
20× assay 5 (100% HEX)	1	120	240	1.1
cDNA	4	480	960	4.4
Total	20 µL	2400 µL	4800 µL	22 µL

### **3.5 Setup and Running Droplet Digital PCR Reactions**

1. Prepare the ddPCR™ Master Mix according to Table 4 and store on ice.
- Prepare master mix for one or more plates, leaving out the cDNA which is added separately.
- If you are using assays rather than 20× assay, the final concentration is typically 900 nM for primers, 250 nM for probes.
2. Distribute 17.6 µL of the ddPCR™ Master Mix into each well of a 96-well PCR plate.
3. Briefly centrifuge the PCR strips containing the RT reaction. Open the caps carefully and transfer 4.4 µL of the cDNA into each well on the 96-well plate containing the ddPCR™ Master Mix. (Note: The remaining cDNA from each cell (also ~4.4 µL) may be frozen or prepared for running in a second ddPCR™ reaction plate with a second 5-plex assay panel as done for the first 5-plex).
4. Seal the plate. Centrifuge it for a few seconds, vortex and centrifuge again.
5. To generate droplets for each ddPCR™ reaction, see Bio-Rad Instruction Manual #10026322 (QX100™) or #10031907 (QX200™):
  - (a) Transfer 20 µL of each 22 µL sample—eight samples at a time, into eight sample wells of a DG8™ droplet generation cartridge (previously secured in a DG8™ cartridge holder).

- (b) Add 70 µL Probes Droplet Generation Oil into the eight companion oil wells.
  - (c) Seal with a gasket and place into the Droplet Generator.
  - (d) Follow manufacturer's guidance in observing proper loading of the DG8™ cartridge—*adding sample first* and avoiding introduction of air bubbles into the chamber well.
  - (e) Start droplet generation within 2' after loading of Droplet Generation Oil using the QX100™ or QX200™ Droplet Generator.
  - (f) Transfer droplets from the DG8™ cartridge output wells to 8 wells of a 96-well PCR plate using an 8-channel P50 pipettor. Aspirate slowly at an ~70 degree angle.
  - (g) Repeat as needed until up to all 96 wells are loaded with droplet reactions.
  - (h) Heat-seal the droplet plate with an adhesive foil seal using the PX1™ PCR Plate Sealer.
6. Load the droplet plate onto a thermal cycler and cycle using the protocol in Table 5.
7. After completion of thermal cycling, transfer the plate to the QX100™/QX200™ Droplet Reader to read both FAM and HEX fluorescence signals from all ddPCR™ wells, following the manufacturer's instructions (see Bulletin #10031906). After entry of the experiment plate layout is completed in QuantaSoft™ Software, select "Run" to begin the droplet reading process. Select the appropriate dye set used and run options when prompted.

**Table 5**  
**Thermal cycling protocol**

Step	Temperature	Time	Ramp	# of cycles
1	95 °C	10 min	2 °C/s	1
2	94 °C	30 sec	2 °C/s	40
3	xx°C*	1 min	2 °C/s	
4	98 °C	10 min	2 °C/s	1
5	4 °C	Hold	2 °C/s	1

\*The annealing temperature (Step 3) could vary depending on the optimal assay annealing temperature, though assays are typically designed for 60 °C

### 3.6 Quantification of ddPCR™ Reactions

After data acquisition, select all samples in the QuantaSoft™ well editor under “Analyze.”

1. Singleplex (or duplex) assays:

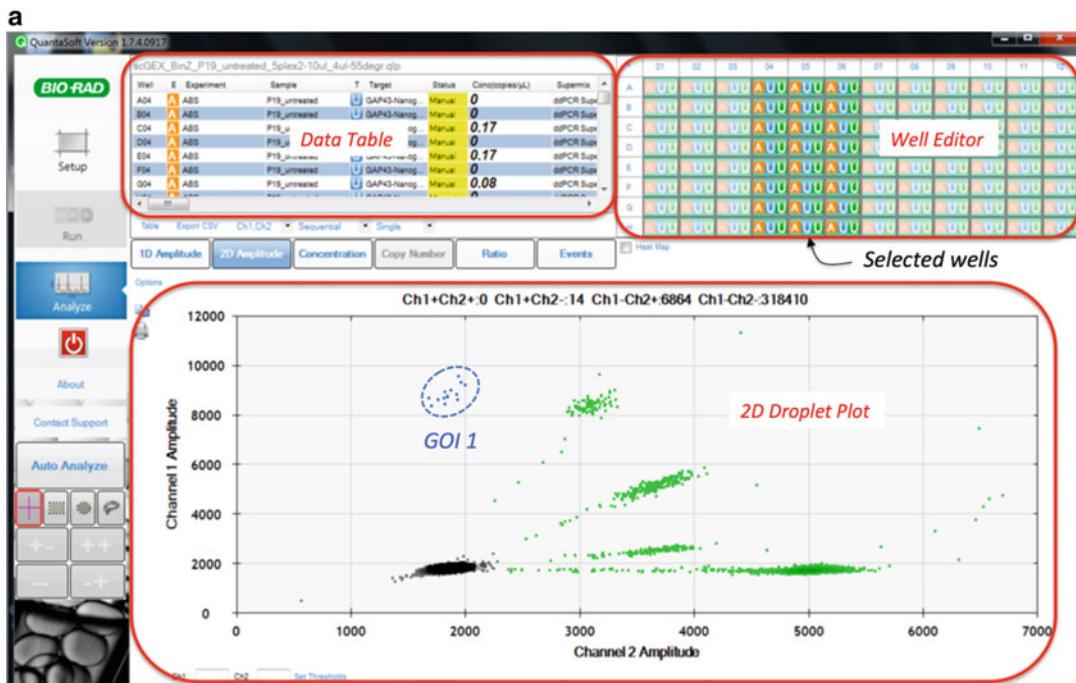
- (a) To evaluate assay quality, examine the automatic thresholding applied to the 1D (or 2D) amplitude data in the “1D Amplitude” or “2D Amplitude” tabs, and if necessary, set the thresholds or clusters manually. Negative and positive clusters should be well-formed and well-separated for assays that will be used to formulate the 5-plex(es) (*See*, e.g., Fig. 2a). If they are not of adequate quality at any of the annealing temperatures tested within the range of the thermal cycler temperature gradient, the assays will need to be redesigned and retested before proceeding.
- (b) The concentration reported in the Data Window or “Concentration” tab is “copies/µL” of the final 20 µL ddPCR™ reaction. For selecting the assays to combine into a 5-plex, be sure all of the assays perform well at a common annealing temperature (i.e., good cluster separation) and that the concentration value given for each assay is consistent over several degrees higher and lower than the annealing temperature selected for use in the 5-plex reactions. *See Note 2.*

2. 5-plex assays:

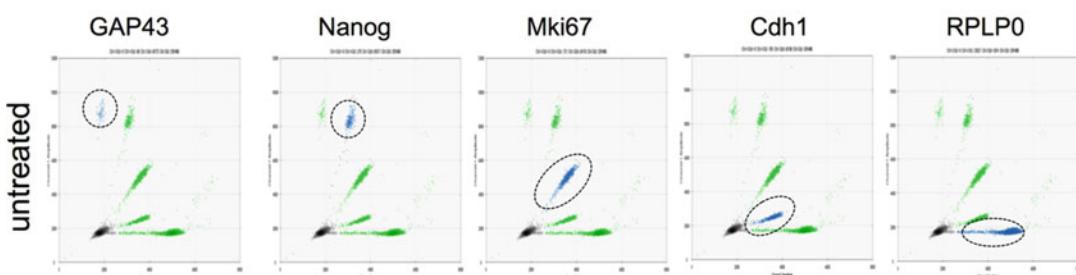
These higher multiplex reactions can be quantified by either of two procedures. If using the newer QuantaSoft™ Analysis Pro software version (QSAP™, available from the download tab at: <https://www.bio-rad.com/en-us/product/qx200-droplet-digital-pcr-system?ID=MPOQQE4VY>), all genes of interest (GOIs) in the multiplex can be quantified at once within the program after defining the target clusters properly under the “Advanced User Options.” This will provide concentrations and confidence intervals for all targets defined in a single analysis run. For further guidance, see Bio-Rad Bulletin #6827.

If using QuantaSoft™ 1.7.4.0917 or earlier software versions, it will be necessary to follow the iterative quantification procedure, below, which identifies one gene of interest at a time (in all wells selected on the plate) and provides the concentration (with 95% confidence interval) and number of transcript copies of that GOI in the 20 µL ddPCR™ reaction. These values can be graphed within QuantaSoft™ and may be exported as a “.csv” file for analysis in other software applications.

- (a) After 5-plex data acquisition, select all wells desired on the plate using the “Well Editor” (*see* Fig. 3a).
- (b) In the 2D droplet plot, using the cluster identification tool, encircle the *primary* (i.e., single) positive cluster for



**b** **positives; negatives, all other droplets** excluded from measurement



**Fig. 3** Quantification of 5-plex transcript levels with QuantaSoft™ software. **(a)** A window from QuantaSoft™ software showing the Well Editor (for choosing wells being analyzed in the other graph and data windows), 2D Droplet Plot (showing FAM and HEX fluorescent droplet intensities), and Data Table (showing various computed values and sample and assay descriptors). **(b)** Iterative quantification of transcript levels. Note that in each of the five 2D plots shown, only 1 primary cluster is labeled in blue and circled (i.e., the GOI being quantified in that iteration), only 1 cluster labeled in grey (complete negatives), and the rest are in green and are ignored for quantification purposes in the iteration shown in that plot

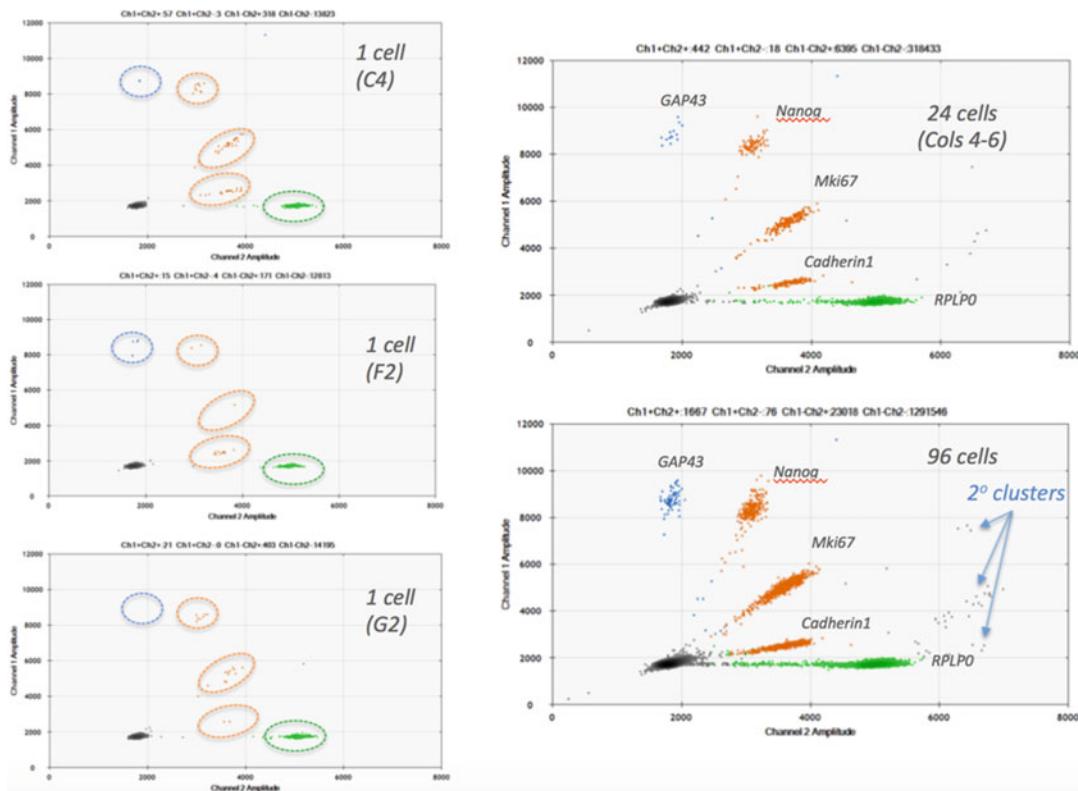
Gene 1 (defined in Fig. 2b) and designate it as “FAM only” (shown as a blue cluster in the 2D plots of Fig. 3b). Next, encircle the “all negative” cluster closest to the origin as your “double-negative” cluster (shown as grey droplets). Finally, use the freehand tool to encircle all other droplets and droplet clusters as a single HEX cluster (shown as green droplets). *See Note 9.*

- (c) QuantaSoft™ will provide you with the concentration of the Gene 1 transcript (per  $\mu\text{L}$  of the ddPCR™ reaction) and the number of transcript copies per well ( $N_{\text{rxn}}$ ) for all selected wells (up to 96 wells per plate). To calculate the copies per cell ( $N_{\text{cell}}$ ) and for further analyses, export the data for all wells as a first “.csv” file. After opening the .csv file (e.g., in Microsoft EXCEL™), locate the column for “CopiesPer20uLWell” (i.e.,  $N_{\text{rxn}}$ ) and divide each value by the fraction of cDNA analyzed in the 20  $\mu\text{L}$  ddPCR™ reaction. Thus,  $N_{\text{cell}} = N_{\text{rxn}}/f$ , where  $f$  is the fraction of a cell’s cDNA analyzed per ddPCR™ well (typically,  $f \approx 0.4$  [4  $\mu\text{L}$  cDNA measured per well/ $\sim 10 \mu\text{L}$  total cDNA per cell]).

Note that typically, the ddPCR™ reaction is formulated to contain 4.4  $\mu\text{L}$  cDNA in a 22  $\mu\text{L}$  sample volume, of which only 20  $\mu\text{L}$  is converted into droplets.

- (d) Repeat **steps (b)** and **(c)** iteratively for each gene of interest in all wells run with the same 5-plex assay (i.e., a total of five times to analyze all five gene transcripts per ddPCR™ reaction). *See Note 10.*
- (e) After quantifying a bulk control RNA by both singleplex and 5-plex assays, confirm that the 5-plex quantification of your target transcripts is in agreement with your singleplex quantifications run on the same sample(s). If there are significant discrepancies, you will need to identify which assay(s) is/are being compromised in the multiplex reaction. You may then either modify the design for that target gene region or chose another target sequence within the transcript. Alternatively, if you are testing two 5-plexes, you may be able to switch an individual assay between the two multiplexes to mitigate the interference.
- (f) Once you have validated the performance of your multiplex(es), you are ready to use them for single-cell analyses according to the workflow described in this chapter.

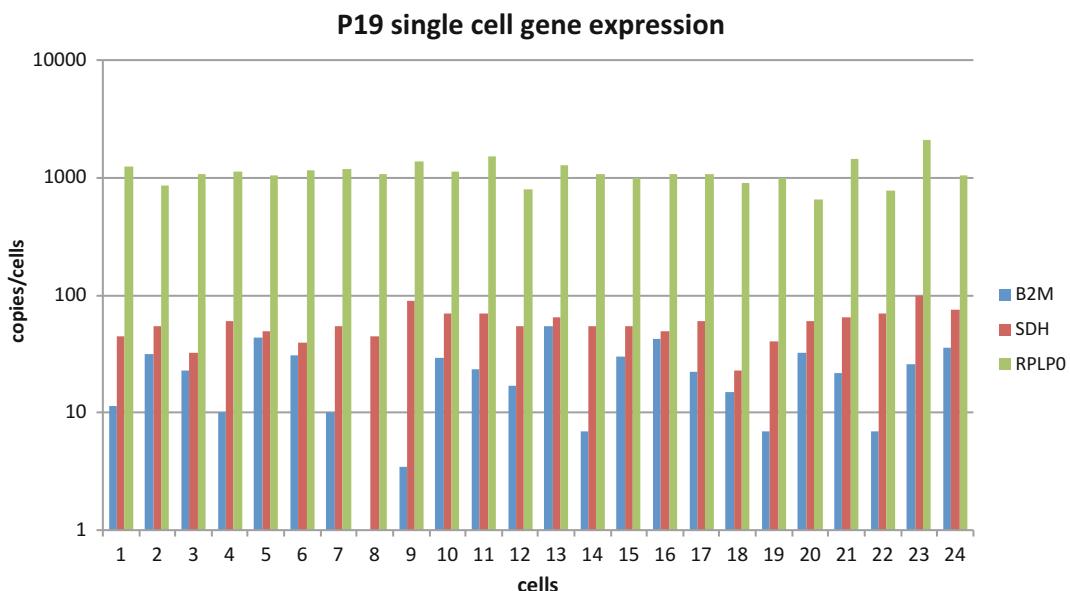
An example of data from a control RNA sample derived from a mixture of both untreated and retinoic acid-treated P19 mouse cells is illustrated in Fig. 2d. Additionally, Fig. 4 compares three 2D droplet plots from individual single cells vs. 2D droplet plots of 24 and 96 single-cell reactions superimposed. It is clear from this example that most transcript levels per cell are low enough such that very few if any transcripts partition into droplets containing more than one gene transcript type (*see Note 11*). Figure 5 shows the result of cell-to-cell variation in transcript levels for 3 genes quantified by this workflow typifying low (B2M), intermediate (SDH) and high (RPLP0) abundance transcripts analyzed in 24 untreated P19 mouse cells (*see Note 12*).



**Fig. 4** Transcript levels for individual single-cells assayed with the 5-plex 1 ddPCR™ assay targeting GAP43, Nanog, Mki67, Cadherin1, and RPLPO. Examples of 2D droplet plots for three P19 single cells are shown assayed with 5-plex 1. Note the differences in transcript level between cells for the various GOIs by comparing the number of droplets in the same encircled primary cluster in each of the typical cells shown. For comparison, see the aggregate level of each transcript from 24 and 96 single cell results superimposed in the adjacent 2D plots. Note some clusters are identified here by orange (rather than green) droplets merely for better visualization of the five genes

#### 4 Notes

1. Avoid low-quality pipette tips that may shed plastic particulates into the DNA and Droplet Digital™ PCR solutions which can then shred droplets. This results in poor quantification.
2. If primers and probes are ordered separately, it is possible to formulate the 5-plex assay so that 1  $\mu\text{L}$  contains 20 $\times$  concentration for all five assays (20 $\times$  = 18  $\mu\text{M}$  each primer and 5  $\mu\text{M}$  each total probe concentration). This would potentially allow the addition of the total amount of cDNA per cell to a single ddPCR™ reaction (~8  $\mu\text{L}$  cDNA/20  $\mu\text{L}$  ddPCR™ reaction) and thus double the sensitivity of the assay, at the expense of only assaying for five genes per cell. However, it would need to be established that the ddPCR™ reaction tolerated this larger



**Fig. 5** Example of varying transcript abundances for three genes across 24 cells using the method described here (but with two duplex assays run per cell). Transcript levels are shown for genes with low (beta-2 microglobulin, B2M), intermediate (succinic dehydrogenase, SDH), and high (Ribosomal Protein Lateral Stalk Subunit P0, RPLPO) abundance in untreated P19 mouse cells. Transcript levels for B2M vary by nearly tenfold across these cells, whereas those for SDH and RPLPO vary by only a fewfold in these same cells and highlight the frequent noncorrelation of transcript levels for various genes across apparently like cells

volume of unpurified cDNA and Lysis Buffer and quantified linearly with the increased sample amount.

3. Be sure the assays do not show signs of overquantification at the selected annealing temperature, due to nonspecific target amplification that may be seen at lower than optimal temperatures; nor underquantification at higher temperatures where the amplitude of positive droplets may still be distinguishable from negative droplets but may give too low a concentration value due to “molecular dropout” (i.e., where a droplet or nano-chamber partition contains a target of interest, but for various reasons, may not achieve sufficient fluorescence amplitude at endpoint to be recognized as a positive partition). Typically, the concentration value measured will remain constant over a range of annealing temperatures and this is where the optimal temperature should be chosen. Although the quantification provided by digital PCR is in absolute copies per transcript per volume of sample analyzed, it must be remembered that there is generally some uncertainty about the efficiency of reverse transcribing any given RNA target region into cDNA. This can be influenced by quality of the RNA, sequence of the target region and its context, the reverse transcriptase used (e.g., RNase H<sup>-</sup> or RNase H<sup>+</sup>) and the conditions of

cDNA synthesis. To determine the true quantity of a transcript present in a sample requires a systematic effort to study these factors for the transcript of interest and the use of multiple assays designed against different regions of the same transcript. Most often, it is sufficient to establish that the assay design chosen gives unambiguous results, and is reproducible from run-to-run and from sample to sample.

4. With regard to deciding which three assays are to be detected with a combination of FAM and HEX probes, if some assays have higher or lower fluorescence amplitude in singleplex than others (potentially due to differences in amplicon lengths, where longer amplicons tend to give lower amplitude droplets with probes chemistry), it may make cluster identification and thus quantification easier if lower amplitude assays are alternated with higher amplitude ones to better spread out the clusters from one another in 2D droplet space plots. Alternatively, total probe concentration for a given target may be increased or decreased to increase or decrease its droplet cluster fluorescence amplitude, respectively. Note that assays yielding short amplicons are generally preferred since they tend to give clusters with higher fluorescence amplitude where neighboring primary clusters will be better separated from one another than if they are all at lower amplitudes.
5. Preparation of single cell suspensions from solid tissue requires mechanical dissociation and/or enzymatic digestion for optimal recovery of cells from the tissue. Conditions and enzyme requirements for digestion of the tissue of interest will need to be determined empirically.
6. It is necessary to filter the cell suspension to avoid clogging the narrow bores of the sample injection needle and tubing on the flow cytometer, which can be easily clogged by aggregated cells and debris. The concentration also influences the rate of sorting, which typically progresses at 2000–20000 cells/second; thus, resuspension of cells at a density of 1–10 million cells/mL in PBS/0.1% BSA is important.
7. If desired, it is also possible to sort more than 1 cell per well to interrogate pools of cells of known number. We have done pools of 10–10,000 cells per well to assess linearity and efficiency of the lysis and overall protocol. Note, however, that when sorting more than ~100 cells per well, the total volume of crude lysate increases due to the sheath volume of sorting buffer (~3.3 nL/sorted cell) and needs to be taken into consideration for calculations of transcript levels per cell.
8. This lysis method efficiently extracts mRNA from single cells as demonstrated by comparing average yields of transcripts per cell (for multiple genes) determined on a population of sorted

single P19 cells, as compared to a bulk preparation of total RNA extracted from 10,000 P19 cells sorted into a single well using a guanidinium isothiocyanate lysis method (Aurum kit from Bio-Rad; manuscript in preparation).

9. It is very important to be aware that in QS 1.7.4.0917 and earlier software versions, all droplets are assigned as negative droplets by default. If they are left as unassigned droplets, they will be included in the total count of negative droplets (i.e., the “All Negatives” cluster) resulting in under-quantification of your gene of interest (GOI) when computed by the Poisson equation. Thus, droplets for all other positive clusters besides that for the gene of interest—including primary clusters for the other gene targets and secondary and tertiary clusters, if present—must be defined as another *single*-positive “meta-cluster” (e.g., in HEX only, shown as green clusters in the Fig. 3b). These green droplets will be ignored in calculation of the GOI’s copy number per well by QuantaSoft™. If they are defined as double-positive (i.e., orange) clusters, they will incorrectly inflate the number of counted copies of the “FAM only” GOI and give it an excessively high transcript copy number per well.
10. If you were to analyze more than one 5-plex reaction per 96-well plate—for example, 48 cells analyzed per plate with two 5-plexes per cell, each ddPCR™ reaction with ~half of a cell’s cDNA—you would define clusters as described in Sub-heading 3.6, step 2 (b) for the first 48 wells together (5-plex 1), then similarly for the second 48 wells together (5-plex 2), then export the transcript concentrations and counts/well for each first gene of interest (in both 5-plexes), followed iteratively by each of the other 4 genes in each 5-plex.
11. This method of radial multiplexing should also be readily applicable to the analysis of a handful of tumor mutations in circulating cell-free (cf)DNA, since such mutations are likely to exist at low concentrations even if more than one of them is present in a single plasma cfDNA sample. Hence, there too, as in single cell analysis, it is unlikely that a significant number of mutant target molecules would colocalize into the same droplet (i.e., in secondary and tertiary clusters) and therefore most targets would be captured in analyzing primary clusters alone.
12. This method is capable of measuring as few as 3–5 transcripts per cell, or approximately five- to tenfold more sensitive than RNA-Seq methods.

## Acknowledgments

We wish to thank our many colleagues who have assisted in the development of the instrumentation and reagents necessary to develop this protocol, especially Shenglong Wang who began earlier single-cell work from which this was developed. Special thanks also to Svilen Tzonev, Doug Hauge, Niels Klitgord, and Dimitri Skvortsov for assistance with the quantification analysis, and to Marcos Oquendo for assistance with the S3e cell sorter.

## References

- Thomas PS (1980) Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc Natl Acad Sci U S A* 77(9):5201–5205
- Karlin-Neumann GA, Sun L, Tobin EM (1988) Expression of light-harvesting chlorophyll a/b-protein genes is phytochrome-regulated in etiolated *Arabidopsis thaliana* seedlings. *Plant Physiol* 88:1323–1331
- Berk AJ, Sharp PA (1977) Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* 12(3):721–732
- Zinn K, DiMaio D, Maniatis T (1983) Identification of two distinct regulatory regions adjacent to the human beta-interferon gene. *Cell* 34(3):865–879
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods* 25:402–408. <https://doi.org/10.1006/meth.2001.1262>
- Livak KJ, Wills QF, Tipping AJ et al (2012) Methods for qPCR gene expression profiling applied to 1440 lymphoblastoid single cells. *Methods* 59(1):71–79. <https://doi.org/10.1016/j.ymeth.2012.10.004>
- Schena M, Shalon D, Davis RW et al (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270(5235):467–470
- DeRisi JL, Iyer VR, Brown PO (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278(5338):680–686
- Iyer VR, Eisen MB, Ross DT et al (1999) The transcriptional program in the response of human fibroblasts to serum. *Science* 283:83–87. <https://doi.org/10.1126/science.283.5398.83>
- Lin Z, Fillmore GC, Um T-H et al (2003) Comparative microarray analysis of gene expression during activation of human peripheral blood T cells and leukemic Jurkat T cells. *Lab Investig* 83(6):765–776
- Canales RD, Luo Y, Willey JC et al (2006) Evaluation of DNA microarray results with quantitative gene expression platforms. *Nat Biotechnol* 24(9):1115–1122. <https://doi.org/10.1038/nbt1236>
- Losick R (2015) A love affair with *Bacillus subtilis*. *J Biol Chem* 290(5):2529–2538. <https://doi.org/10.1074/jbc.X114.634808>
- Sorlie T, Tibshirani R, Parker J et al (2003) Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* 100(14):8418–8423. <https://doi.org/10.1073/pnas.0932692100>
- Tobin EM, Silverthorne J (1985) Light regulation of gene expression in higher plants. *Annu Rev Plant Physiol* 36:569–593. <https://doi.org/10.1146/annurev.pp.36.060185.003033>
- Sanders R, Mason DJ, Foy CA et al (2013) Evaluation of digital PCR for absolute RNA quantification. *PLoS One* 8(9):e75296. <https://doi.org/10.1371/journal.pone.0075296>
- Hindson CM, Chevillet JR, Briggs HA et al (2013) Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat Methods* 10(10):1003–1005. <https://doi.org/10.1038/nmeth.2633>
- Lemos DR, Babaeijandaghi F, Low M et al (2015) Nilotinib reduces muscle fibrosis in chronic muscle injury by promoting TNF-mediated apoptosis of fibro/adipogenic progenitors. *Nat Medicine* 21(7):786–794. <https://doi.org/10.1038/nm.3869>
- Nagalakshmi U, Wang Z, Waern K et al (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* 320(5881):1344–1349. <https://doi.org/10.1126/science.1158441>
- Wilhelm BT, Marguerat S, Watt S et al (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution.

- Nature 453(7199):1239–1243. <https://doi.org/10.1038/nature07002>
20. Mortazavi A, Williams BA, McCue K et al (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5 (7):621–628. <https://doi.org/10.1038/nmeth.1226>
  21. Ståhlberg A, Bengtsson M (2010) Single-cell gene expression profiling using reverse transcription quantitative real-time PCR. Methods 50:282–288. <https://doi.org/10.1016/jymeth.2010.01.002>
  22. Tang F, Barbacioru C, Wang Y et al (2009) mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods 6:377–382. <https://doi.org/10.1038/nmeth.1315>
  23. Shalek A, Satija R, Shuga J et al (2014) Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. Nature 510 (7505):363–369. <https://doi.org/10.1038/nature13437>
  24. Macosko EZ, Basu A, Satija R et al (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell 161:1202–1214. <https://doi.org/10.1016/j.cell.2015.05.002>
  25. Klein AM, Mazutis L, Akartuna I et al (2015) Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 161:1187–1201. <https://doi.org/10.1016/j.cell.2015.04.044>
  26. Whale A, Huggett J, Tzonev S (2016) Fundamentals of multiplexing with digital PCR. Biomol Detect Quantif 10:15–23. <https://doi.org/10.1016/j.bdq.2016.05.002>
  27. Monzo HJ, Park TIH, Montgomery JM et al (2012) A method for generating high-yield enriched neuronal cultures from P19 embryonal carcinoma cells. J Neurosci Methods 204:87–103. <https://doi.org/10.1016/j.jneumeth.2011.11.008>
  28. Zhong Q, Bhattacharya S, Kotsopoulos S et al (2011) Multiplex digital PCR: breaking the one target per color barrier of quantitative PCR. Lab Chip 11(13):2167–2174. <https://doi.org/10.1039/c1lc20126c>
  29. Pender A, Garcia-Murillas I, Rana S et al (2015) Efficient genotyping of KRAS mutant non-small cell lung cancer using a multiplexed droplet digital PCR approach. PLoS One 10 (9):e0139074. <https://doi.org/10.1371/journal.pone.0139074>
  30. Hughesman CB, XJD L, Liu KYP et al (2016) A robust protocol for using multiplexed droplet digital PCR to quantify somatic copy number alterations in clinical tissue specimens. PLoS One 11(8):e0161274. <https://doi.org/10.1371/journal.pone.0161274>
  31. Kinz E, Leiherer A, Lang AH et al (2015) Accurate quantitation of JAK2 V617F allele burden by array-based digital PCR. Int J Lab Hematol 37(2):217–224. <https://doi.org/10.1111/ijlh.12269>
  32. Madic J, Zocevic A, Senlis V et al (2016) Three-color crystal digital PCR. Biomol Detect Quantif 10:34–46. <https://doi.org/10.1016/j.bdq.2016.10.002>



# Chapter 25

## Quantification of Circulating MicroRNAs by Droplet Digital PCR

Manuela Ferracin and Massimo Negrini

### Abstract

MicroRNAs (miRNAs) are released in the blood as cell-free molecules either linked to Ago proteins and LDL or enveloped inside exosomes and microvesicles. The amount of specific circulating microRNAs has been discovered to change accordingly to a disease state and to be potentially used as a disease biomarker. Sensitive and accurate methods for circulating microRNA quantification using probe-based or dye-based digital PCR technology have been developed. With a digital PCR system it is possible to obtain the absolute quantification of specific miRNAs, bypassing several issues related to low abundance targets and miRNA normalization. This chapter addresses the workflow and methods for miRNA assessment in biological fluids using EvaGreen-based droplet digital PCR as well as how to analyze and interpret results.

**Key words** MicroRNA, Serum, Plasma, Droplet digital PCR, Diagnostics, Cancer, Biomarkers

---

### 1 Introduction

The discovery of stable and assessable miRNA molecules in human body fluids, including blood, saliva, urine, cerebrospinal fluid, and synovial fluid, opened novel possibilities in their use as disease biomarkers [1–3]. These miRNAs are detected outside the cells and for this reason they are better known as cell-free or circulating miRNAs; their amount can be extremely low but they present an unexpected resistance to degradation. Despite the promising and exciting features of circulating miRNAs, the low amount and lack of known endogenous reference genes in body fluids provide a real challenge for every reliable translational application [4, 5]. In order to obtain trustworthy miRNA quantification—even for the less abundant miRNA species—several protocols for the global circulating miRNA profiling and miRNA-specific validation have been developed.

The methods applied to perform body fluid miRNA profiling include probe-based or LNA-based quantitative PCR array, microarray, Nanostring nCounter technology, smallRNA sequencing [6].

For each of these technologies, specific protocol variants have been developed in order to manage these low-abundance RNA samples, although satisfactory results are not always guaranteed [7].

In addition, the absence of true endogenous control miRNAs outside cells generated novel issues to deal with during the validation step. Some laboratories decided to use the more stable circulating miRNAs detected during the initial high-throughput screening for sample normalization. Others adopted an approach aimed at quantifying the amount of each miRNA normalized to plasma, serum, or any other body fluid volume, using a sort of absolute quantification [7, 8]. Digital PCR technology proved to be very effective in the second approach, even exceeding the performance of traditional quantitative PCR [8]. Indeed, digital PCR target amount calculation was comparable to the use of raw CT (cycle threshold) in qPCR experiments, with the advantage of providing the exact count of miRNA molecules in the solution and a better accuracy in low-abundance miRNA quantification. The EvaGreen-based ddPCR efficiency has been previously determined [9]. The assay is precise, reproducible over a range of concentrations of four orders of magnitude, and sensitive, detecting a target miRNA at levels down to 1 copy/ $\mu$ L.

Thus, this chapter focuses on how to perform miRNA quantification using EvaGreen-based droplet digital PCR technology with QX200 system (Bio-Rad), specifically considering serum and plasma clinical samples [9].

---

## 2 Materials

List of reagents, instrumentation, and support materials to run an experiment.

### 2.1 Sample Preparation

1. EDTA Tubes (Vacutette or BD Vacutainer).
2. Tabletop centrifuge.
3. 1.5 mL polypropylene tubes.

### 2.2 RNA Extraction

1. miRNeasy Mini Kit (Qiagen).
2. 100 nmole RNA oligo Cel-miR-39-3p (Integrated DNA Technologies). Sequence: UCACCGGGUGUAAAUCAGCUUG.
3. Tabletop centrifuge.

### 2.3 cDNA Synthesis

1. Universal cDNA synthesis kit II (Exiqon).
2. 200  $\mu$ L PCR tubes or plates.
3. Thermal cycler.

## 2.4 EvaGreen-Based Droplet Digital PCR

1. MicroRNA LNA PCR primer set (Exiqon).
2. QX200 droplet generator (Bio-Rad).
3. QX200 droplet reader (Bio-Rad).
4. QuantaSoft software (Bio-Rad).
5. PX1 PCR plate sealer (Bio-Rad).
6. DG8 droplet generator cartridges and gaskets (BioRad).
7. QX200 ddPCR EvaGreen supermix (BioRad).
8. QX200 droplet generator oil for EvaGreen dye (BioRad).
9. 200  $\mu$ L PCR plates (Eppendorf).
10. Pierceable foil plate seals.
11. Thermal cycler.

## 3 Methods

### 3.1 Sample Preparation

Plasma and serum processing is a relevant step in circulating miRNA quantification. There is no preferred procedure for plasma and serum preparation. To be comparable, all the samples from the same experiment must be processed using exactly the same workflow.

1. *Plasma preparation.* Collect 5 mL blood in EDTA (ethylene-diaminetetraacetic acid) tubes (Vacutte or BD Vacutainer); centrifuge samples at  $1000 \times g$  for 10 min at room temperature to remove blood cells, and dispense the supernatant plasma in aliquots. Aliquots can be stored at  $-80^{\circ}\text{C}$  until use.
2. *Serum preparation.* Collect 5 mL blood in serum tubes (Vacutte or BD Vacutainer), keep at room temperature to clot for at least 60 min, and then spin at  $1000 \times g$  for 10 min; remove he serum and dispense in aliquots. Aliquots can be stored at  $-80^{\circ}\text{C}$  until use.

### 3.2 Protocol for Total RNA (Including miRNA) Isolation

We recommend starting from 200  $\mu$ L serum or plasma. Total RNA can be isolated from serum or plasma using commercially available kits. We suggest using miRNeasy Mini Kit (Qiagen) as described by the supplier (*see Note 1*).

To monitor the extraction and reverse transcription reaction occurrence, add 3  $\mu$ L of a 4.16 nM solution of the synthetic miRNA cel-miR-39-3p from *C. elegans* (e.g., custom synthesized by Integrated DNA Technologies, sequence: UCACCGGGU-GUAAAUCAGCUUG) to 1 mL QIAzol Lysis Reagent (Qiagen) before adding the lysis reagent to the sample (*see Note 4*).

This is the protocol for total RNA extraction using miRNeasy Mini Kit (Qiagen).

1. Thaw serum/plasma samples on ice.
2. Add 1 mL of QIAzol Lysis Reagent (Qiagen) to 200  $\mu$ L serum/plasma.
3. Mix by vortexing and place the tube containing the homogenate on the bench at room temperature for 5 min.
4. Add 3  $\mu$ L of a 4.16 nM solution of the synthetic miRNA cel-miR-39-3p from *C. elegans*.
5. Add 200  $\mu$ L of chloroform. Shake the tube vigorously for 15 s. Place the tube on the bench at room temperature for 2 min.
6. Centrifuge for 15 min at 12,000  $\times g$  at 4 °C to obtain phases separation: the upper aqueous phase contains RNA.
7. Transfer the aqueous phase to a new tube, about 700  $\mu$ L, avoid transfer of any white interphase material.
8. Add 1 mL of 100% ethanol and mix by inversion.
9. Pipet up 700  $\mu$ L of the sample into a Mini spin column placed on 2 mL collection tube. Close the lid and centrifuge at 12,000  $\times g$  for 15 s. Discard the flow-through.
10. Repeat this step using the remaining sample.
11. Add 700  $\mu$ L Buffer RWT to the Mini spin column, close the lid and centrifuge for 15 s at 12,000  $\times g$  to wash the column. Discard the flow-through.
12. Pipet 500  $\mu$ L Buffer RPE into the Mini spin column. Close the lid and centrifuge for 15 s at 12,000  $\times g$ . Discard the flow-through. Repeat this step again.
13. Centrifuge the Mini spin column at full speed for 2 min to dry the spin column membrane from residual ethanol.
14. Transfer the Mini spin column to a new 1.5 mL collection tube and pipet 35  $\mu$ L RNase-free water on the membrane of the column.
15. Centrifuge for 1 min at 12,000  $\times g$  to elute the RNA.

Note: Since RNA concentration cannot be determined accurately, we suggest using fixed volumes as a measure for input amount.

### **3.3 MicroRNA Reverse Transcription**

The conversion of RNA, including miRNA, to cDNA can be performed using the Universal cDNA synthesis kit II (Exiqon) following the company's guidelines for miRNA profiling in serum and plasma (*see Note 2*). The resulting cDNA has to be diluted at least 1:50 prior to amplification.

**Table 1**  
**Reverse transcription reagent mix components**

Reagent	Volume ( $\mu$ L)
5x reaction buffer	4
Nuclease-free water	11
Enzyme mix	2
Template total RNA	3
Total volume	20

Protocol for miRNA reverse transcription using the Universal cDNA synthesis kit II (Exiqon).

1. Thaw the reverse transcription (RT) reagent mix components: 5× reaction buffer, nuclease-free water. Mix gently inverting the tubes and place on ice. Immediately before use, remove the enzyme mix from the freezer and place on ice. Spin down all reagents.
2. Prepare the RT reagent mix on ice as described in Table 1. Prepare at least 10% exceeding mix.
3. Mix the reaction by pipetting, and then spin down.
4. Incubate for 60 min at 42 °C.
5. Heat-inactivate the reverse transcriptase for 5 min at 95 °C.
6. Immediately cool to 4 °C.
7. Store the cDNA at –20 °C.

### 3.4 cDNA Dilution

1. Dilute the amount of cDNA template needed for the planned ddPCR reactions in nuclease free water. We recommend diluting the cDNA reaction between 1:50 and 1:500 in water, dependent on target miRNA abundance. Store the diluted cDNA at –20 °C. The diluted cDNA proved to be stable for at least 3 months at –20 °C.

### 3.5 Droplet Generation and PCR

Locked Nucleic Acid (LNA)-based miRNA-specific primers (Exiqon) are used with a green fluorescent DNA-binding dye (Eva-Green) in the QX200 droplet digital PCR system (Bio-Rad). Droplet generation should be performed on 8 samples at a time. Technical replicates are not required, due to the high reproducibility of this technology [8, 9]. A No Template Control (NTC) sample should be always run in every plate for each different ddPCR condition.

**Table 2**  
**ddPCR reagent mix components**

Reagent	Volume ( $\mu\text{L}$ )
2x EvaGreen Supermix	10
LNA primer set	Variable (0.5–1) <sup>a</sup>
Nuclease-free water	Variable
Diluted cDNA template	8
<b>Total volume</b>	<b>20</b>

<sup>a</sup>See Note 5

For miRNA quantification, a 20  $\mu\text{L}$  PCR mixture is prepared containing 10  $\mu\text{L}$  2X EvaGreen supermix (Bio-Rad), 8  $\mu\text{L}$  diluted cDNA, and 0.25–1  $\mu\text{L}$  of miRNA-specific miRCURY LNA PCR primer set (Exiqon).

#### Protocol for miRNA quantification.

1. Thaw and equilibrate the EvaGreen Master mix (Bio-Rad), microRNA primer set (Exiqon) and cDNA at room temperature.
2. Mix thoroughly EvaGreen Master mix by inverting the tube several times. Spin down all the reagents.
3. Prepare the ddPCR mix as described in Table 2. When multiple ddPCR reactions are performed with the same microRNA primer set, it is recommended to prepare a ddPCR mix working-solution. Prepare at least 10% exceeding mix.
4. Once assembled, thoroughly mix and spin down the ddPCR mix.
5. Dispense 12  $\mu\text{L}$  of ddPCR mix into a 96-well PCR plate or PCR tubes.
6. Add 8  $\mu\text{L}$  of diluted cDNA template to each tube/well and mix by pipetting.
7. Insert the droplet generator cartridge (DG8, Bio-Rad) into the holder.
8. Transfer 20  $\mu\text{L}$  of each prepared sample to the sample wells (middle row) of the DG8 cartridge, carefully while avoiding air bubbles at the bottom of the well.
9. Fill each oil well (bottom row) with 70  $\mu\text{L}$  of EvaGreen droplet generator oil (Bio-Rad).
10. Hook the gasket over the cartridge holder and insert into the QX200 droplet generator (Bio-Rad).

**Table 3**  
**Thermal cycling conditions for droplet digital PCR**

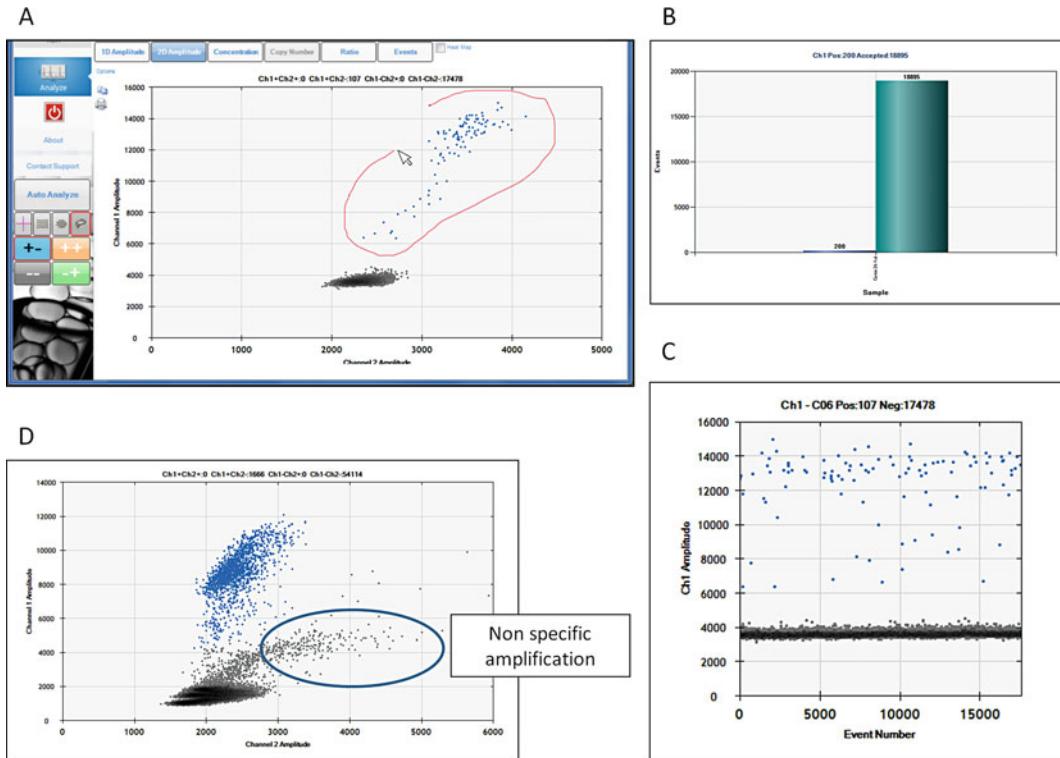
Cycling step	Temperature (°C)	Time	Ramping rate	Cycles
Enzyme activation	95	5 min	~2 °C/s	1
Denaturation	95	30 s		40
Annealing/extension	56–60 <sup>a</sup>	1 min		
Signal stabilization	4 90	5 min 5 min		1 1
Hold (optional)	4	Infinite		1

<sup>a</sup>Optimal temperature should be determined for each primer set. See Note 5

11. Close the lid to start droplet generation. When droplet generation has finished, open the lid, remove the disposable gasket. The top wells of the cartridge contain the droplets.
12. Pipet slowly and smoothly 40 µL of the contents of the eight top wells (the droplets) into a single column of a 96-well PCR plate.
13. Seal the PCR plate with foil immediately after transferring droplets to avoid evaporation. Use pierceable foil plate seals that are compatible with the needles in the QX200 droplet reader.
14. Begin thermal cycling (PCR) within 30 min of sealing the plate.
15. Perform the cycling protocol according to Table 3.

### 3.6 Droplet Reading and Data Analysis

1. Power on the QX200 droplet reader (Bio-Rad).
2. Move the plate from the thermal cycler to the droplet reader.
3. Place the 96-well PCR plate containing the post-PCR droplets into the base of the plate holder.
4. Place the top of the plate holder on the PCR plate. Firmly press both release tabs down to secure the PCR plate in the holder.
5. Start the QuantaSoft software from the system PC.
6. In QuantaSoft software, click Setup to define your experiment then click Run to start the droplet reading.
7. When droplet reading is complete, click “Analyze” button to open and analyze the data.
8. Use the 2D amplitude plot in QuantaLife software (Bio-Rad) to select the positive droplets (lasso tool) (Fig. 1a).
9. Use the Events tab to check the number of positive and total droplets. A total number of 18,000–21,000 droplets is usually achieved with EvaGreen ddPCR (Fig. 1b, c).



**Fig. 1** Positive droplets selection. (a) Positive droplet selection from the 2D Plot using the *lasso* function and manually drawing a circle around the proper cloud. (b) Total number of positive (blue) and negative (green) droplets visualized in the QuantaLife software for each sample. (c) 1D Plot representation of positive and negative droplets. (d) Some LNA primer set (miR-125a-5p in the panel) displays a nonspecific product formation, which is evident also in NTC well. A miRNA-specific concentration can still be obtained using the 2D plot and excluding the off-target amplification product (see Note 3)

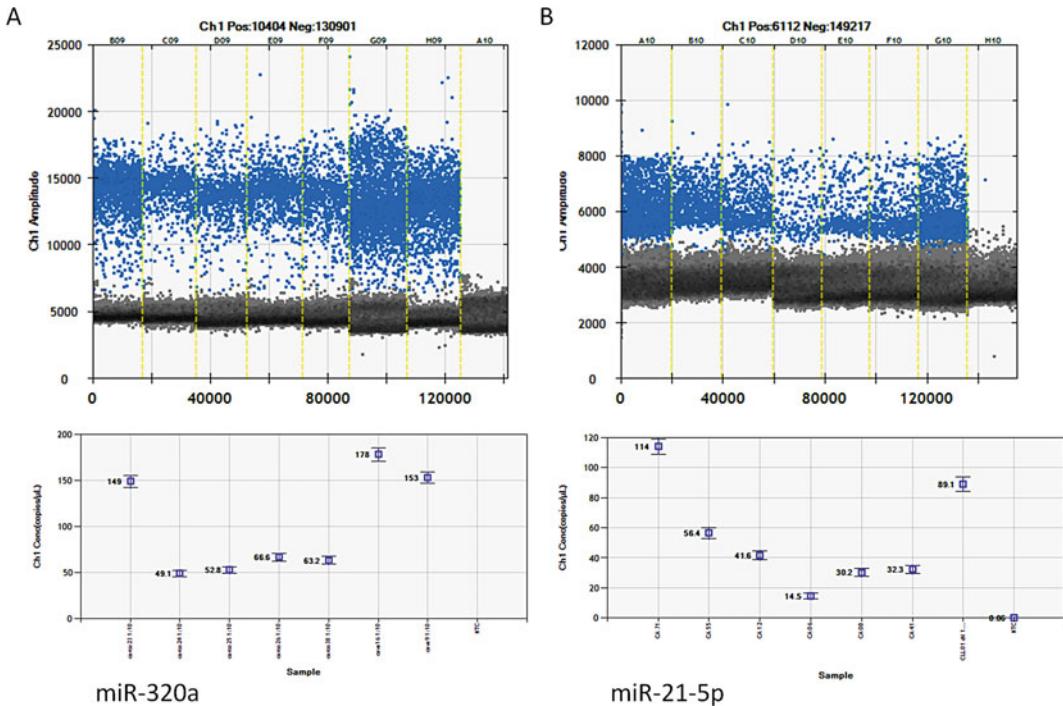
10. Once the positive droplets are selected, export the miRNA concentration values using the export .csv option from the Concentration tab.
11. Assuming that each miRNA molecule is reverse-transcribed in a cDNA molecule, we can calculate the absolute copies in 1  $\mu$ L of plasma or serum multiplying the obtained concentration in ddPCR reaction value for a dilution factor (DF = 145.83 for 1:50 cDNA dilution).

#### 4 Notes

1. Plasma can be prepared using different protocols. The most common protocol suggests to use two centrifugations at  $1200 \times g$  either at room temperature or at +4 °C. The

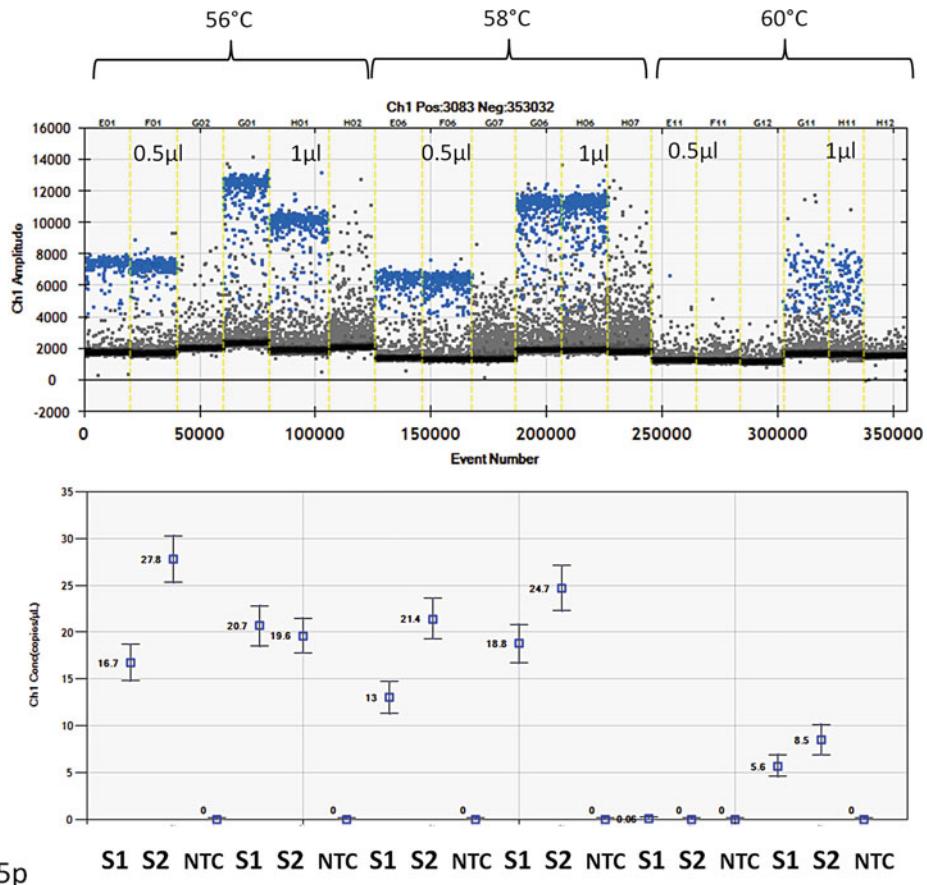
increasing of the spin rate reduces the number of exosomes, microvesicles, and platelets in the sample and therefore changes the miRNA composition [10]. It is therefore mandatory to compare samples that have been prepared following the same protocol.

2. When a certain number of circulating miRNAs needs to be assessed, a universal cDNA system that reverse-transcribes all the miRNAs in one reaction (e.g., the systems proposed by Exiqon, Qiagen, and now also by Applied Biosystems/Thermo Fisher) is preferable to the still largely used system developed by Applied Biosystems/Thermo Fisher consisting of a miRNA-specific reverse transcription with stem-loop primers. Indeed, a universal reverse transcription system provides the flexibility to subsequently select the subset of miRNAs to test starting from the same RT, therefore reducing the amount of RNA required for multiple miRNAs assessment.
3. The NTC (No Template Control) sample needs to be run in every plate to verify the expected fluorescent signal of the individual droplets. There is always the possibility that due, to the short length of the miRNAs, nonspecific products are amplified. Since the size of targeted miRNA amplicons and that of primer dimers are relatively close in size, it may be difficult to discriminate them using fluorescence amplitude only. Analysis using 2D plots adds a second fluorescent dimension to the droplets and sometimes allows separation of two distinct populations. If the “cloud” of nonspecific target amplification is not overlapping with the true signal (Fig. 1d), the miRNA can still be quantified selecting only the true positive droplets. In this case, the NTC sample is essential for droplet selection. In the case of overlapping true- and false-positive droplet distributions, the assay cannot be used with the ddPCR system. A different assay, from a different supplier, should be considered.
4. Working in every step with fixed input volumes it is possible to obtain the absolute quantification of miRNA copies, without the need of additional normalization steps. Therefore, results obtained using droplet digital PCR can be easily used to calculate the number of each miRNA molecule that are present in 1 mL of plasma or serum. Spiked-in exogenous miRNAs (e.g., cel-miR-39) can be used to monitor the occurrence of RNA extraction or RT reaction but not to normalize the miRNA levels, because their recovery is more variable than that of endogenous miRNAs [9].

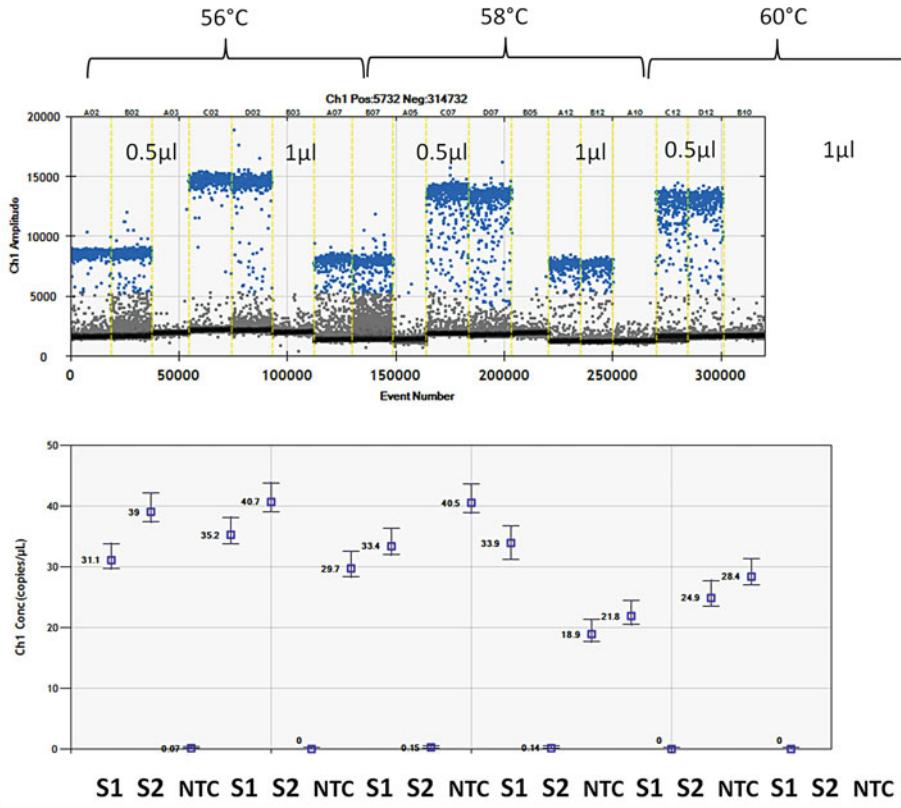


**Fig. 2** Droplet distribution for two different miRNA assays. miR-320a (**a**) and miR-21-5p positive and negative droplet amplitudes (upper panels) and concentration (lower panels) in seven representative plasma samples and NTCs. The fluorescence amplitude can change according to the assay. Results are presented as copies per microliter in the amplification reaction. Error bars represent the Poisson 95% confidence interval

5. The amount of each miRNA in the blood is very different and some miRNAs result to be more abundant than others. Using a 1:50 diluted cDNA in a ddPCR reaction is usually sufficient to obtain an adequate number of positive and negative droplets. In case of positive droplet saturation (i.e., no negative droplets) a further dilution of the cDNA sample is necessary.
6. Different EvaGreen-based miRNA assays can generate different positive and negative droplet amplitudes (Fig. 2). Every LNA primer set should be optimized before running all the samples. The result of miR-125a-5p and miR-425-5p optimization is represented in Figs. 3 and 4 respectively. Specifically, by changing the amount of primer (usually in the range of 0.25–1 μL per ddPCR reaction) and the annealing temperature (usually in the range of 56–60 °C), it is possible to identify the combination that determines a better separation between positive and negative droplets.



**Fig. 3** Influence of primer concentration and annealing temperature on the fluorescence amplitude of droplets in EvaGreen-based ddPCR. 1D Plots and concentration measurements from the miR-125a-5p obtained at annealing temperatures of 56, 58, and 60 °C and using 0.5 or 1 μL primer for two different plasma samples (S1, S2) and NTC. This assay does not work at 60 °C in EvaGreen master mix. The best positive-negative droplet separation can be obtained performing the PCR at 56° and using either 0.5 or 1 μL primer. The nonspecific product is still visible in 1D Plots (see Note 6)



**Fig. 4** Influence of primer concentration and annealing temperature on the fluorescence amplitude of droplets in EvaGreen-based ddPCR. 1D Plots and concentration measurements from the miR-425-5p obtained at annealing temperatures of 56, 58, and 60 °C and using 0.5 or 1 μL primer for two different plasma samples (S1, S2) and NTC. This assay work properly at all temperatures in EvaGreen master mix. The best positive-negative droplet separation can be obtained performing the PCR at 56° and using 1 μL primer

## Acknowledgments

The work was supported by funding from the Italian Association for Cancer Research (AIRC) to MF (MFAG 11676) and to MN (Special Program Molecular Clinical Oncology - 5 per mille n. 9980, 2010/15) and from the Italian Ministry of Instruction, University and Research FIRB 2011 (Project RBAPIIBYNP) and University of Ferrara (FAR 2012-14) to MN.

## References

- Chen X, Ba Y, Ma L, Cai X, Yin Y, Wang K et al (2008) Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res* 18 (10):997–1006. PubMed PMID: 18766170. Epub 2008/09/04. eng.

2. Lawrie CH, Gal S, Dunlop HM, Pushkaran B, Liggins AP, Pulford K et al (2008) Detection of elevated levels of tumour-associated microRNAs in serum of patients with diffuse large B-cell lymphoma. *Br J Haematol* 141(5):672–675. PubMed PMID: 18318758. Epub 2008/03/06. eng.
3. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL et al (2008) Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A* 105(30):10513–10518. PubMed PMID: 18663219. Pubmed Central PMCID: 2492472. Epub 2008/07/30. eng.
4. Jarry J, Schadendorf D, Greenwood C, Spatz A, van Kempen LC (2014) The validity of circulating microRNAs in oncology: five years of challenges and contradictions. *Mol Oncol* 8(4):819–829. PubMed PMID: 24656978. Epub 2014/03/25. eng.
5. Witwer KW (2015) Circulating MicroRNA biomarker studies: pitfalls and potential solutions. *Clin Chem* 61(1):56–63. PubMed PMID: 25391989. Epub 2014/11/14. Eng.
6. Moldovan L, Batte KE, Trgovcich J, Wisler J, Marsh CB, Piper M (2014) Methodological challenges in utilizing miRNAs as circulating biomarkers. *J Cell Mol Med* 18(3):371–390. PubMed PMID: 24533657. Pubmed Central PMCID: 3943687. Epub 2014/02/19. eng.
7. Ferracin M, Lupini L, Salamon I, Saccenti E, Zanzi MV, Rocchi A et al (2015) Absolute quantification of cell-free microRNAs in cancer patients. *Oncotarget* 6(16):14545–14555. PubMed PMID: 26036630. Epub 2015/06/04. Eng.
8. Hindson CM, Chevillet JR, Briggs HA, Galli- chotte EN, Ruf IK, Hindson BJ et al (2013) Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat Methods* 10(10):1003–1005. PubMed PMID: 23995387. Epub 2013/09/03. eng.
9. Miotto E, Saccenti E, Lupini L, Callegari E, Negrini M, Ferracin M (2014) Quantification of circulating miRNAs by droplet digital PCR: comparison of EvaGreen- and TaqMan-based chemistries. *Cancer Epidemiol Biomark Prev* 23(12):2638–2642. PubMed PMID: 25472671. Epub 2014/12/05. eng.
10. Cheng HH, Yi HS, Kim Y, Kroh EM, Chien JW, Eaton KD et al (2013) Plasma processing conditions substantially influence circulating microRNA biomarker levels. *PLoS One* 8(6):e64795. PubMed PMID: 23762257. Pubmed Central PMCID: 3676411. Epub 2013/06/14. eng.



# Chapter 26

## Droplet Digital PCR for Absolute Quantification of Extracellular MicroRNAs in Plasma and Serum: Quantification of the Cancer Biomarker hsa-miR-141

Maria D. Giraldez, John R. Chevillet, and Muneesh Tewari

### Abstract

Droplet-based digital PCR provides high-precision, absolute quantification of nucleic acid target sequences with wide-ranging applications for both research and clinical diagnostic applications. Droplet-based digital PCR enables absolute quantification by counting nucleic acid molecules encapsulated in discrete, volumetrically defined water-in-oil droplet partitions. The current available systems overcome the previous lack of scalable and practical technologies for digital PCR implementation. Extracellular microRNAs in biofluids (plasma, serum, urine, cerebrospinal fluid, etc.) are promising noninvasive biomarkers in multiple diseases and different clinical settings (e.g., diagnosis, early diagnosis, prediction of recurrence, and prognosis). Here we describe a protocol that enables highly precise and reproducible absolute quantification of extracellular microRNAs using droplet digital PCR.

**Key words** Digital PCR, MicroRNA, Biofluids, Plasma, Serum, Absolute quantification, qPCR, Reproducibility

---

### 1 Introduction

Digital PCR is a method that provides high-precision, absolute quantification of nucleic acid target sequences based on multiple partitioning of individual analyte molecules at limiting dilution, in its original form described as resulting in one or zero molecules in most reactions [1]. After endpoint PCR, the starting concentration of template is determined by Poisson statistical analysis of the number of positive (containing amplified target) and negative (no amplified target detected reactions). The digital PCR concept [2] has many potential advantages over real-time PCR, including the capability to obtain absolute quantification without external references, robustness to variations in PCR efficiency [3] and potentially higher resilience to PCR inhibitors [4, 5]. Importantly, technology has become commercially available that overcomes the previous lack of scalable and practical technologies for digital PCR.

implementation [6]. In this sense, the current commercial systems allow reactions to be partitioned into nanoliter-size (Bio-Rad) [7, 8] or picoliter-size (RainDance) [9], aqueous droplets in oil rather than multiwell plates. Rapid microfluidic analysis of thousands or millions of droplets per sample (Bio-Rad and RainDance, respectively) makes droplet-based digital PCR feasible for routine use. In addition, the practical dynamic range of the system is substantially improved with highly uniform droplets (e.g., Bio-Rad QX100/200), which (with Poisson correction for multiple target molecules per droplet) enable the precise calculation of concentrations even above conditions of limiting dilution [7, 8] (e.g., up to and sometimes even beyond an average of five copies per droplet).

We have developed a protocol for highly precise and reproducible absolute quantification of extracellular microRNAs using the Bio-Rad Droplet Digital™ PCR (ddPCR™) system and compared its performance against RT-qPCR [10]. Extracellular microRNAs in biofluids (plasma, serum, urine, cerebrospinal fluid, etc.) are being actively studied as potential noninvasive biomarkers in multiple diseases and different clinical settings (e.g., diagnosis, early diagnosis, prediction of recurrence, and prognosis) [11]. For example, abundance of circulating, cell-free hsa-miR-141 in plasma or serum has been found to be a biomarker for several epithelial cancer types, including prostate, colorectal, and breast cancers [12–19]. It is being studied as a biomarker for prognosis, as well as for minimally invasive (i.e., via blood sample) diagnosis and potentially even early detection of cancer. In our hands, the use of ddPCR for microRNA quantification provided higher precision and less day-to-day variation than RT-qPCR [10]. This greater precision was associated with a somewhat greater sensitivity of ddPCR than qRT-PCR for some of the microRNA assays evaluated, often when sensitivity was assessed as limit of quantification, but also in some cases when assessed as limit of detection [10]. These features along with the possibility of obtaining absolute quantification without any standard reference (i.e., without requiring a standard curve) make ddPCR a very attractive technique for microRNA research. In any case, the better reproducibility provided by ddPCR along with the possibility of direct comparison of results intra-lab (not only day-to-day results from the same project, but also potentially current results to archival data from older projects) and potentially across different labs, is expected to enable increased reproducibility in research, multi-institutional data comparisons and collaboration, and ultimately faster research progress.

## 2 Materials

### 2.1 Instruments

1. QX100™ or QX200™ Droplet Digital™ PCR (ddPCR™) systems (each consists of two instruments, the Droplet Generator and the Droplet Reader) (Bio-Rad).
2. C1000 Touch™ thermal cycler with a gradient-enabled 96-deep well module (Bio-Rad).
3. PCR Plate Heat Sealer (Eppendorf) or PX1™ PCR Plate Sealer (Bio-Rad).
4. Single channel manual pipettes, 2, 20, 200, and 1000 µL (Pipet Lite LTS, Rainin).
5. Eight-channel manual pipettes 10, 20, 50, and 200 µL (Pipet Lite LTS, Rainin).

### 2.2 Reagents for Creating Dilution Series (Optional)

1. miRNA synthetic template/s (IDT). See Note 1.
2. Nuclease-free water.
3. MS2 carrier RNA 0.8 µg/µL (Roche, PN: 10165948001).
4. Eppendorf twin.tec PCR plate 96 (Eppendorf, PN: 951020362).
5. 15 mL conical tubes (Corning, PN: 352196).
6. 1.7 mL low adhesion microcentrifuge tubes (GeneMate, PN: C-3302-1).

### 2.3 Reagents for Reverse Transcription

1. TaqMan® MicroRNA Reverse Transcription Kit (Applied Biosystems, PN: 4366596 or 4366597) Kit Contents: 10× RT Buffer, dNTP mix, RNase Inhibitor, Multiscribe™ reverse transcriptase (RT) enzyme.
2. TaqMan® MicroRNA Assays 5X Reverse Transcription primer for your specific target (Applied Biosystems, PN: 4427975).
3. Eppendorf twin.tec PCR plate 96 (Eppendorf, PN: 951020362).
4. Adhesive PCR foils (Thermo Scientific, PN: AB-0626).
5. 1.7 mL low adhesion microcentrifuge tubes (GeneMate, PN C-3302-1).

### 2.4 Reagents for Preparation of PCR- Ready Samples, Droplet Generation and PCR Amplification of Droplets

1. TaqMan® MicroRNA 20X PCR assay (Applied Biosystems, PN: 4427975).
2. 2× ddPCR™ Supermix for Probes No dUTP (Bio-Rad, PN: 186-3024).
3. Droplet Generation Oil for Probes (Bio-Rad, PN: 186-3005).
4. DG8™ Cartridges for QX100/QX200 Droplet Generator (Bio-Rad, 24 cartridges PN: 186-4008).

5. DG8™ Gaskets for QX100/QX200 Droplet Generator (Bio-Rad, 24 gaskets PN: 186-3009).
6. Eppendorf twin.tec semiskirted PCR plate 96 (Eppendorf, PN: 951020362).
7. Pierceable foil heat seal (Bio-Rad, PN:1814040).
8. 1.7 mL low adhesion microcentrifuge tubes (GeneMate, PN: C-3302-1).
9. Sterile disposable reagent reservoir (Costar, PN: 07-200-127).

## **2.5 Reagents for Droplet Reading**

1. ddPCR Droplet Reader Oil (Bio-Rad, PN:186-3004).

---

## **3 Methods**

### **3.1 Blood Specimen Processing for Plasma and RNA Isolation**

Methods for sample processing and RNA isolation have been described by us elsewhere [12, 20, 21]. Briefly, in the latest iteration of the plasma collection protocol, we collect whole blood samples in K2EDTA tubes and separate the plasma within 2 h using a two-stage centrifugation protocol to ensure that it is free not only of cells but also of significant numbers of contaminating platelets [21]. We discard plasma samples that show visual evidence of hemolysis. Total RNA containing microRNAs is purified from 200 µL of plasma or serum using the miRNeasy kit (Qiagen) with slight modification as previously described [20] although we have also successfully employed an alternative protocol using the miR-Vana PARIS RNA isolation kit (Ambion) [12]. We find that it is not uncommon for the concentration of purified RNA to be too low for quantification using absorption spectrophotometry. We use a fixed volume of RNA eluate rather than a fixed mass of RNA as input into the subsequent reverse transcription reaction.

### **3.2 Create Dilution Series (Optional Note 2)**

All the following steps should be performed on ice unless otherwise specified. *See Note 3.*

1. Thaw synthetic miRNA template/s (10 µM) and MS2 carrier RNA.
2. Label four low adhesion 1.7 mL microcentrifuge tubes I–IV for each synthetic template. *See Note 4.*
3. Prepare MS2 diluent in a 15 mL conical tube as shown in Table 1.
4. Mix gently by inversion.
5. Pipet the volumes of MS2 diluent shown in column 3 of Table 2 into the microcentrifuge tubes prelabeled as I–IV for each microRNA dilution series.

**Table 1**  
**MS2 diluent**

	$1 \times$
Nuclease-free water	9.585 mL
MS2 carrier RNA (0.8 $\mu$ g/ $\mu$ L)	415 $\mu$ L
Total volume	10 mL

**Table 2**  
**microRNA synthetic template dilutions**

Dilution # (Tube Label)	Transferred miRNA ( $\mu$ L)	Diluent H <sub>2</sub> O/ MS2 ( $\mu$ L)	miRNA copies/ $\mu$ L	Copies/ $\mu$ L in RT assay when 2 $\mu$ L of dilution is added to a 10 $\mu$ L RT assay	Copies/ $\mu$ L in ddPCR when 1 $\mu$ L of synthesized cDNA is added to 20 $\mu$ L ddPCR assay
Tube or well	Supplied			$6.02 \times 10^{12}$	
I	10	990	$6.02 \times 10^{10}$		
II	10	990	$6.02 \times 10^8$		
III	10	990	$6.02 \times 10^6$		
IV	10	990	$6.02 \times 10^4$		
1 (96-well plate starts)	41	59	24,691	4936	250
2	50	50	12,345	2468	125
3	50	50	6173	1234	62.5
4	50	50	3086	617	31.3
5	50	50	1543	309	15.6
6	50	50	772	154	7.8
7	50	50	386	77	3.9
8	50	50	193	39	1.95
9	50	50	96	19	0.98
10	50	50	48	10	0.49
11	50	50	24	5	0.25
12	0	50	0	0	0

6. Pipet 59  $\mu$ L of MS2 diluent in the first well of a 96-well plate and 50  $\mu$ L for the 2nd–12th consecutive wells.
7. Spin briefly the synthetic miRNA template/s (10  $\mu$ M), mix well by pipetting up and down several times, and briefly spin down again.
8. Transfer 10  $\mu$ L of the corresponding synthetic miRNA into the microcentrifuge tube prelabeled as I.
9. Mix this solution 5 $\times$  with the full volume by pipetting up and down to ensure complete mixing and spin down briefly.
10. Transfer 10  $\mu$ L of the volume contained in the microcentrifuge tube I into the microcentrifuge tube prelabeled as II.
11. Repeat steps 9–10, changing the volume transferred as indicated in Table 2, until the entire dilution series has been completed in tubes II–IV and then in wells 1 through 12 of the microplate. See Note 5.

### 3.3 Reverse Transcription

All the following steps should be performed on ice unless otherwise specified. See Note 3.

1. Thaw 10X RT buffer, dNTP mix, and 5X RT primer for the miRNA of interest. Keep RNase inhibitor and RT enzyme cold on a portable cooler. See Notes 6 and 7.
2. Vortex RT primer, 10X RT buffer, and dNTP mix to ensure adequate mixing and spin down briefly. Spin down briefly RNase inhibitor and RT enzyme.
3. Add the RT Master Mix components listed in Table 3 to a low adhesion microcentrifuge tube. Mix well by pipetting up and down several times before adding RNase inhibitor and reverse transcription enzyme. Then mix again by pipetting up and down five times.

**Table 3**  
**Reverse transcription master mix**

	1 $\times$
Nuclease-free water	4.12
10x RT buffer	1
100 mM dNTPs	0.1
RT primer (5 $\times$ )	2
RNase inhibitor	0.12
Multiscribe RT	0.66
RNA	2
Total volume	10 $\mu$ L

**Table 4**  
**Reverse transcription thermal conditions**

16 °C	30 min
42 °C	30 min
85 °C	5 min
4 °C	Hold

4. Add 8 µL of RT master mix into each relevant well of a 96-well plate using a multichannel pipette.
5. Pipet 2 µL of the synthetic microRNA template dilutions or extracellular RNA sample into the reverse transcription mix. Mix with full solution volume (10 µL) using a multichannel pipette. Please note that we have tested this protocol for plasma and serum. It may also be useful for analysis of extracellular RNAs from other sources such as urine, sputum, and CSF but would need to be validated for each source biofluid type.
6. Cover the 96-well plate with foil.
7. Spin down briefly the 96-well plate.
8. Place the 96-well plate onto a compatible thermal cycler. Close the lid and perform reverse transcription using the thermal conditions specified in Table 4.

### 3.4 Preparation of PCR-Ready Samples

All the following steps should be performed on ice unless otherwise specified.

1. Thaw 2X Bio-Rad supermix and the 20X primer/probe mix that is specific for your target miRNA. See Notes 6 and 7.
2. Spin down briefly 2X Bio-Rad PCR supermix and mix by pipetting up and down several times (a concentration gradient may form during –20 °C storage). See Note 8.
3. Spin down primers and vortex to ensure an appropriate mixing.
4. Prepare the ddPCR Master Mix by pipetting the reagents shown in Table 5 into a low adhesion microcentrifuge tube and mix thoroughly by pipetting up and down several times. Table 4 provides volumes of components needed for one reaction of Master Mix. In practice, the volumes should be multiplied by the number of reactions needed and increased by another 10% each to have an excess of master mix to compensate for pipetting losses which are inevitable when subsequently drawing multiple aliquots from the Master Mix tube.
5. Add 21 µL of the master mix into each relevant well of a new 96-well plate on ice.
6. Dilute the cDNA solution 1:1 with nuclease-free water.

**Table 5**  
**ddPCR master mix**

	1 ×
Nuclease-free water	8.8
20× primers/probe TM	1.1
2× Bio-Rad MM v1.2	11.1
Total master mix volume for 1 reaction	21 µL

Note that the total volume of master mix and cDNA for each reaction is 23 µL, even though only 20 µL is loaded onto the DG8 chip. The excess is included to cover for any pipetting losses.

### 3.5 Droplet Generation and Amplification of Droplets

7. Pipet 2 µL of diluted cDNA into the master mix.
8. Seal the 96-well plate and spin down briefly (e.g., 1 min at  $600 \times g$ )
  1. Dispense droplet-generating oil into a reservoir. *See Note 9.*
  2. Turn plate thermo sealer on to warm up at least 10 min before use.
  3. Let reaction mixes stand at room temperature for 3 min before transferring into cartridge.
  4. Load a DG8 cartridge in a DG8 cartridge holder placing the notch in the cartridge at the upper left of the holder.
  5. Transfer the reaction mixes to the middle rows of the DG8 cartridge using a multichannel pipette and pipetting slowly to avoid the generation of air bubbles. *See Notes 10 and 11.*
  6. Add 70 µL of the droplet generator oil into each of the bottom wells of the DG8 cartridge using a multichannel pipette.
  7. Attach a gasket on top of the DG8 cartridge by hooking it on both sides of the cartridge holder.
  8. Open the droplet generator by pressing the button on the green top and place the cartridge holder into the instrument.
  9. Close the droplet generator by pressing the same button. The instrument will start generating droplets.
  10. After droplet generation is completed (all three indicator lights on the droplet generator are solid green), open the droplet generator and take out the cartridge holder with the cartridge in place. *See Note 12.*
  11. Remove the disposable gasket from the holder and discard it. Keep the cartridge in the holder.
  12. Transfer 40 µL droplets from the top wells of the cartridge to a 96-well PCR plate (*see Note 13*) by pipetting gently with a

**Table 6**  
**ddPCR cycling conditions**

95 °C	10 min
40 cycles of	
94 °C	30 s
60 °C	60 s
98 °C	10 min
Then	
4 °C	Hold

multichannel pipette. *See Notes 14 and 15.* Cover the plate during the process to prevent evaporation and contamination.

13. Seal the plate immediately after transferring the droplets with a pierceable foil seal to avoid evaporation. *See Notes 16 and 17.*
14. Place the sealed 96-well plate on a thermal cycler and perform the cycling conditions specified on Table 6. *See Notes 18 and 19.*

### 3.6 Droplet Reading

1. When the PCR amplification is completed remove the 96-well plate from the thermal cycler and load it onto a droplet reader. *See Note 20.* Press the button on the green lid of the reader to open the instrument and place the 96-well plate in the plate holder (well A1 of the PCR plate must be in the top left position).
2. Confirm the three indicator lights of the droplet reader are solid green. *See Note 21.*
3. Open QuantaSoft software and click *Setup* in the left navigation bar to define your experiment. Use the well selector to define your settings.

#### Sample:

- Name. Enter the sample ID corresponding to each well.
- Experiment. Select Abs Quant from the drop-down menu.
- Mastermix. ddPCR supermix for probes (no dUTP).

#### Target 1:

- Name. Enter the name of your target microRNA.
- Type. Select Ch1 unknown or NTC (non template control) for each well.

#### Target 2:

- As we are analyzing only one target for experiment, you do not need to fill out any information here.

Click Apply or OK to save your settings (by clicking Apply you save settings without exiting the well editor. In contrast by clicking OK you save changes and close the well editor).

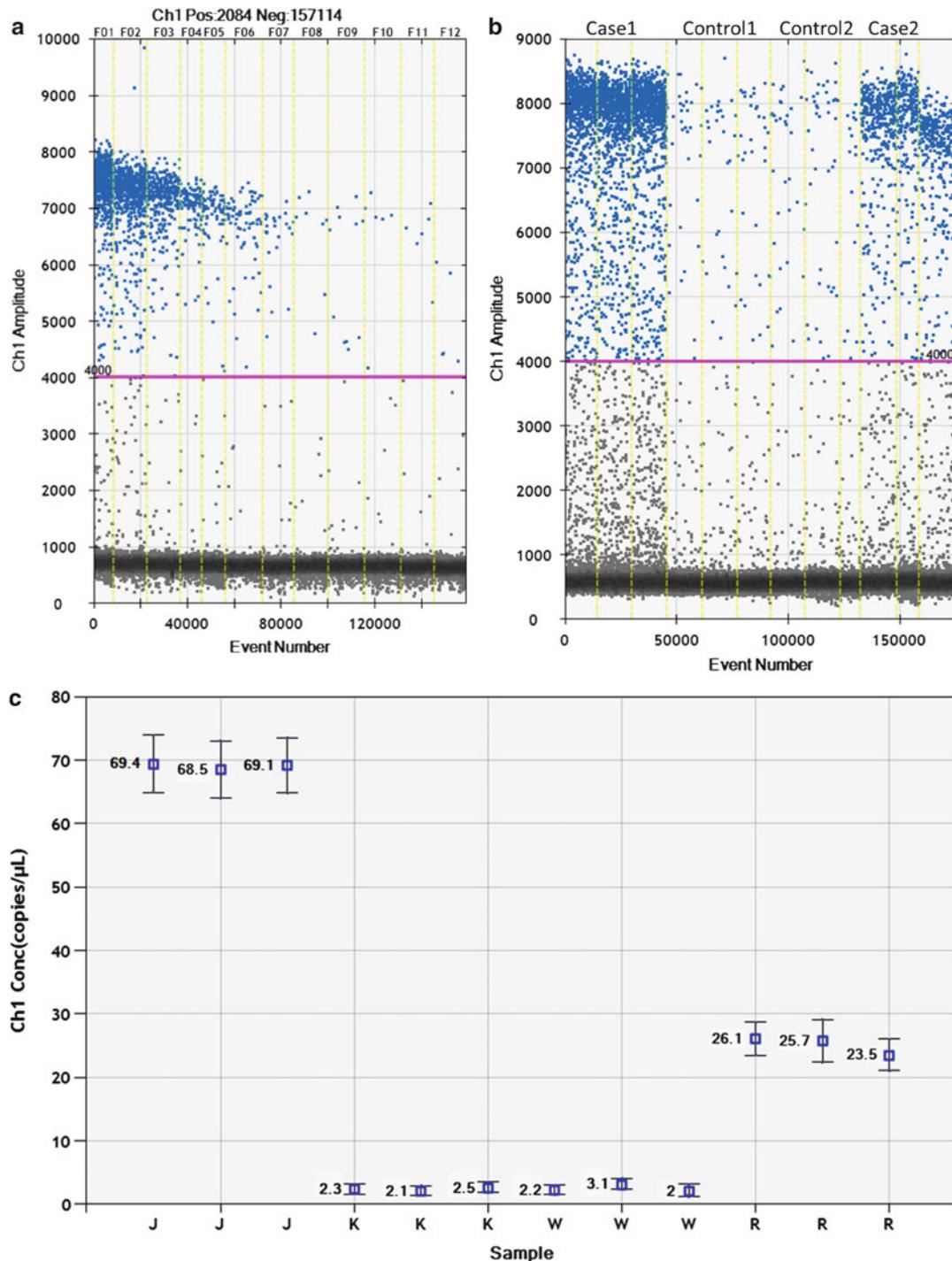
4. Once you have defined your experiment settings click **Run** on the left navigation bar to start the run. In the Run Options window that will appear select the detection chemistry (under set probe dye, click FAM/VIC from the drop-down menu). Click OK. After that, a green circle will appear on the left navigation bar and flash periodically to indicate the run is in progress. *See Note 22.*
5. When droplet reading is complete (all four indicator lights are solid green) open the door by pressing the button on the lid and remove the plate holder from the unit. *See Note 23.* Remove the 96-well PCR plate from the holder and discard it.

### **3.7 Analyze Results**

1. Visualize the data collected from each channel clicking the **1-D amplitude** tab in the QuantaSoft software. In our case only one channel (channel 1, FAM) is used. The software plots each droplet in a sample on a graph of fluorescence intensity vs. droplet number (events). *See Fig. 1a.* You can also visualize the data as a histogram of events vs. amplitude. This tab provides options for adjusting the thresholds used in assigning positives and negatives for the channel.
2. Select the same threshold for all your samples (a threshold of 4000 has typically performed well for the plasma/serum miRNAs we have analyzed, although the appropriate threshold could vary depending on the intensity of positive and negative events) using the multisample threshold tool. *See Note 24.*
3. Click the tab **Events** to review the number of droplet events (you can see positive, negative, or total droplet counts, or any combination of these) counted for each well/sample. Figure *1b*. Consider excluding from the analysis samples with a total number of droplets <10,000 to achieve more narrow 95% confidence intervals.
4. Click the tab **Concentration** to see the estimated copies/ $\mu$ L of your target for each sample as well as the Poisson 95% confidence limits. Figure *1c*.
5. Export your results to an Excel file by clicking on **Export CSV**.

### **3.8 Interpretation of Results—Biomarker Analysis**

The interpretation of data with respect to biomarker analysis for plasma or serum microRNAs is still an active research area, and it depends not only on the specific biomarker but also on the disease context. Thus, a universal, specific “threshold” at which miR-141 level is indicative of disease has not been established. In the research setting where two groups of individuals are compared (e.g., cases which have a disease vs. healthy controls), both parametric and



**Fig. 1** (a) Quantification of serial dilution of synthetic miR-141 using ddPCR. (b) and (c) Quantification of hsa-miR-141 in plasma samples from two healthy controls (K and W) and two patients with advanced prostate cancer (J and R). Three technical replicates per sample were measured. In ddPCR, the QuantaSoft software measures the numbers of droplets that are positive and negative for each fluorophore. Each droplet in a

nonparametric statistical tests are frequently used. In addition, Receiver Operating Characteristic (ROC) curve analysis is an important biomarker analysis approach, which can characterize the relationship between sensitivity and specificity of the biomarker with respect to distinguishing case specimens from controls.

---

#### 4 Notes

1. Prepare individual single-use 10  $\mu\text{M}$  aliquots of the synthetic microRNAs and keep them at  $-80^\circ\text{C}$  until use.
2. Creation of a synthetic miRNA dilution series is not mandatory for ddPCR, as quantification with this technology does not rely on a standard curve. However, the performance of a synthetic miRNA dilution series is advisable to evaluate the efficiency of your assays. Such a dilution series provides a quality control to confirm that the assays and protocol are working well, and it is advisable to include this regularly in runs for such quality control purposes. Using synthetic miRNA dilution series experiments, it is possible to detect assays with low efficiency that would benefit from further optimization before their use in ddPCR experiments. Assay optimization can sometimes be needed to find proper annealing conditions, and in some other cases to discriminate between similar targets. Some factors that could contribute to low apparent assay efficiency include low target accessibility (e.g., primers not being able to anneal due to nonseparation of the target DNA strands), and degraded DNA (i.e., if the DNA is fragmented in the region where primer(s) would bind). It is important to note, however, that for most assays once optimized, normalization for assay efficiency is typically not required (except in extreme cases) even when the efficiency of the assays that you plan to use for an experiment varies somewhat, as the effect of this variation is typically minimal on ddPCR quantification in contrast to qPCR.

---

◀

**Fig. 1** (continued) sample is plotted on a 1D intensity graph (**a**) of fluorescence intensity versus droplet number (events). Vertical yellow lines in the plot show where droplet data from each well start and end, and the assigned threshold appears as a horizontal pink line. All positive droplets (those above the threshold intensity indicated by the horizontal pink line) are scored as positive, and each is assigned a value of one. All negative droplets (those below the threshold) are scored as negative, and each is assigned a value of 0 (zero). The number of positive or negative droplets vs. total number of droplets can be visualized in a bar graph (**b**). This counting technique provides a digital signal from which to calculate the starting target concentration by a statistical analysis of the numbers of positive and negative droplets in a given sample. The software fits the fraction of positive droplets to a Poisson distribution to determine the absolute starting copy number in units of copies/ $\mu\text{L}$  input sample, and then plots the target concentration of each sample as copies per  $\mu\text{L}$  (**c**). Error bars in the concentration plot reflect total error or Poisson 95% confidence limits

3. Working with RNA requires special precautions because of the chemical instability of the RNA and the ubiquitous presence of RNases. Ensure good laboratory practice for handling and storage of RNA to achieve optimal performance of the protocol:
  - Always wear disposable gloves, and work in a nuclease-free environment.
  - Treat surfaces of benches and glassware with commercially available RNase inactivating agents such as RNaseZap (Life Technologies).
  - Use nuclease-free, low nucleic acid binding plasticware and filter barrier pipette tips.
  - Always ensure that all reagents and chemicals purchased commercially are guaranteed to be RNase-free.
  - Keep tubes capped when possible, always spin tubes before opening.
  - Work on ice.
  - For long-time storage, RNA may be stored at –80 °C.
  - Avoid repeated freeze–thaw cycles.
4. Create an individual dilution series for each microRNA that you plan to evaluate in your experiment.
5. Change the micropipette tip between every dilution step to avoid adsorptive carryover.
6. It is advisable to prepare the master mix in a pre-PCR space working in a PCR hood if possible, to minimize chances of contamination during PCR setup.
7. Prepare an individual specific master mix for each microRNA to be tested.
8. After thawing, store the leftover 2X Bio-Rad PCR supermix at 4 °C for 1 to 2 weeks, avoid refreezing.
9. The oil requirement depends on the number of samples that you are processing. If you are not processing a whole 96-well plate, instead of dispensing the total volume contained in the oil generating bottle (7 mL), adjust the volume to the number of samples that you have as listed in Table 7.
10. Air bubbles can cover the bottom of the well and result in 2500–7000 fewer droplets and poor data quality.
11. All eight sample wells in the DG8 droplet generator cartridge must contain sample (create nontemplate control reactions or use Bio-Rad 1× ddPCR buffer control if you are not processing enough samples to occupy the full chip), and all eight oil wells must contain droplet generation oil.

**Table 7**  
**Droplet generator oil requirements**

# Wells	Volume of oil ( $\mu\text{L}$ ) <sup>a</sup>
8	700
24	1820
48	3500
96	6860

<sup>a</sup>Note that the volume of 700  $\mu\text{L}$  includes excess to compensate for losses when pipetting using a multichannel pipettor

12. The time required for droplet generation is typically 2 min for eight samples.
13. As this 96-well plate will be later placed in the droplet reader, it is important to use a plate compatible with this instrument. (Semiskirted plates such as twin.tec semiskirted plates from Eppendorf not only fit properly in the instrument but also provide an optimal rigidity required during the reading process.)
14. Use a 50  $\mu\text{L}$  manual multichannel pipette and pipet gently to avoid damaging the droplets.
15. Do not aspirate  $>40 \mu\text{L}$ , as this causes air to percolate through the droplet solution which can shear and/or cause coalescence of droplets prior to thermocycling.
16. Make sure that you are using pierceable foil heat seal compatible with the sample needle in the droplet reader. (The instrument will pierce the plate seal to aspirate each sample during the droplet reading.) Do not use self-adhesive PCR films.
17. Begin thermal cycling (PCR) within 30 min of sealing the plate, or store the plate at 4 °C for up to 4 h prior to thermal cycling.
18. Use a 2.5 °C/s ramp rate to ensure each droplet reaches the correct temperature for each step during the cycling.
19. 40 cycles of PCR is enough for an optimized ddPCR assay. Do not exceed 50 cycles.
20. Once droplets have undergone PCR-amplification the product is quite stable. You can leave the plate in the thermal cycler overnight at 10 °C or store at 4 °C for up to 3–4 days before proceeding to read the droplets.
21. The first green solid light indicates that the droplet reader is on, the second one that the level of the reader oil and waste bottle are adequate to perform the run, and the third one that there is a plate in place.

22. It might take up to 1 min for the green circle to appear after clicking OK.
23. The QX100 droplet reader can process 32 wells/h.
24. Although changing the threshold typically has a minimal effect on ddPCR results (in contrast to qPCR), we recommend setting the same threshold for each sample in all your experiments to make the results as comparable as possible.

---

## Acknowledgements

M.D.G acknowledges initial support from a Rio Hortega Fellowship and later from a Martin Escudero Fellowship. M.T. acknowledges support from the Department of Defense Peer-Reviewed Cancer Research Program Award CA100606 and NIH US National Institutes of Health Transformative R01 grant R01DK085714.

## References

1. Vogelstein B, Kinzler KW (1999) Digital PCR. Proc Natl Acad Sci U S A 96(16):9236–9241
2. Sykes PJ, Neoh SH, Brisco MJ, Hughes E, Condon J, Morley AA (1992) Quantitation of targets for PCR by use of limiting dilution. BioTechniques 13(3):444–449
3. Bustin SA, Nolan T (2004) Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. J Biomol Tech 15(3):155–166
4. Rački N, Dreо T, Gutierrez-Aguirre I, Blejec A, Ravnikar M (2014) Reverse transcriptase droplet digital PCR shows high resilience to PCR inhibitors from plant, soil and water samples. Plant Methods 10(1):42
5. Dingle TC, Sedlak RH, Cook L, Jerome KR (2013) Tolerance of droplet-digital PCR vs real-time quantitative PCR to inhibitory substances. Clin Chem 59:1670–1672
6. Baker M (2012) Digital PCR hits its stride. Nat Methods 9:541–544
7. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC, Kitano TK, Hodel MR, Petersen JF, Wyatt PW, Steenblock ER, Shah PH, Bousse LJ, Troup CB, Mellen JC, Wittmann DK, Erndt NG, Cauley TH, Koehler RT, So AP, Dube S, Rose KA, Montesclaros L, Wang S, Stumbo DP, Hodges SP, Romine S, Milanovich FP, White HE, Regan JF, Karlin-Neumann GA, Hindson CM, Saxonov S, Colston BW (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Anal Chem 83(22):8604–8610
8. Pinheiro LB, Coleman VA, Hindson CM, Herrmann J, Hindson BJ, Bhat S, Emslie KR (2012) Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. Anal Chem 84 (2):1003–1011
9. Kiss MM, Ortoleva-Donnelly L, Beer NR, Warner J, Bailey CG, Colston BW, Rothberg JM, Link DR, Leamon JH (2008) High-throughput quantitative polymerase chain reaction in picoliter droplets. Anal Chem 80 (23):8975–8981
10. Hindson CM, Chevillet JR, Briggs H, Galli-chotte EN, Ruf IK, Hindson BJ, Vessella RL, Tewari M (2013) Absolute quantification by droplet digital PCR versus analog real-time PCR. Nat Methods 10:1003–1005
11. Weiland M, Gao XH, Zhou L, Mi QS (2012) Small RNAs have a large impact: circulating microRNAs as biomarkers for human diseases. RNA Biol 9(6):850–859
12. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A, Lin DW, Urban N, Drescher CW, Knudsen BS, Stirewalt DL, Gentleman R, Vessella RL, Nelson PS, Martin DB, Tewari M (2008) Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A 105 (30):10513–10518

13. Cheng H, Zhang L, Cogdell DE, Zheng H, Schetter AJ, Nykter M, Harris CC, Chen K, Hamilton SR, Zhang W (2011) Circulating plasma MiR-141 is a novel biomarker for metastatic colon cancer and predicts poor prognosis. *PLoS One* 6(3):e17745
14. Bryant RJ, Pawlowski T, Catto JW, Marsden G, Vessella RL, Rhees B, Kuslich C, Visakorpi T, Hamdy FC (2012) Changes in circulating microRNA levels associated with prostate cancer. *Br J Cancer* 106(4):768–774
15. Madhavan D, Zucknick M, Wallwiener M, Cuk K, Modugno C, Scharpf M, Schott S, Heil J, Turchinovich A, Yang R, Benner A, Riethdorf S, Trumpp A, Sohn C, Pantel K, Schneeweiss A, Burwinkel B (2012) Circulating miRNAs as surrogate markers for circulating tumor cells and prognostic markers in metastatic breast cancer. *Clin Cancer Res* 18 (21):5972–5982
16. Nadal E, Truini A, Nakata A, Lin J, Reddy RM, Chang AC, Ramnath N, Gotoh N, Beer DG, Chen G (2015) A novel serum 4-microRNA signature for lung cancer detection. *Sci Rep* 5 (12464)
17. Kelly BD, Miller N, Sweeney KJ, Durkan GC, Rogers E, Walsh K, Kerin MJ (2015) A circulating MicroRNA signature as a biomarker for prostate cancer in a high risk group. *J Clin Med* 4(7):1369–1379
18. Madhavan D, Peng C, Wallwiener M, Zucknick M, Nees J, Schott S, Rudolph A, Riethdorf S, Trumpp A, Pantel K, Sohn C, Chang-Claude J, Schneeweiss A, Burwinkel B (2016) Circulating miRNAs with prognostic value in metastatic breast cancer and for early detection of metastasis. *Carcinogenesis* 37 (5):461–470
19. Sun Y, Liu Y, Cogdell D, Calin GA, Sun B, Kopetz S, Hamilton SR, Zhang W (2016) Examining plasma microRNA markers for colorectal cancer at different stages. *Oncotarget* 7(10):11434–11449
20. Kroh EM, Parkin RK, Mitchell PS, Tewari M (2010) Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). *Methods* 50(4):298–301
21. Cheng HH, Yi HS, Kim Y, Kroh EM, Chien JW, Eaton KD, Goodman MT, Tait JF, Tewari M, Pritchard CC (2013) Plasma processing conditions substantially influence circulating microRNA biomarker levels. *PLoS One* 8 (6):e64795

# **Part VI**

## **Other Uses of Partitioning**



# Chapter 27

## Droplet Digital™ PCR Next-Generation Sequencing Library QC Assay

Nicholas J. Heredia

### Abstract

Digital PCR is a valuable tool to quantify next-generation sequencing (NGS) libraries precisely and accurately. Accurately quantifying NGS libraries enable accurate loading of the libraries on to the sequencer and thus improve sequencing performance by reducing under and overloading error. Accurate quantification also benefits users by enabling uniform loading of indexed/barcoded libraries which in turn greatly improves sequencing uniformity of the indexed/barcoded samples. The advantages gained by employing the Droplet Digital PCR (ddPCR™) library QC assay includes the precise and accurate quantification in addition to size quality assessment, enabling users to QC their sequencing libraries with confidence.

**Key words** Droplet, Digital PCR, NGS, Sequencing library quantification, Sequencing library qualification, Library balancing, Library loading

---

### 1 Introduction

Next-generation sequencing (NGS) has unlocked vast genetic information and accelerated discoveries in the life sciences and medicine which were unobtainable before. In order to utilize NGS technology, sequencing libraries are constructed typically with adapter sequences which bind to sequencing surfaces and are clonally amplified to boost signal. An important feature of such libraries is to ensure that the adapted library fragments are at a concentration which allows for efficient clonal amplification to occur without mixing with another adapted library fragment. This is achieved in various forms such as creating an emulsion where droplets of the emulsion contain a sequencing surface such as a bead particle and a single adapted library fragment which can bind to oligonucleotides complementary to the adapter sequence and thus be clonally PCR amplified in the droplets from the single starting template. Examples of the emulsion methodology include Ion Torrent™ and Roche 454® sequencing technologies [1, 2]. If the emulsions contain two or more adapted library molecules, then

the resulting bead sequencing surfaces will generate mixed reads which cannot be differentiated. Thus, quantification of the input of adapted library molecules in to the emulsion is a critical component of the sequencing workflow.

Another example of achieving clonal amplification is to create adapted library fragments and flow them on a flow cell which has complementary oligonucleotide adapter sequences arrayed on the surface which can bind the adapted molecules [3]. An important aspect of this method is that the adapted library molecules are at a concentration where they are diffuse enough on the flow cell surface such that when they amplify on the surface they can remain distinct from other adapted library molecules and not merge together during cluster generation. An Example of this method of clonal amplification is Illumina sequencing technology. If clusters merge together, the instrument is not capable of distinguishing and deconvoluting the signal generated during sequencing by synthesis, where fluorescently labeled nucleotides are incorporated into the growing strands on the sequencing clusters. Thus, it is important to load the sequencing flow cell surface at a concentration where most adapted library molecules will be spatially distinct from neighboring adapted library molecules during cluster generation, without cluster merging. Again, this is an example of a critical point in the workflow to have accurate quantification of the adapted library molecules to prevent cluster merging if the concentration were too high, but also not to underload the sequencing surface and lose valuable potential sequencing reads. This protocol will focus on using droplet digital PCR and the Illumina sequencing workflow for NGS library quantification to achieve optimal loading of the sequencing flow cell and to gain equal representation of sequencing libraries by precisely balancing indexed/barcoded libraries when they are combined in sequencing runs.

The Droplet Digital PCR Next-Generation Sequencing Library QC Assay enables the user get highly accurate library quantification, and also provides information on the relative library size and amount of adapter dimers generated in the library construction process. These additional metrics give the user confidence in the quality of their library prior to sequencing which typically required separate qPCR and gel analysis in the past. Another improvement of digital PCR is that the absolute quantification is not dependent on a set of standards with specific sizes and sequences which may or may not be relevant to the PCR efficiencies of the user's actual samples. Digital PCR provides better accuracy, confidence, and reproducibility in the user's measurements [4].

---

## 2 Materials

Set up all ddPCR reactions at room temperature after thawing the various components of the ddPCR™ Library Quantification Kit for Illumina TruSeq, from Bio-Rad part number 1863040. Accurate library quantification is achieved by assaying the NGS libraries at between 100 and 5000 copies/ $\mu$ L.

Kit components include:

ddPCR Supermix for Probes (no dUTP) at 2 $\times$  concentration, ddPCR library quantification assay at 20 $\times$  concentration.

Other materials required:

Nuclease-free water, TE buffer (20 mM Tris-HCl, 0.1 mM EDTA, pH 8), 1.5 mL LoBind eppendorf tubes, Twin-Tec® semi-skirted 96-well PCR plates (Eppendorf 951020362), PX1™ PCR Plate Sealer (Bio-Rad 1814000), Pierceable Foil Heat Seal (Bio-Rad #1814040), QX200™ Droplet Digital™ PCR System (Bio-Rad 1864001), Droplet Generation Oil for Probes (Bio-Rad 1863005), ddPCR™ Droplet Reader Oil (Bio-Rad 1863004), sterile reagent trough, LTS Rainin pipette tips with filters (Rainin GP-L200F, GP-L10F), benchtop microcentrifuge, Microseal® ‘B’ Adhesive Seals (Bio-Rad #MSB1001).

---

## 3 Methods

1. Thaw kit components at room temperature. Vortex the supermix and assay vials thoroughly and spin briefly.
2. Generate a dilution series of Illumina library from the user's unknown stock TruSeq library to be quantified by digital PCR. Dilutions to achieve quantification in the dynamic range of digital PCR will depend on factors such as which library preparation kit or method was used, starting concentration of nucleic acids, and the number of preamplifications cycles used, if any. Suggestions for achieving a working range of dilutions based on the original illumine TruSeq LT library preparation are shown in Table 1. In general, a good starting dilution point would be 10<sup>-7</sup>. Repeat dilution series for additional samples. Note: additional dilution points may be necessary for different library construction methods in order to achieve the dynamic range of ddPCR which currently spans about four orders magnitude. Samples should be diluted in low adsorbing plastic tubes or plates, such as Eppendorf lobind 1.5 mL. Surfactants should not be added to samples to prevent adsorption as they can interfere with droplet making.
3. Pipet 16  $\mu$ L of 2 $\times$  mastermix generated in Table 2 into each well of the 96-well plate.

**Table 1**  
**Library dilution scheme**

Dilution Step	Tube 1	Tube 2	Tube 3	Tube 4	Tube 5
Sample into	5 µL of stock into 495 µL TE	5 µL of Tube 1 into 495 µL TE	5 µL of Tube 2 into 495 µL TE	10 µL of Tube 3 into 90 µL TE	10 µL of Tube 4 into 90 µL TE
TE buffer					
Dilution	100×	100×	100×	10×	10×
Final dilution	10 <sup>-2</sup>	10 <sup>-4</sup>	10 <sup>-6</sup>	10 <sup>-7</sup>	10 <sup>-8</sup>

**Table 2**  
**Reaction setup**

Component	Volume/Reaction	Final Concentration
2× ddPCR supermix for probes (no dUTP)	10 µL	1×
20× ddPCR library quantification assay	1 µL	1×
RNase-/DNase-free water	5 µL	—
Total volume	16 µL	1×

- Pipet 4 µL of the appropriate dilution series generated from Table 1 into the corresponding wells according to a plate layout design such as Table 3. Include no template controls (NTCs) by adding 4 µL of nuclease-free water in place of actual sample.
- Seal the plate with adhesive seal, such as Bio-Rad Microseal “B,” and vortex the plate and briefly centrifuge.
- Remove the adhesive seal and transfer 20 µL of the mastermix into the sample inlet wells on a droplet generation cartridge, add 70 µL of droplet generation oil for probes to the oil inlets wells on the cartridge, place a cartridge gasket on the cartridge and insert the cartridge in to the droplet generator instrument.
- After droplet generation, carefully transfer the newly formed droplets to a new 96-well plate, repeat droplet generation for remaining samples and transfer to the plate, place a foil seal on the plate and heat seal the plate using the PX1 heat sealing instrument.
- Transfer the sealed plate to a thermal cycler and run a thermal cycling protocol according to Table 4.

**Table 3**  
**ddPCR 96 well plate map**

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>
A	NTC	NTC	NTC	NTC	NTC	NTC						
B	Lib 1 (10 <sup>-8</sup> )	Lib 1 (10 <sup>-8</sup> )	Lib 1 (10 <sup>-7</sup> )	Lib 1 (10 <sup>-7</sup> )	Lib 1 (10 <sup>-6</sup> )	Lib 1 (10 <sup>-6</sup> )	Lib 8 (10 <sup>-8</sup> )	Lib 8 (10 <sup>-8</sup> )	Lib 8 (10 <sup>-7</sup> )	Lib 8 (10 <sup>-6</sup> )	Lib 8 (10 <sup>-6</sup> )	Lib 8 (10 <sup>-6</sup> )
C	Lib 2 (10 <sup>-8</sup> )	Lib 2 (10 <sup>-8</sup> )	Lib 2 (10 <sup>-7</sup> )	Lib 2 (10 <sup>-7</sup> )	Lib 2 (10 <sup>-6</sup> )	Lib 2 (10 <sup>-6</sup> )	Lib 9 (10 <sup>-8</sup> )	Lib 9 (10 <sup>-8</sup> )	Lib 9 (10 <sup>-7</sup> )	Lib 9 (10 <sup>-6</sup> )	Lib 9 (10 <sup>-6</sup> )	Lib 9 (10 <sup>-6</sup> )
D	Lib 3 (10 <sup>-8</sup> )	Lib 3 (10 <sup>-8</sup> )	Lib 3 (10 <sup>-7</sup> )	Lib 3 (10 <sup>-7</sup> )	Lib 3 (10 <sup>-6</sup> )	Lib 3 (10 <sup>-6</sup> )	Lib 10 (10 <sup>-8</sup> )	Lib 10 (10 <sup>-8</sup> )	Lib 10 (10 <sup>-7</sup> )	Lib 10 (10 <sup>-6</sup> )	Lib 10 (10 <sup>-6</sup> )	Lib 10 (10 <sup>-6</sup> )
E	Lib 4 (10 <sup>-8</sup> )	Lib 4 (10 <sup>-8</sup> )	Lib 4 (10 <sup>-7</sup> )	Lib 4 (10 <sup>-7</sup> )	Lib 4 (10 <sup>-6</sup> )	Lib 4 (10 <sup>-6</sup> )	Lib 11 (10 <sup>-8</sup> )	Lib 11 (10 <sup>-8</sup> )	Lib 11 (10 <sup>-7</sup> )	Lib 11 (10 <sup>-6</sup> )	Lib 11 (10 <sup>-6</sup> )	Lib 11 (10 <sup>-6</sup> )
F	Lib 5 (10 <sup>-8</sup> )	Lib 5 (10 <sup>-8</sup> )	Lib 5 (10 <sup>-7</sup> )	Lib 5 (10 <sup>-7</sup> )	Lib 5 (10 <sup>-6</sup> )	Lib 5 (10 <sup>-6</sup> )	Lib 12 (10 <sup>-8</sup> )	Lib 12 (10 <sup>-8</sup> )	Lib 12 (10 <sup>-7</sup> )	Lib 12 (10 <sup>-6</sup> )	Lib 12 (10 <sup>-6</sup> )	Lib 12 (10 <sup>-6</sup> )
G	Lib 6 (10 <sup>-8</sup> )	Lib 6 (10 <sup>-8</sup> )	Lib 6 (10 <sup>-7</sup> )	Lib 6 (10 <sup>-7</sup> )	Lib 6 (10 <sup>-6</sup> )	Lib 6 (10 <sup>-6</sup> )	Lib 13 (10 <sup>-8</sup> )	Lib 13 (10 <sup>-8</sup> )	Lib 13 (10 <sup>-7</sup> )	Lib 13 (10 <sup>-6</sup> )	Lib 13 (10 <sup>-6</sup> )	Lib 13 (10 <sup>-6</sup> )
H	Lib 7 (10 <sup>-8</sup> )	Lib 7 (10 <sup>-8</sup> )	Lib 7 (10 <sup>-7</sup> )	Lib 7 (10 <sup>-7</sup> )	Lib 7 (10 <sup>-6</sup> )	Lib 7 (10 <sup>-6</sup> )	Lib 14 (10 <sup>-8</sup> )	Lib 14 (10 <sup>-8</sup> )	Lib 14 (10 <sup>-7</sup> )	Lib 14 (10 <sup>-6</sup> )	Lib 14 (10 <sup>-6</sup> )	Lib 14 (10 <sup>-6</sup> )

*Lib* library, *NTC* no template control

**Table 4**  
**Cycling protocol for Bio-Rad's C1000 Touch™ Thermal Cycler<sup>a</sup>**

Cycling Step	Temperature, °C	Time	Ramp Rate	Number of Cycles
Enzyme activation	95	10 min		1
Denaturation	94	30 s		
Annealing/extension	60	1 min	Approximately 2.0°C/s	40
Enzyme deactivation	98	10 min		1
Hold (optional)	12 or 4	∞		1

<sup>a</sup>Use a heated lid set to 105°C and set the sample volume to 40 µL

9. Transfer the thermal cycled plate to the Droplet Reader instrument, create a plate setup for the run using QuantaSoft software, and read the droplets.
10. Verify the correct automatic thresholding for each well data set which was read and adjust manually if necessary.
11. Calculate the concentration of the NGS library stock by the following method: multiply the ddPCR concentration value of copies/µL by the factor of the dilution used for that value. Then multiply this new value by the factor to account for the dilution in the ddPCR master mix reaction, for example a factor of 5 if 4 µL of diluted library was placed into 16 µL of supermix and assay. The stock concentration of each library can be backcalculated from this method if the ddPCR concentration measurements falls between 100 and 5000 copies/µL. If quantification is outside this range, it will be necessary to assay additional dilution points.

Example:

A value of 253 copies/µL was obtained by ddPCR for a library sample that was assayed at the  $1 \times 10^{-6}$  dilution point. To backcalculate, multiply 253 copies/µL by  $1 \times 10^6$  to account for the dilution point, and then multiply by a factor of 5 for the ddPCR reaction dilution. This equals  $1.265 \times 10^9$  copies/µL. To obtain nM concentration, take  $1.265 \times 10^9$  copies/µL and multiply by  $1 \times 10^6$  µL/L and divide this by  $6.023 \times 10^{23}$  copies per mole =  $2.1 \times 10^{-9}$  M or 2.1 nM

12. Access NGS library quality by quantitating adapter dimers. It is possible to assess adapter dimer concentrations if there are two or more distinct clusters and the user is confident, based on expected sizes of the library that there is an additional unaccounted for cluster which should be assigned to adapter dimers. The TaqMan assay employed in this library QC would yield a cluster, if adapter dimers are present, as the highest fluorescent cluster on the plot. Figure 5a demonstrates

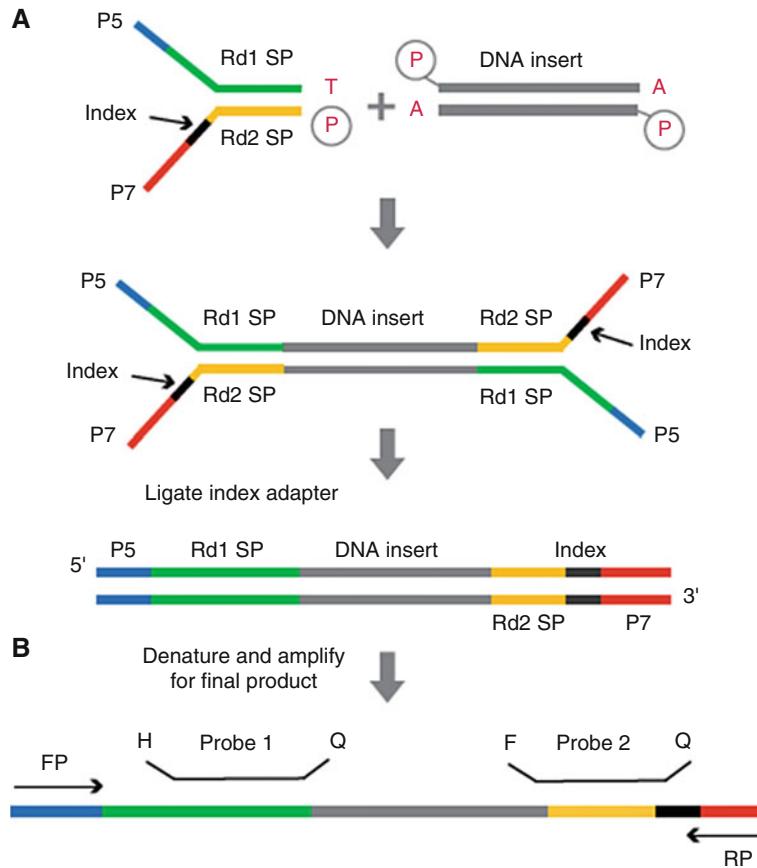
a fluorescent two-dimensional plot of such an adapter dimer cluster. By using the lasso tool in QuantaSoft, the user should assign the negative droplet cluster as double negative (gray), the intermediate cluster or clusters FAM positive only (Blue), and the highest cluster as double positive (orange) such as in Fig. 5a. This enables the user to get a total concentration call of the adapter dimers and the library using the concentration in the FAM channel, and the HEX channel concentration will be the adapter dimer concentration. The user can then take ratios of adapter dimer concentration to the total to get a percentage of the library that is adapter dimer.

---

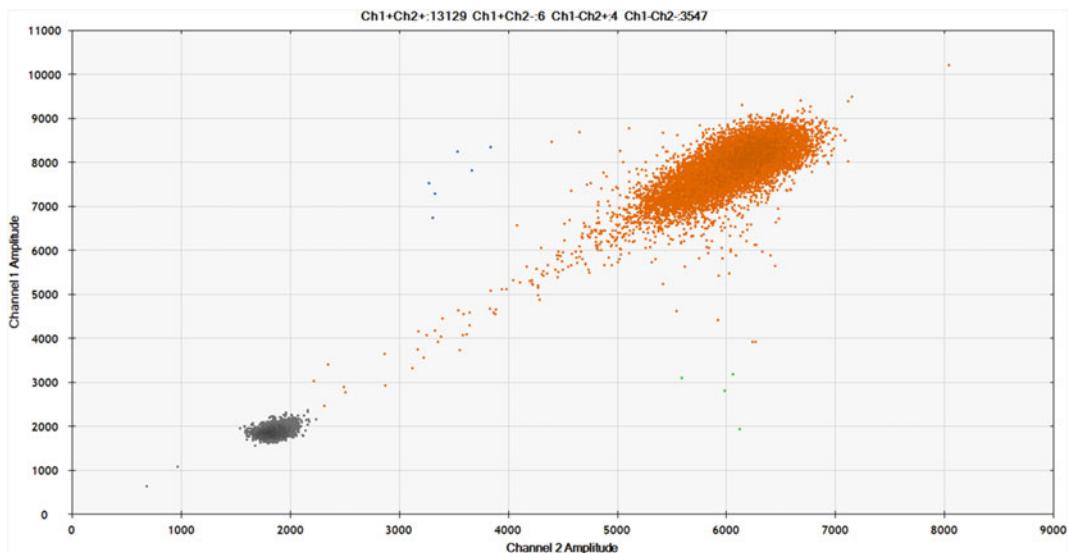
## 4 Data Analysis

The TruSeq library quantification kit contains a 5'-nuclease assays designed to detect and quantify both the P5 and P7 adaptor arms (Fig. 1a, b). The combined signal from each assay is thus used to confirm the formation of bona fide library fragments, which display themselves as double-positive clusters within the 2-channel plot (Fig. 2). Additionally, by exploiting the ability of the QX-100 platform to detect subpopulations of templates with differing amplification efficiencies and with different combinations of P5 and P7 moieties a qualitative measure of library purity can be obtained.

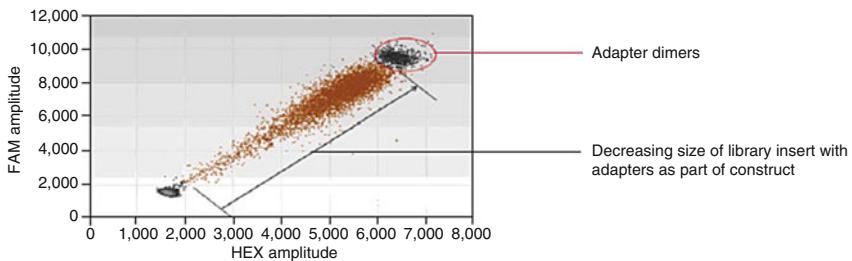
Once data has been acquired on a Droplet Reader instrument, it is important to view the 2-channel plot of the data. This plot enables the user to visualize fluorescent intensity from the two probes in the library quantification kit, the FAM intensity is plotted on the Y-Axis, and the HEX/VIC intensity is plotted on the X-Axis. The 2-channel plot also gives a qualitative sense of the library due to the position of the fluorescent intensity of the clusters. Since the assay is based on two hydrolysis probes, library molecules having both a P5 and P7 adapter will be double positive. If there is no insert in the library molecule and hence it only has adapter-dimers, this product is the most efficient at PCR amplification in the droplet. Thus, droplets containing only adapter-dimers will have the property of being the highest fluorescent amplitude on a 2-channel plot, and they typically form a secondary cluster to the cluster which is comprised of droplets with library molecules having the P5 and the P7 adapters, as well as, library insert. Likewise, the longer the library insert between the two adapters, the less efficient that PCR reaction will be, and thus these droplets will have a lower fluorescent intensity or clustering on the 2-channel plot. Due to these fluorescent intensity differences, the user is able to get a qualitative sense of the library insert size distribution which is inversely related to the fluorescent intensity of the droplets, the larger the insert, the lower the intensity. This effect is shown in



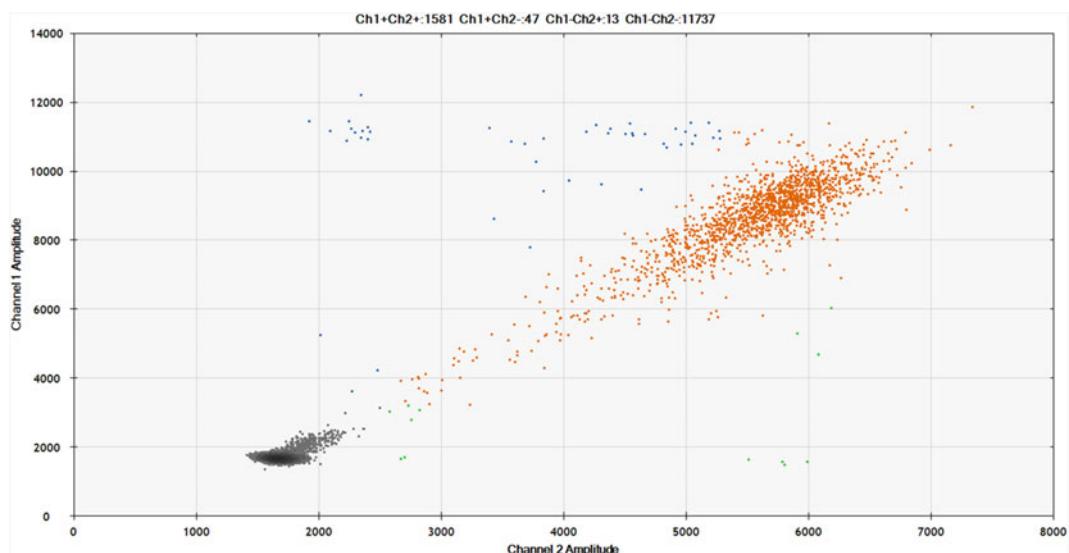
**Fig. 1 (a-b)** Diagrams representing the illumina library architecture of the adapter structures (a), and the hydrolysis probes and primers for assaying library molecules (b)



**Fig. 2** A 2-channel plot showing well-formed library molecules in droplets (orange cluster), and negative droplets (Gray cluster)

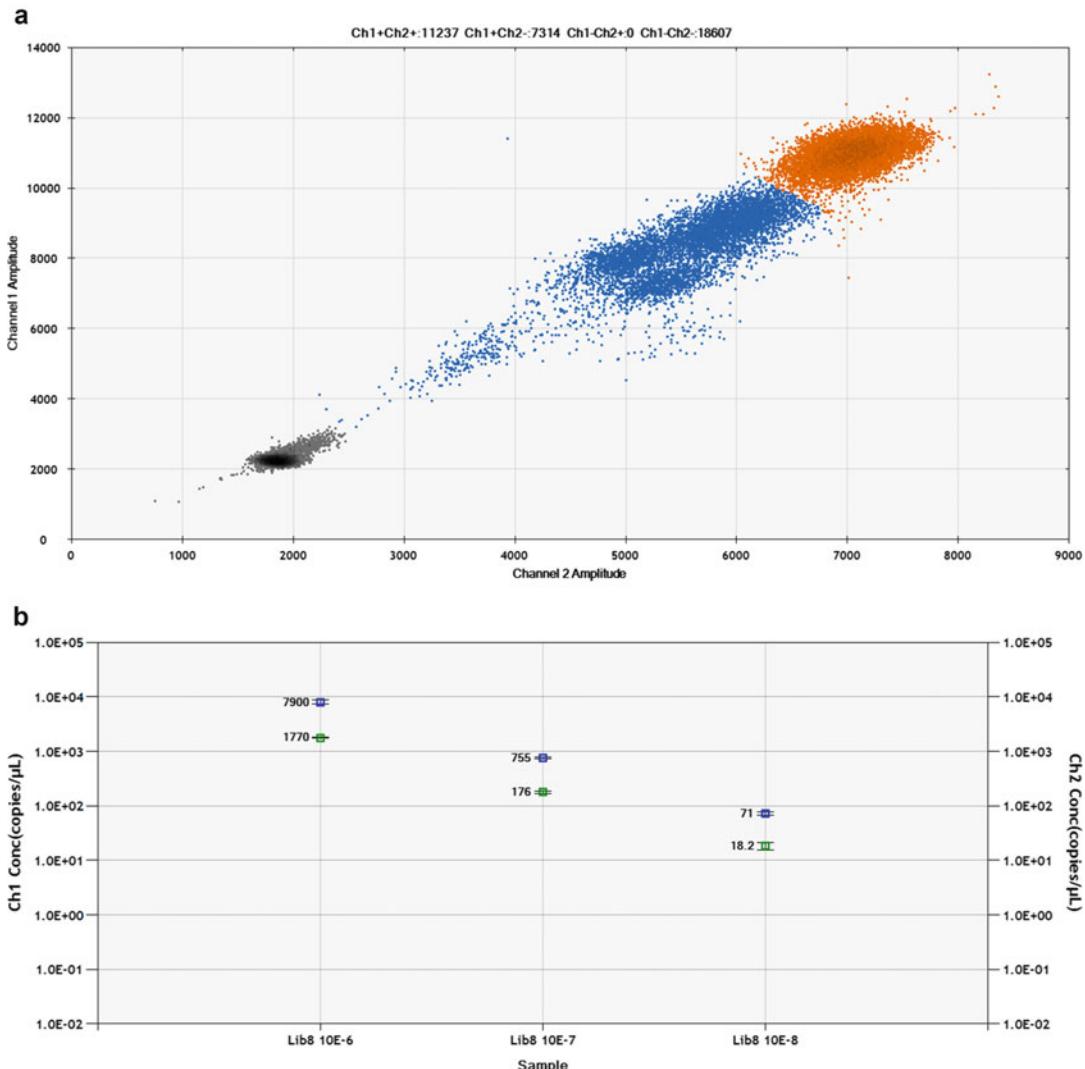


**Fig. 3** A two-channel plot representing the inverse relationship between library insert size and fluorescent intensity. The highest fluorescent intensity droplets are highlighted within the outlined red oval which correspond to library molecules with no inserts and are adapter dimers



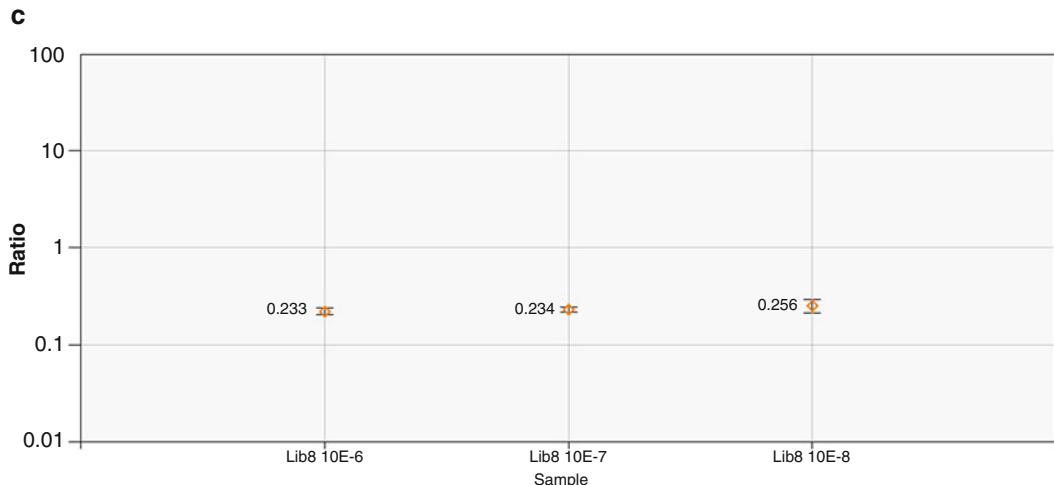
**Fig. 4** Highlighted in blue oval, are ill-formed library molecules in droplets which potentially have multiple P7 adapters attributing a more FAM fluorescent intensity to those droplets compared to the well-formed library molecules in the orange, double positive, cluster

Fig. 3. Figure 4 demonstrates poor library formation of a few droplets with potentially multiple P7 adapters added to these library molecules in these droplets as they are not along the double positive cluster, but rather possess more FAM fluorescence compared to the vast majority of well-formed library molecules in the double positive cluster. Figures 5 and 6 represent three ddPCR well's worth of data for a dilution series of library molecules spanning  $10^{-6}$ – $10^{-8}$  dilutions. Figure 5a is the 2-channel plot which has been manually thresholded using the lasso tool in QuantSoft software. The main cluster of library molecules which have insert are intentionally selected as FAM positive only (blue), and the empty library molecules which are adapter dimers, are intentionally selected as double positive for FAM and HEX (orange). These intentional assignments for the library clusters enable the user to

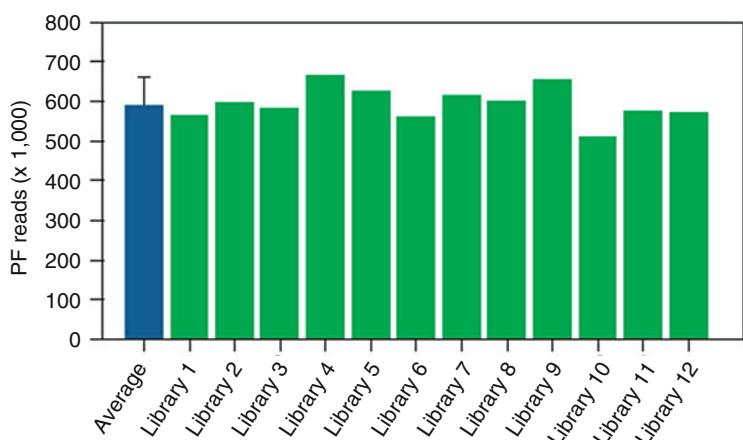


**Fig. 5** (a) Coloring scheme of library clusters which enable quantification of well-formed library molecules (blue), and adapter-dimers (orange). (b) Quantification of concentration of a dilution series from the data shown in Fig. 5 which demonstrates good linearity over the tenfold dilution series. The blue points represent well-formed library plus adapter dimers, and the green points represent only the adapter dimers quantification. (c) QuantaSoft can take the ratio of the HEX concentration to FAM concentration for the data in Fig. 5b. This essentially calculates what percent of the adapter-dimers constitute the total library being assayed

quantify the total number of library molecules which could occupy space on sequencing flow cell, specifically those library molecules which contain either insert or empty adapter-dimers and these are quantified in the FAM channel data in QuantaSoft. By assigning the adapter dimers cluster as double positive, it is possible to simultaneously quantify what percent of the total library molecules are empty adapter dimers, which can be quantified in the HEX channel



**Fig. 5** (continued)



**Fig. 6** Twelve TruSeq LT DNA libraries were quantified by ddPCR and subsequently pooled together at equimolar concentrations based on the ddPCR results. The combined libraries were then loaded on a single sequencing run and exhibit excellent balancing within less than 15% difference (total number of reads passing filter for each library). *PF* passing filter

on QuantaSoft. Figure 5b plots the concentration values across the dilutions series in the FAM and HEX channels according to the assignments in Fig. 5a. The data obtained in Fig. 5c allows the user to compute the percent of adapter-dimers (HEX) compared to the total library molecules (FAM) values. Figure 6 demonstrates the power of ddPCR library QC by quantitating 12 different sequencing libraries, pooling them together based on the results at equimolar concentrations, and sequencing them to show excellent concordance between ddPCR balancing and the number of reads obtained from each library on the illumina sequencer.

## 5 Conclusion

Digital PCR offers exceptional sensitivity for NGS library quantification and can improve sequencing results by providing precise measurements of library loaded on to the sequencing surface. Digital PCR also offers unique capabilities such as library QC in the same measurement which give additional valuable information such as a relative library insert sizing visualization and library quality. By using digital PCR to evaluate NGS libraries, it is possible to measure such quality metrics as adapter dimers in a library and overall library construction quality. Undoubtedly as digital PCR applications expand, the synergy between digital PCR and NGS will likewise grow, especially in the areas of rare species quantification and verification, and absolute counting.

## References

1. Rothberg JM et al (2011) An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475:348–352
2. Margulies M (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437(7057):376–380
3. Bentley DR et al (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456(7218):53–59
4. Hindson BJ (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem* 83 (22):8604–8610

## Further Reading

- Laurie MT, Bertout JA, Taylor SD, Burton JN, Shendure JA, Bielas JH (2013) Simultaneous digital quantification and fluorescence-based size characterization of massively parallel sequencing libraries. *BioTechniques* 55 (2):61–67
- Buehler B, Hogrefe HH, Scott G, Ravi H, Pabón-Peña C, O'Brien S, Formosa R, Happe S (2010) Rapid quantification of DNA libraries for next-generation sequencing. *Methods* 50:S15–S18
- QPCR NGS Library Quantification Kit (illumina GA), Part Number G4880A, Publication number 5990–3065. (2012). Agilent

Quail MA, Kozarewa I, Smith F, Scally A, Stephens PJ, Durbin R, Swerdlow H, Turner DJ (2008) A large genome center's improvements to the Illumina sequencing system. *Nat Methods* 5:1005–1010

Sequencing Library qPCR Quantification Guide Catalog # SY-930-1010, Part #11322363 Rev. C (2011). Illumina

White RA III, Blainey PC, Fan HC, Quake SR (2009) Digital PCR provides sensitive and absolute calibration for high throughput sequencing. *BMC Genomics* 10:116–128



# Chapter 28

## Phasing DNA Markers Using Digital PCR

John Regan and George Karlin-Neumann

### Abstract

Besides quantifying the absolute number of copies of known DNA targets, digital PCR can also be used to assess whether two nonpolymorphic gene sequences or two heterozygous markers reside on the same DNA molecule (i.e., are physically linked). Some useful linkage applications include: phasing variants to define a haplotype; genotyping of inversions; determining the presence of multimarker pathogenic bacteria in a metagenomic sample; and assessing DNA integrity. This chapter describes an efficient and cost-effective method for analyzing linkage of any two genetic sequences up to at least 200 Kb apart, including phasing of heterozygous markers such as that which occur abundantly in the cystic fibrosis transmembrane conductance regulator (CFTR) gene.

**Key words** Phasing, Linkage, Haplotype, Compound heterozygote, Complex alleles, Co-localization, Genotyping inversions, DNA integrity, DNA sizing, Large DNA isolation, Digital PCR, Droplet Digital PCR

---

### 1 Introduction

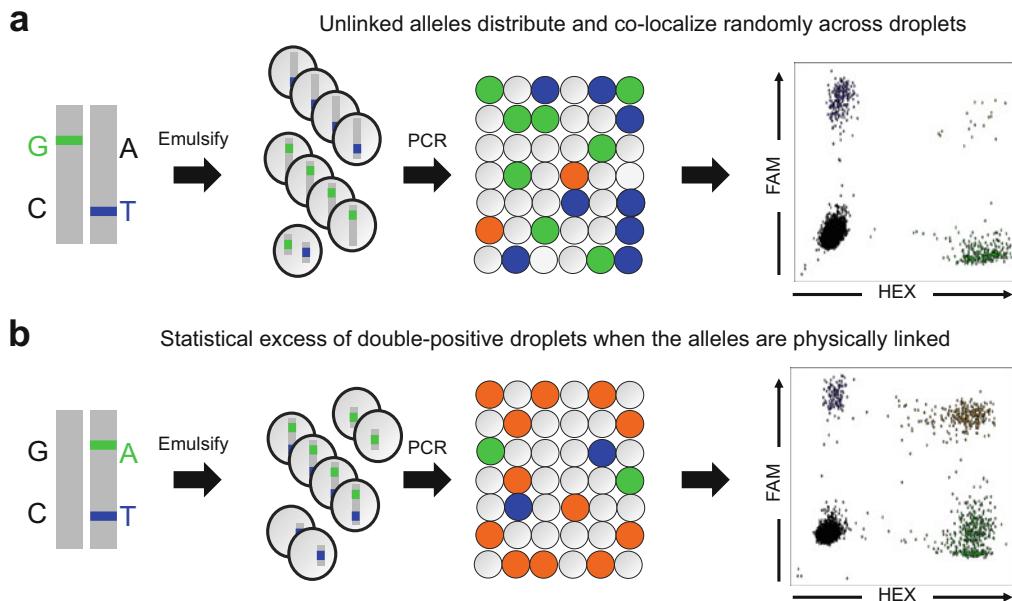
A number of biological and clinical questions require not only knowing whether given sequences are present in a sample—something capably done by digital PCR—but also whether these sequences (or markers) exist on the same chromosome or homologue, or how close they are to one another in a chromosomal region. Some examples include: in clinical microbiology, determining whether multimarker pathogenic bacteria such as Shiga toxin-producing *Escherichia coli* or methicillin-resistant *Staphylococcus aureus* are present [1, 2]; for personalized medicine, deciding whether an advanced-stage non-small-cell lung carcinoma harboring both T790M and C797S mutations in the EGFR oncogene, is likely to be responsive (if they are in *trans*) or insensitive (if they are in *cis*) to a combination of tyrosine kinase inhibitors (TKIs) [3]; and for diagnosing cystic fibrosis in a patient with two sequenced mutations in their CFTR gene, determining whether the patient has two pathogenic CFTR alleles (i.e., a compound heterozygote, where they are in *trans*) and thus will have the disease, or whether they

have one pathogenic allele harboring both mutations (a complex allele, where they are in *cis*) and one wild-type allele, and will likely not have the disease [4]. Besides the last two examples, assessment of phasing also facilitates the investigation of allelic imbalance for understanding the transcriptional basis of various phenotypes (See chapter 23 entitled “Using droplet digital PCR to analyze allele-specific RNA expression” by Kamataki et al.), and enables the verification and genotyping of chromosomal inversions which may underlie both disease and nondisease phenotypes [5]. Lastly, digital PCR has been used to assess the intactness of a DNA sample to determine its fitness for a subsequent analysis, such as phasing itself [4] or genotyping [6].

An underappreciated feature of digital PCR is that the very act of partitioning a sample into many small microreactions to make digital copy number measurements, also enables the determination of whether two targets being queried are physically linked (i.e., found on the same molecule or chromosome). For diploid organisms, markers (or targets) can be either homozygous or heterozygous in nature. In both cases, physical linkage can be established if the DNA sample(s) in question has sufficiently large molecules to contain both marker sequences. If both markers are heterozygous (e.g., Aa and Bb), then in principle it is possible to phase the markers using digital PCR, that is, to determine whether the sample is AB/ab or Ab/aB (the underline denotes markers are on the same chromosome). Prior to digital PCR, obtaining linkage or phase information was extremely difficult. Commonly used approaches including single-molecule dilution (SMD) followed by amplification with either gel analysis or sequencing [7–9], long-range PCR [10], linking-emulsion PCR [11, 12], and cloning strategies [13] are either very laborious and/or greatly limited in their ability to phase distant markers. As a result, studies that required long-range phase information (5–200 kb) were very hard to carry out.

The advent of two-color commercial digital PCR instruments with partitions created by either fixed nanochambers or droplets has greatly simplified carrying out linkage experiments. By labeling each of the two markers in question with distinguishable fluorophores, digital PCR systems are able to determine linkage status *since linked markers increase the number of double-positive partitions that contain both markers of interest above what would be expected from random co-localization of two unlinked markers*. Figure 1 provides an overview of how digital PCR quantifies the presence of linked markers. This inference can be tested by appropriately digesting the DNA sample with a restriction enzyme to separate the two markers into now unlinked DNA molecules.

The first utilization of a commercial digital PCR system to determine haplotype was recently published [4]. The authors focused on phasing common mutations in the cystic fibrosis gene (CFTR). Their work demonstrated phasing of markers between



**Fig. 1** Measuring linked species. Marker-specific fluorescent probes (FAM: blue, HEX: green) are used to detect the markers of interest in undigested genomic DNA. Following endpoint PCR, droplets are positive for one fluorophore (blue or green), both fluorophores (orange), or neither (black). **(a)** Markers on DNA from different chromosomes (or DNA molecules) partition independently into droplets (including some that randomly colocalize into the same droplets and become double positive droplets). **(b)** Markers on the same unbroken strand systematically colocalize into the same droplets because they are physically linked, resulting in an overabundance of double-positive droplets above what would be expected from chance co-localization of unlinked markers

110 kb and 116 kb apart with 100% concordance to inheritance information. The method, called “Drop-Phase,” requires less than 4 h, is scalable to hundreds of samples, is effective at genomic distances up to ~200 kb, and is low cost.

It is important to stress that the distance over which linkage measurements can be made is a function of the intactness of DNA being used. Published work has shown there to be a significant difference in the ability of some DNA extraction kits to preserve the intactness of DNA during the purification process [4]. Currently, it is possible to establish linkage from cultured cells out to ~200 kb. Other extraction kits or new advances in sample preparation could extend the distance at which linkage between two markers can be established. However, at this time, it is unclear if mega-based length DNA will require larger partition sizes to enable very long-distance linkage measurements.

This chapter presents two protocols to assess linkage. The first, the standard linkage protocol, is used to determine whether two nonpolymorphic sequences are linked. This protocol is independent of zygosity. The second, the phasing protocol, is particularly aimed at diploid organisms and is used to phase heterozygous markers.

---

## 2 Materials

The protocols below specifically reference Bio-Rad's Droplet Digital™ (ddPCR™) system. However, the principles presented are likely applicable to other commercial and noncommercial digital PCR systems.

### 2.1 Equipment

1. Droplet Digital PCR System composed of a droplet generator (manual or AutoDG) and droplet reader (QX100 or QX200) (Bio-Rad).
2. PXI™ PCR Plate Sealer (Bio-Rad).
3. C1000 Touch™ Thermal Cycler (Bio-Rad).

### 2.2 DNA Samples

1. DNA samples should be extracted with a sample preparation chemistry that is known to sufficiently preserve the intactness of the DNA and used as soon as possible, preferably before freezing the DNA (*see Notes 1–3*).
2. Optional: High molecular weight intact genomic control DNA (*see Note 4*).

### 2.3 Reagents and Consumables

1. ddPCR SuperMix for Probes (no dUTP) (Bio-Rad).
2. DNA Suspension Buffer/10 mM Tris, 0.1mM EDTA, pH 8.0 (Teknova).
3. Nuclease-free water (Ambion).
4. 20× FAM- and HEX-labeled TaqMan probe assays (Final 1× concentration of each primer is 900 nM, and of each probe is 250 nM).
5. Droplet Generation Oil for Probes (Bio-Rad).
6. Appropriate restriction enzyme for confirming linkage (choice depends on target sequences and cuts the sample DNA between the two markers being evaluated but not within the assay amplicons themselves).
7. DG8™ Droplet Generator Cartridges and Gaskets (for manual droplet generation) or DG32™ Automated Droplet Generator Cartridges (both from Bio-Rad).
8. Rainin normal-bore P20 and P200 filter pipette tips (Note: Lower quality pipette tips may result in shredding of droplets).
9. ddPCR 96-well semiskirted plates (Bio-Rad, Cat. #12001925). These plates are specially formulated to maximize stability and recovery of droplets and their use is strongly recommended.
10. PCR Plate Heat Seal, foil, pierceable (Bio-Rad, #1814040).

### 3 Methods

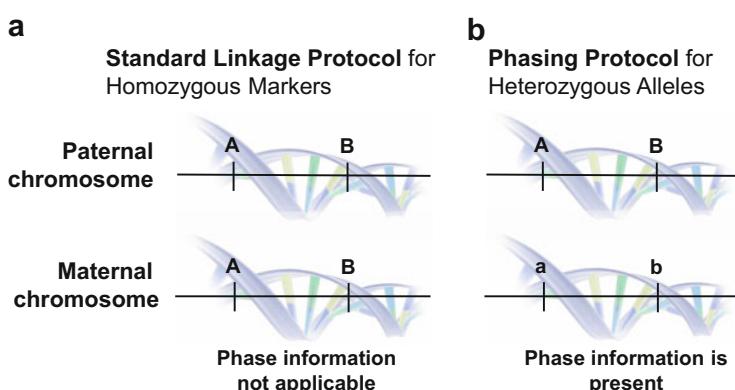
Two protocols are provided: a **standard linkage protocol** where phase is not applicable and a **phasing protocol** for determining the *cis-trans* relationship between two heterozygous markers (Fig. 2). The standard linkage protocol is used to establish whether two DNA sequences are linked. It can be applied to any markers regardless of ploidy or zygosity. In contrast, the phasing protocol, is most frequently applied to diploid organisms to determine whether linkage exists between two heterozygous sites, but in principle can be applied to organisms with higher order zygosities.

#### 3.1 Preparation of Genomic DNA

Choosing the correct extraction methodology and sample type is critical (*see Note 5*). When starting from cultured cells or cells from whole blood, silica membrane methods generally only allow for linkage measurements out to 30 kb. In contrast, Life Technologies' PrepFiler Forensic DNA Extraction Kit regularly yields DNA that is sufficiently intact to permit linkage measurements out to 200 kb. Others have reported success with a phenol-chloroform protocol (personal communication).

For the PrepFiler kit:

1. Suspend the cells at  $\sim 7 \times 10^6$  cells/mL in 1× PBS and remove 40  $\mu$ L for extraction.
2. Follow the manufacturer's recommendations, except reduce the lysis step down to 10 min at 70 °C with no shaking.
3. In addition, use gentle end-over-end inversion rather than vortexing, except after the addition of 180  $\mu$ L isopropyl alcohol, when the lowest rpm setting of a vortex mixer should be used.
4. Generally, minimize the speed and duration of the centrifugation steps.
5. Elute the extracted DNA in 50  $\mu$ L of kit-provided elution buffer.



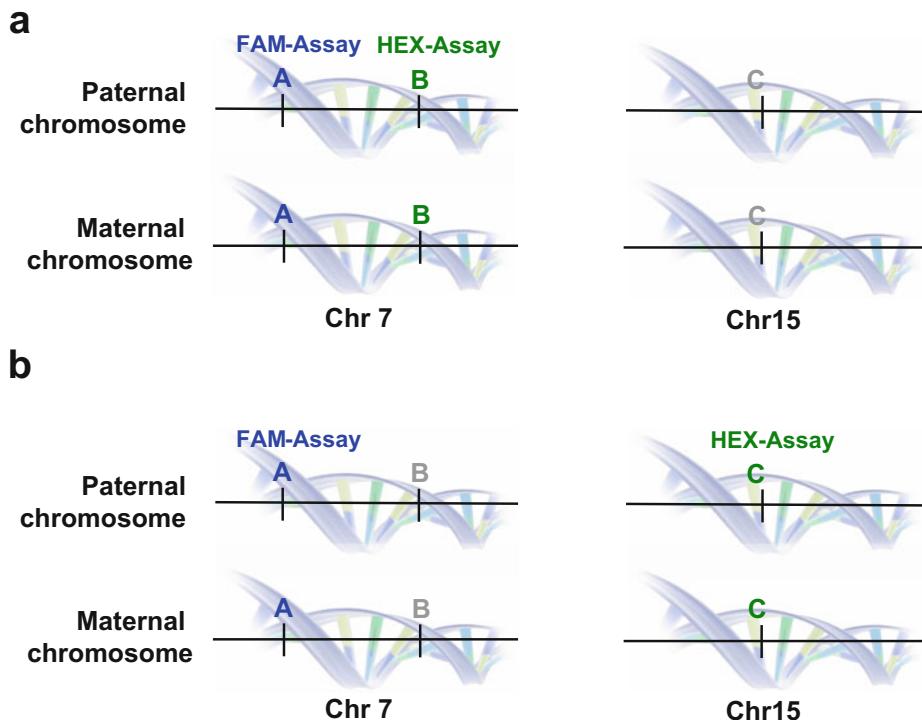
**Fig. 2** Use cases for linkage studies. (a) Standard linkage protocol is used when querying two loci that are homozygous in nature. (b) Phasing protocol is used when querying two loci that are heterozygous in nature

### 3.2 Standard Linkage Protocol

#### 3.2.1 Design/Selection and Optimization of Standard Linkage Assays

The standard linkage protocol should be used when verifying whether two genes are present in the same organism (i.e., pathogen virulence testing), to genotype inversions [5], and to test whether the DNA sample is intact enough (*see Note 1*) to permit successful phasing determination (Fig. 3).

For standard linkage experiments, a pair of conventional FAM-labeled and HEX-labeled TaqMan probe assays are combined in a single tube. Suitable wet-lab validated assays for this purpose can be obtained through the Bio-Rad assay design website (<https://www.bio-rad.com/digital-assays/#/>), or custom designs can be provided from raw sequence at both this site and that of IDT (<http://www.idtdna.com/calc/analyzer>). (*See Subheading 3.3.1 below for additional design information.*) These assays must



**Fig. 3** Duplex assays for standard linkage protocol. (a) The test duplex includes an FAM-labeled assay targeting the “A” locus and a HEX-labeled assay targeting the “B” locus, which is presumed to be on the same chromosome as the “A” locus and close enough such that sufficient intact DNA remains for a positive linkage measurement. (b) The control duplex includes an FAM-labeled assay targeting the “A” locus and a HEX-labeled assay targeting the “C” locus, which is known to be on a different chromosome than the “A” locus. This control duplex provides a realistic background estimate of linkage caused by nonrandomly partitioned DNA due to entanglement of large DNA molecules (as occurs in the test reaction). This approach is preferred over use of a restriction enzyme to separate “A” from “B” in the test duplex reaction because digested DNA more easily distributes randomly into droplets than undigested DNA and may underestimate the amount of background linkage

produce sufficient fluorescence amplitude separation between positive and negative droplet clusters for unambiguous assignment of droplets into clusters and provide specific target detection. Generally, standard ddPCR thermal cycling parameters are used with annealing temperatures in the 55–60 °C range. Optimal cycling conditions for the assays should be determined in advance by running a thermal gradient experiment from ~5 °C below to ~5 °C above the predicted  $T_m$  (melting temperature) of the assay primers.

### 3.2.2 Assembling Standard Linkage Reactions

1. Create 3 technical replicates for both test and control ddPCR reactions, by assembling the triplicate reaction mixtures detailed in Tables 1a and 1b, respectively.
2. Vortex both 60 µL reaction mixtures, and then briefly centrifuge the contents.

**Table 1a**  
**Assay components for test reactions**

Test reaction component	Volume (µL)
ddPCR SuperMix for Probes (no dUTP)	37.5
20× FAM Assay for Marker A	3.75
20× HEX Assay for Marker B	3.75
Nuclease Free Water	15
Total	60

**Table 1b**  
**Assay components for control reactions (markers known not to be linked)**

Control reaction component	Volume (µL)
ddPCR SuperMix for Probes (no dUTP)	37.5
20× FAM Assay for Marker A	3.75
20× HEX Assay for Marker C*	3.75
Nuclease Free Water	15
Total	60

\*Marker C should reside on a different chromosome from Marker A. If the organism is a haploid organism (e.g., bacterium), instead of assembling a different duplex assay for the control, maintain the original A and B duplex, but instead add a restriction enzyme at 10 U/60 µL reaction to specifically cut the DNA between A and B (see Note 6).

3. Dispense 18  $\mu\text{L}$  of each reaction mixture into 3 wells of a 96-well plate (6 wells total).
4. Individually add 4.5  $\mu\text{L}$  of undigested DNA (5 ng/ $\mu\text{L}$ ) to each well using a P-20 pipette tip (*see Note 7*). Do not pipet up and down with the P-20 pipette tip (*see Note 8*).
5. Using a normal-bore pipette tip and a P-200 with the volume set to 18  $\mu\text{L}$ , slowly pipet up and down 20 times, making sure not to create bubbles (*see Note 9*). If bubbles are created, spin down the plate at 1500 RCF for 1 min.
6. Generate droplets by either:
  - (a) Adding 20  $\mu\text{L}$  of the ddPCR reaction mixture into a droplet generation cartridge using a multichannel pipettor and manually generate droplets according to the manufacturer's instructions, followed by their careful transfer to a 96-well PCR plate, using a P-50 multichannel pipettor. Repeat until all samples have been partitioned into droplets and transferred to the PCR plate, 8 samples at a time.
  - (b) Or by leaving the entire 22.5  $\mu\text{L}$  reaction in the plate, sealing the plate, and transferring the sample plate to an AutoDG for automated droplet generation.
  - (c) Note that regardless of whether droplets are generated by the manual or automated systems, all 8 wells of a DG chip must be filled with either sample reaction mix or control buffer, along with DG oil in the corresponding DG8 wells.
7. Seal the droplet plate using a heat-seal foil and the PX1 heat sealer set to 180 °C for 5 s.
8. Thermal-cycle the droplets using conditions previously determined to be optimal for all assays on the plate. Standard ddPCR cycling conditions (95 °C for 10 min; 40 cycles of 94 °C for 30 s followed by 60 °C for 1 min; 98 °C for 10 min; 12 °C hold; 2.5 °C per cycle ramp rate for all steps) should be followed, except for using the optimal anneal/extension temperature if it differs from 60 °C.
9. Read the droplets on the QX100 or QX200 Droplet Reader.

### 3.2.3 Analyzing the Data for Linkage of Markers

1. In QuantaSoft Analysis software, manually examine each well for droplet quality in 2D fluorescent amplitude droplet plots. Discard any wells that have an abundance of droplets (i.e., >5%) that align on a diagonal running through the double negative cluster, which is indicative of sheared droplets. (If this occurs excessively, pay attention to pipetting the droplets more slowly when transferring them to the PCR plate and use high quality pipette tips which do not shed plastic

Total concentration of *A*, *B*, and *C*, including unlinked and linked molecules

Concentrations of linked *AB* and linked *AC* only

Duplex #1	T	Target	Status	Conc(copies/µL)	Duplex #2	Ch1-Ch2-	Linkage	AcceptedDroplets
		Marker A	Manual	<b>105</b>		13336	58.8	15288
Duplex #2		Marker B	Manual	<b>104</b>		13336	58.8	15288
		Marker A	Manual	<b>116</b>		11710	1.5	14344
		Marker C	Manual	<b>108</b>		11710	1.5	14344

**Fig. 4** Screen captures from a QuantaSoft table showing the total concentration of *A*, *B*, and *C* markers, regardless of linkage status, and the “Linkage” column, which displays only the concentration of linked molecules (*AB* and *AC*) in copies/µL

particulates into DNA preps and ddPCR reactions.) Select the remaining wells and classify the droplets while viewing the 2D fluorescent amplitude plot.

- Using the data available in the QX100/200 QuantaSoft Analysis table (Fig. 4), calculate the percentage of linked *AB* molecules for the wells containing assays for both *A* and *B* using Eq. 1. Handle each well separately, rather than merging the wells, to protect against the possibility that one poorly mixed well skews the analysis. If the replicates are not within 20% of each other for the percentage of linked molecules, then the mixing of the DNA sample into the reactions was not sufficient and caution should be exercised when interpreting the findings.

$$\% \text{ Linked } \underline{\text{AB}} = \left( \frac{2\lambda_{\underline{\text{AB}}}}{(\lambda_A + \lambda_B)} \right) 100 \quad (1)$$

- Where  $\lambda_{\underline{\text{AB}}}$  is the value found in the “Linkage” column, which is expressed in copies/µL.
- Where  $\lambda_A$  is the value taken from the “Conc (copies/µL)” column for the concentration of Marker *A*.  $\lambda_A$  is a summation of the linked and unlinked *A* molecules.
- Where  $\lambda_B$  is the value taken from the “Conc (copies/µL)” column for the concentration of Marker *B*.  $\lambda_B$  is a summation of the linked and unlinked *B* molecules.
- Using the numbers shown for Duplex #1 in Fig. 4, 56.3% of *A* and *B* molecules are linked :  $\underline{\text{AB}} = \left( \frac{2*58.8}{(105+104)} \right) 100$ .

- Using the data available in the QuantaSoft table, calculate the percentage of linked *AC* molecules for the *AC* duplexes using Eq. 2.

$$\% \text{ Linked } \underline{\text{AC}} = \left( \frac{2\lambda_{\underline{\text{AC}}}}{(\lambda_A + \lambda_C)} \right) 100 \quad (2)$$

- Where  $\lambda_{\underline{\text{AC}}}$  is the value found in the “Linkage” column, which is expressed in copies/ $\mu\text{L}$ .
  - Where  $\lambda_A$  is the value taken from the “Conc(copies/ $\mu\text{L}$ )” column for the concentration of marker A.  $\lambda_A$  is a summation of the linked and unlinked A molecules.
  - Where  $\lambda_C$  is the value taken from the “Conc (copies/ $\mu\text{L}$ )” column for the concentration of marker C, *here representing the unlinked control locus*.  $\lambda_C$  is a summation of the linked and unlinked C molecules.
  - Using the numbers shown for Duplex #2 in Fig. 4, 1.3% of A and C molecules are linked :  $\underline{\text{AC}} = \left( \frac{2*1.5}{(116+108)} \right) 100$ . In this case, the 1.3% of linked  $\underline{\text{AC}}$  molecules is artificial linkage due to incomplete mixing (*see Notes 8 and 9*).
4. For the three test and three control reactions, calculate the combined mean and standard deviation for each.
  5. Multiply each standard deviation by 2 and compare the lower confidence bar of the test samples with the upper confidence bar of the control samples. If these bars overlap, linkage cannot be claimed. If they do not overlap, then linkage can be claimed at >95% confidence.

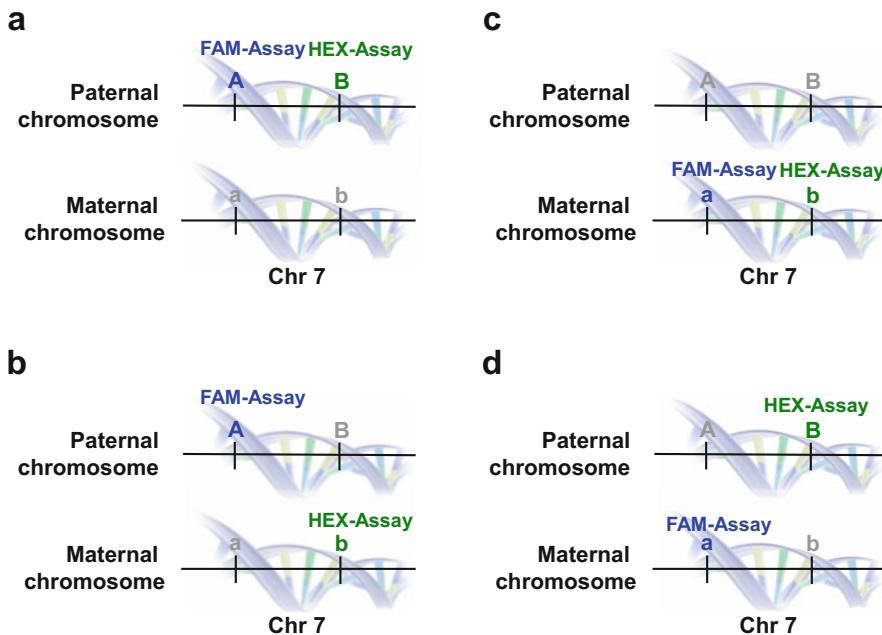
### 3.3 Phasing Protocol

This phasing protocol is designed to determine if alleles “a” and “b” are on the same chromosome. Across two heterozygous sites, there are four unique FAM-HEX duplexes that can be assembled to determine linkage (Fig. 5). A minimum of two of these duplexes needs to be assembled, tested, and compared as one serves as a test duplex and the other a control duplex for artificial linkage. If the percentage of linked molecules is expected to be less than 25%, all four duplexes should be assembled, tested and compared.

#### 3.3.1 Design/Selection and Optimization of Phasing Assays

Assay design and data interpretation for phasing experiments is slightly more complicated than for the Standard Linkage Protocol due to the heterozygous nature of the targeted loci. This can result in cross-reactivity of an assay with the nontargeted allele as well as competition for reagents in droplets containing both alleles, and can produce up to 16 different clusters on a 2D fluorescent amplitude plot (Fig. 6).

The droplets in the plot shown in Fig. 6 are easily classified—as indicated in the figure, but it is not uncommon to observe plots where the clusters are not as orthogonal due to cross-reactivity, and even experts in digital PCR can easily misclassify the clusters.

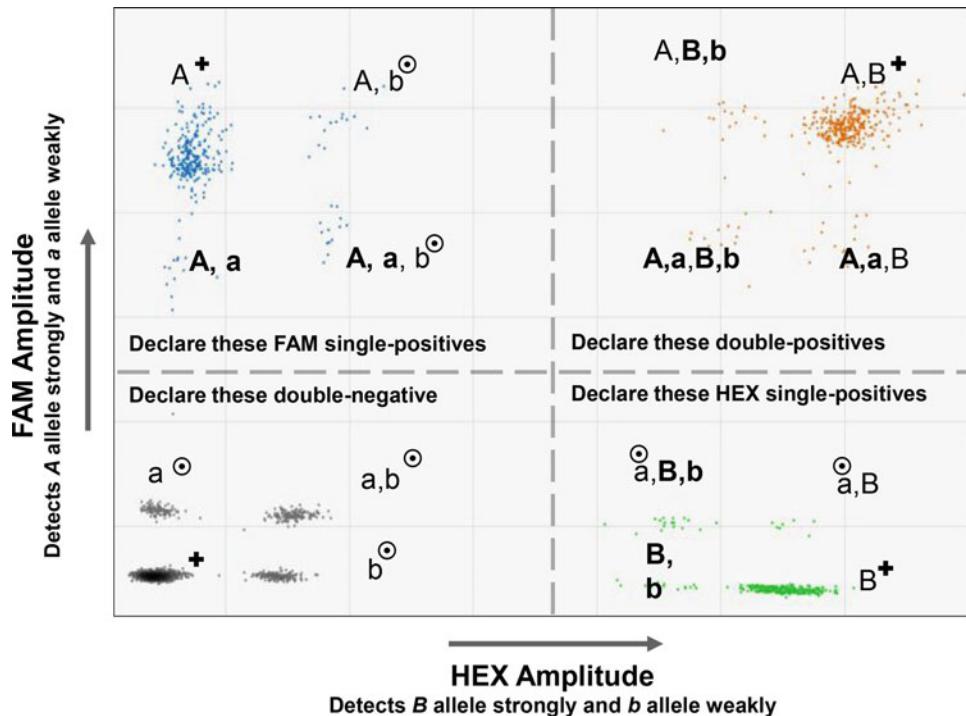


**Fig. 5** Duplexes assembled in phasing protocol. **(a)** Assays targeting the “A” and “B” alleles are combined. **(b)** Assays targeting the “A” and “b” alleles are combined. **(c)** Assays targeting the “a” and “b” alleles are combined. **(d)** Assays targeting the “a” and “B” alleles are combined

Accordingly, phasing assays must not only have sufficient amplitude for easy positive droplet classification, but steps should be taken to minimize cross-reactivity. At a minimum, the probes should either have a  $T_m \sim 1\text{--}2$  °C above the 55 °C melting temperature of the assay primers or should be designed with a  $T_m$  enhancer, such as linked nucleic acids (LNAs from Exiqon) or a minor groove binder (MGB from ThermoFisher Scientific). In addition, adding a dark competitor probe at 250 nM final concentration (equivalent to the concentration of the probe with which it is competing) is strongly recommended (see Figs. 7 and 8, and Note 10).

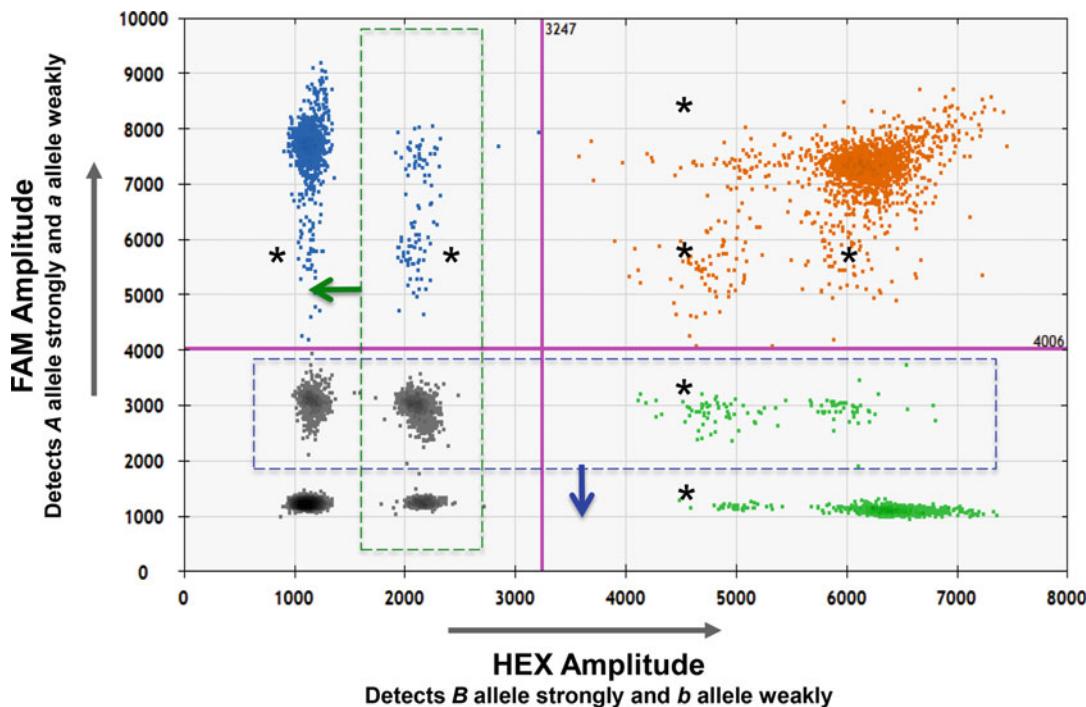
For common clinically relevant variants, Bio-Rad offers many wet-lab validated and *in silico* designed assays for digital PCR. In addition, their website allows for users to input COSMIC ID, gene symbol and/or amino acid change, and/or nucleotide change. Alternatively, raw sequence can be submitted to their website for custom assay design (<https://www.bio-rad.com/digital-assays/#/>). If ordering from this site, it is still necessary to design a dark competitor probe for the nontargeted alleles. To do so:

1. Go to IDT’s oligo analyzer (<http://www.idtdna.com/calc/analyzer>) and input ~27 bases of the nontargeted sequence with the base of interest centrally located.
2. Adjust the default parameters such that Mg<sup>++</sup> is 3.8 mM and the dNTPs concentration is 0.8 mM.



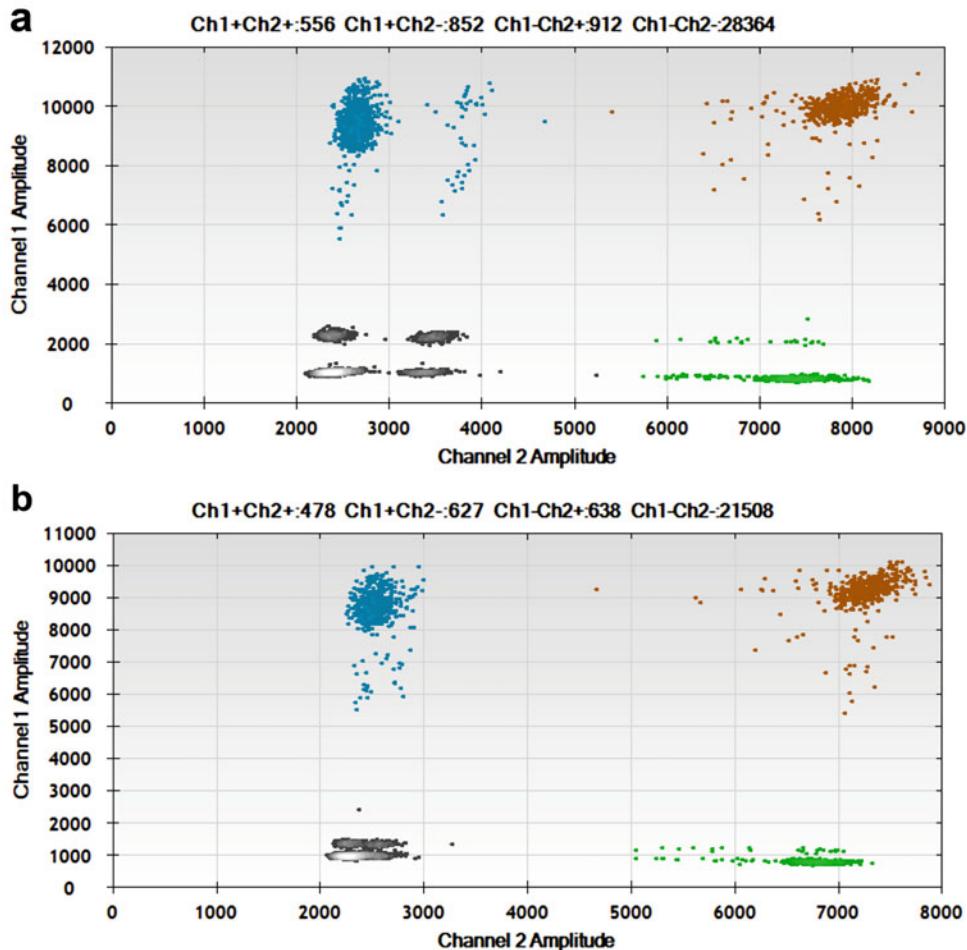
**Fig. 6** Droplet cluster identification and classification in the presence of assay competition and cross-reacting fluorescent probes (i.e., NO dark competitor probes present). Major clusters (designated as “ $\oplus$ ” in the figure) are comprised of single- and double-positive droplets that contain the FAM (A)- and/or HEX (B)-targeted alleles (without the nontargeted “a” and “b” alleles), as well as the double-negative clusters (−, −). Additionally, at high enough DNA loads per reaction, some droplets contain both targeted and nontargeted alleles (A and a or B and b). When both alleles are present in the same droplet (alleles shown in **bold** on the plot), the two templates compete for the same primer pair resulting in droplets with reduced fluorescent amplitude since only half of the end-product is specific to the included probe (and this is not mitigated by the presence of dark competitor probes). Droplets that lack the targeted allele, but contain the nontargeted allele (denoted by the  $\odot$  symbol), experience probe cross-reactivity that generates as many as 7 additional clusters that are offset from the four major clusters ( $\oplus$ ). If dark competitor probes had been added to the reaction these offset clusters are often driven toward the axes to overlap with the main FAM and HEX clusters, including the competition clusters (see Fig. 7)

3. Click “Analyze” and examine the Melt Temperature. Ideally, design the dark competitor to have a melting temperature which is equivalent to the  $T_m$  of the fluorescent probe with which it is competing (~57 °C; *see* above and **Note 10**).
4. If the melting temperature is too high, remove bases equally from the ends such that the allele of interest remains centrally located within the competitor probe.
5. After settling on the design, the probe should be ordered with a 3' BHQ quencher (/3BHQ\_1/), but no fluorophore.



**Fig. 7** Influence of dark competitor probes (see Note 10). Phasing experiments without competitor probes (as in this 2D plot) can have as many as 16 clusters both because the probes for each targeted sequence will often cross-react with the nontargeted allele when present and because the presence of both alleles in a single droplet will result in competition that lowers the targeted alleles endpoint yield and thus fluorescence amplitude creating additional subclusters. Adding dark competitor probes into the reaction drives the cross-reacting clusters surrounded by a blue box, (those clusters having targets cross-reacting with the “A” probe) and a green box (those clusters having targets cross-reacting with the “B” probe) into the conventional clusters that lie closest to the HEX and FAM axes, respectively. The seven clusters denoted with an asterisk (\*) are present due to PCR competition, since a single primer pair amplifies both targeted and nontargeted alleles and thus results in only about ½ of the final PCR product in these droplets with commensurately reduced hydrolysis of the corresponding fluorescent probe. To reduce the number of droplets in these competition-based clusters, load less DNA into the reaction

6. Spike this dark competitor probe into the fluorescent assay targeting the allele of interest such that the final concentration in the ddPCR reaction mixture is 250 nM (see Note 10).
- 3.3.2 Assembling Phasing Protocol Reactions**
1. Plan to create three technical replicates as in the Standard Linkage Protocol above.
  2. If a high degree of linkage is expected for the alleles of interest (e.g., >25%), only two of the four reactions (e.g., Tables 2a and 2b) are required. However, if <25% of the alleles are expected to be linked, then all four duplexes should be run (see Tables 2a, 2b, 2c & 2d), as indicated in Fig. 5a–d, to mitigate the risk that artificial linkage confounds the results (see Note 9). In such cases, two of the duplexes should have measurable linkage for the alleles in question and the other two should not.



**Fig. 8** Influence of adding dark competitor probes to phasing assay. **(a)** Phasing assay without the addition of dark competitor probes. **(b)** Phasing assay with the addition of dark competitor probes. Notice that the dark competitor probes do not eliminate the additional minor clusters resulting from the simultaneous presence of both targeted and nontargeted alleles in droplets (i.e., the A,a; B,b; A,B,b; A,a,B and A,a,B,b clusters seen in Fig. 6 still remain)

**Table 2a**  
**Assay components for AB reactions**

<b>AB test reaction component</b>	<b>Volume (<math>\mu\text{L}</math>)</b>
ddPCR SuperMix for Probes (no dUTP)	37.5
20× FAM Assay for Allele A	3.75
20× HEX Assay for Allele B	3.75
Nuclease Free Water	15
Total	60

**Table 2b**  
**Assay components for *Ab* reactions**

<b><i>Ab</i> test reaction component</b>	<b>Volume (<math>\mu</math>L)</b>
ddPCR SuperMix for Probes (no dUTP)	37.5
20× FAM Assay for Allele <i>A</i>	3.75
20× HEX Assay for Allele <i>b</i>	3.75
Nuclease Free Water	15
Total	60

**Table 2c**  
**Assay components for *ab* reactions (only required if low % of linkage expected)**

<b><i>ab</i> test reaction component</b>	<b>Volume (<math>\mu</math>L)</b>
ddPCR SuperMix for Probes (no dUTP)	37.5
20× FAM Assay for Allele <i>a</i>	3.75
20× HEX Assay for Allele <i>b</i>	3.75
Nuclease Free Water	15
Total	60

**Table 2d**  
**Assay components for *aB* reactions (only required if low % of linkage expected)**

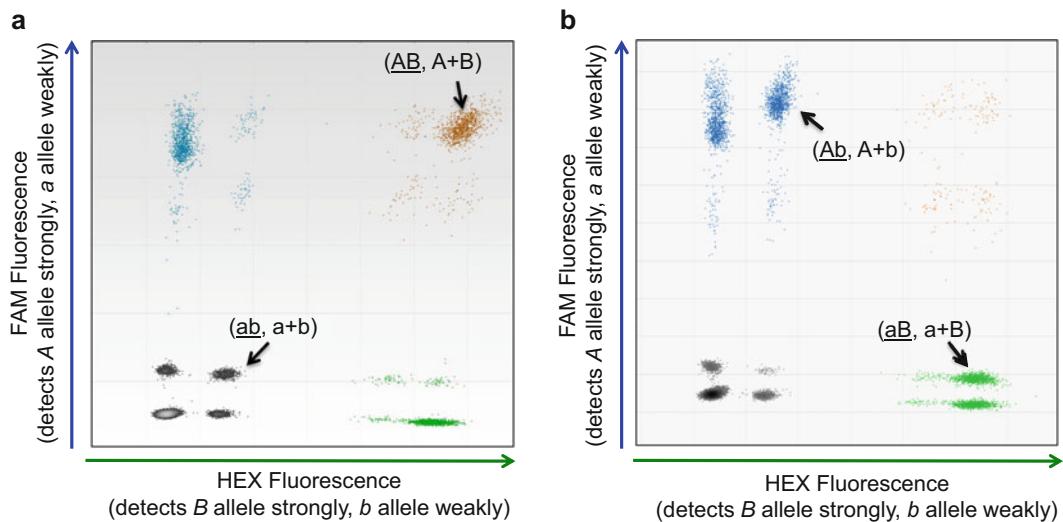
<b><i>aB</i> test reaction component</b>	<b>Volume (<math>\mu</math>L)</b>
ddPCR SuperMix for Probes (no dUTP)	37.5
20× FAM Assay for Allele <i>a</i>	3.75
20× HEX Assay for Allele <i>B</i>	3.75
Nuclease Free Water	15
Total	60

3. Vortex the 60  $\mu$ L reaction mixtures, and then briefly centrifuge the contents.
4. Dispense 18  $\mu$ L of each reaction mixture into 3 wells of a 96-well plate (12 wells total, if running all 4 duplexes, 6 wells, if running only 2 of the duplexes).

5. Individually add 4.5  $\mu\text{L}$  of undigested DNA (10 ng/ $\mu\text{L}$ ) to each well using a P-20 pipette tip (*see Note 7*). Do not pipet up and down with the P-20 pipette tip (*see Note 8*).
6. Using a normal-bore pipette tip and a P-200 with the volume set to 18  $\mu\text{L}$ , slowly pipet up and down 20 times, making sure not to create bubbles (*see Note 9*). If bubbles are created, spin down the plate at 1500 RCF for 1 min.
7. Generate droplets by either:
  - (a) Transferring 20  $\mu\text{L}$  of the reaction mixture into a droplet generation cartridge and manually generating droplets according to the manufacturer's instructions, followed by their careful transfer to a 96-well PCR plate, using a P-50 multichannel pipettor. Repeat until all samples have been partitioned into droplets and transferred to the PCR plate, 8 samples at a time.
  - (b) Or by leaving the entire 22.5  $\mu\text{L}$  reaction in the plate, sealing the plate, and transferring it to an AutoDG for automated droplet generation.
8. Seal the droplet plate using a heat-seal foil and the PX1 heat sealer set to 180 °C for 5 s.
9. Thermal-cycle the droplets using conditions previously determined to be optimal for all assays on the plate. Standard ddPCR cycling conditions (95 °C for 10 min; 40 cycles of 94 °C for 30 s followed by 55 °C for 1 min; 98 °C for 10 min; 12 °C hold; 2.5 °C per cycle ramp rate for all steps) should be followed, except for using the optimal anneal/extension temperature if it differs from 55 °C.
10. Read the droplets on the QX100 or QX200 Droplet Reader.

### *3.3.3 Analyzing the Data for Linkage of Markers*

1. Barring significant differences between the three wells per duplex (e.g., obvious shearing of the droplets in one well), select all three wells and classify the droplets while viewing the 2D fluorescent amplitude plot. (*See Subheading 3.2.3, step 1*, for further guidance on well analysis). In phasing experiments, PCR competition between alleles and probe cross-reactivity can make classifying droplets difficult. For guidance on understanding these principles and their influence on droplet cluster location within the 2D fluorescent plot *see Figs. 7 and 9*.
2. If running the AB, Ab, ab, and aB duplexes, use the data available in the QX100/200 QuantaSoft Analysis table (as seen in Fig. 10), to calculate the percentage of linked molecules.



**Fig. 9** Shifting cluster location. For linked targets, the location of the cluster(s) with an excess of droplets shifts depending on whether the included duplex is a perfect match for the linked alleles or if one of the assays only cross-reacts with the linked allele. **(a)** When the sample (AB, ab) contains linked alleles that are specifically targeted by the included assays (FAM-labeled assay detects “A” and HEX-labeled assay detects “B”), then there is an overabundance of the major clusters denoted with arrowheads. **(b)** When a different sample of the opposite genotype (Ab, aB) is queried using the same duplex without dark competitor probes, the cluster with excess double-positive droplets shifts to the ones where the nontargeted allele is detected through cross-reactivity (denoted with arrowheads). If dark competitor probes were used, these cross-reactive droplets would ideally merge with the conventional major FAM (A)-only and HEX (B)-only clusters in Fig. 6, effectively causing an overabundance of single-positive droplets and an underabundance of double-positive droplets and the expected number of double-positive droplets containing the two targeted alleles (all of the A,B-containing droplets in the upper right of Fig. 6) would be no greater than by chance alone

Total concentration of A, B, a, and b alleles, including unlinked and linked versions				Concentrations of linked <u>AB</u> , <u>Ab</u> , <u>ab</u> , and <u>aB</u> only		
T	Target	Status	Conc(copies/ $\mu$ L)	Ch1-Ch2-	Linkage	AcceptedDroplets
Duplex #1	Allele A	Manual	104	13469	60.7	15430
	Allele B	Manual	106	13469	60.7	15430
Duplex #2	Allele A	Manual	104	13406	0	16246
	Allele b	Manual	107	13406	0	16246
Duplex #3	Allele a	Manual	109	13717	59.3	15799
	Allele b	Manual	106	13717	59.3	15799
Duplex #4	Allele a	Manual	108	11342	0.931	13789
	Allele B	Manual	107	11342	0.931	13789

**Fig. 10** Screen captures from a QuantaSoft table showing the total concentration of A, B, a, and b alleles, regardless of linkage status, and the “Linkage” column, which displays only the concentration of linked molecules (AB, Ab, ab, and aB) in copies/ $\mu$ L

3. For calculating the percentage of linked AB molecules use Eq. 3.

$$\% \text{ Linked } \underline{\text{AB}} = \left( \frac{2\lambda_{\underline{\text{AB}}}}{(\lambda_A + \lambda_B)} \right) 100 \quad (3)$$

- Where  $\lambda_{\underline{\text{AB}}}$  is the value found in the “Linkage” column, which is expressed in copies/ $\mu\text{L}$ .
- Where  $\lambda_A$  is the value taken from the “Conc(copies/ $\mu\text{L}$ )” column for the concentration of allele A.  $\lambda_A$  is a summation of the linked and unlinked A molecules.
- Where  $\lambda_B$  is the value taken from the “Conc(copies/ $\mu\text{L}$ )” column for the concentration of allele B.  $\lambda_B$  is a summation of the linked and unlinked B molecules.
- Using the numbers shown for Duplex #1 in Fig. 10, 57.8% of A and B molecules are linked:  $\underline{\text{AB}} = \left( \frac{2*60.7}{(104+106)} \right) 100$ .

4. For calculating the percentage of linked Ab molecules use Eq. 4.

$$\% \text{ Linked } \underline{\text{Ab}} = \left( \frac{2\lambda_{\underline{\text{Ab}}}}{(\lambda_A + \lambda_b)} \right) 100 \quad (4)$$

- Where  $\lambda_{\underline{\text{Ab}}}$  is the value found in the “Linkage” column, which is expressed in copies/ $\mu\text{L}$ .
- Where  $\lambda_A$  is the value taken from the “Conc (copies/ $\mu\text{L}$ )” column for the concentration of allele A.  $\lambda_A$  is a summation of the linked and unlinked A molecules.
- Where  $\lambda_b$  is the value taken from the “Conc (copies/ $\mu\text{L}$ )” column for the concentration of allele b.  $\lambda_b$  is a summation of the linked and unlinked b molecules.
- Using the numbers shown for Duplex #2 in Fig. 10, 0% of A and b molecules are linked:  $\underline{\text{Ab}} = \left( \frac{2*0}{(104+107)} \right) 100$ .

In this case, no linkage is found between the A and b markers (see Notes 8 and 9).

5. For calculating the percentage of linked ab molecules use Eq. 5.

$$\% \text{ Linked } \underline{\text{ab}} = \left( \frac{2\lambda_{\underline{\text{ab}}}}{(\lambda_a + \lambda_b)} \right) 100 \quad (5)$$

- Where  $\lambda_{\underline{\text{ab}}}$  is the value found in the “Linkage” column, which is expressed in copies/ $\mu\text{L}$ .
- Where  $\lambda_a$  is the value taken from the “Conc (copies/ $\mu\text{L}$ )” column for the concentration of allele a.  $\lambda_a$  is a summation of the linked and unlinked a molecules.
- Where  $\lambda_b$  is the value taken from the “Conc (copies/ $\mu\text{L}$ )” column for the concentration of allele b.  $\lambda_b$  is a summation of the linked and unlinked b molecules.

- Using the numbers shown for Duplex #3 in Fig. 10, 55.2% of *a* and *b* molecules are linked:  $\underline{ab} = \left( \frac{2*59.3}{(109+106)} \right) 100$ .
6. For calculating the percentage of linked *aB* molecules use Eq. 6.

$$\% \text{Linked } \underline{aB} = \left( \frac{2\lambda_{aB}}{(\lambda_a + \lambda_B)} \right) 100 \quad (6)$$

- Where  $\lambda_{aB}$  is the value found in the “Linkage” column, which is expressed in copies/ $\mu\text{L}$ .
  - Where  $\lambda_a$  is the value taken from the “Conc (copies/ $\mu\text{L}$ )” column for the concentration of allele *a*.  $\lambda_a$  is a summation of the linked and unlinked *a* molecules.
  - Where  $\lambda_B$  is the value taken from the “Conc (copies/ $\mu\text{L}$ )” column for the concentration of allele *B*.  $\lambda_B$  is a summation of the linked and unlinked *B* molecules.
  - Using the numbers shown for Duplex #4 in Fig. 10, 0.87% of *a* and *B* molecules are linked  $\underline{aB} = \left( \frac{2*0.93}{(108+107)} \right) 100$ . In this case, the 0.87% of linked *aB* molecules is artificial linkage due to incomplete mixing (see Notes 8 and 9).
7. For the *AB*, *Ab*, *ab*, and *aB* reactions, calculate the mean and standard deviation across triplicate reactions.
8. Compare the % of linked markers and using the criteria in steps 9 and 10 below, determine which of the following three options is most likely:
- Alleles AB and ab are linked.*
  - Alleles Ab and aB are linked.*
  - The Aa and Bb loci are either on different chromosomes or the DNA linking the two loci is too fragmented to allow for phasing of the alleles.*
9. In the case where *AB* and *ab* are linked, these duplexes should yield similar percentages of linked alleles. Similarly, if *Ab* and *aB* are linked, these duplexes should yield similar percentages of linked alleles. In contrast, the other two duplexes serve as controls and any linkage observed with them is artificial linkage due to insufficient mixing. It is acceptable to have artificial linkage in the control wells, so long as sufficient technical replicates were performed and the test wells have statistically more linked molecules than the control wells.
10. If the difference in the percent linked between the four duplexes is small, determine the mean and standard deviation for the percentage of linked molecules for each of the four duplexes. For the two most linked (which should both be for the phased alleles) and two least linked duplexes (which should

both be for the unphased alleles), recalculate the mean and standard deviation to include the six wells across each of these higher and lower linkage duplexes, respectively. Next, subtract two standard deviations from the mean of the higher linkage duplexes and add two standard deviations to the mean of the lower linkage duplexes and check to see if the error bars overlap (i.e., if the first value is greater than the second). If the error bars overlap, either the alleles are not on the same chromosome or the DNA sample is too fragmented for making a linkage determination.

---

#### 4 Notes

1. When starting from cultured cells, DNA extracted using silica-based columns (e.g., Qiagen's DNeasy Blood & Tissue Kit) and polysaccharide precipitation-based chemistry (e.g., ThermoFisher Scientific's PrepFiler Forensic DNA Extraction Kit) permit linkage assessment out to ~ 25 kb and ~200 kb, respectively [4]. If working with an untested extraction chemistry, first perform a "mile marker series" as described in Regan et al. [4], where a series of duplex assays are assembled, each having the same HEX-labeled assay specific to an anchor sequence, and each with a unique FAM-labeled assay specific to a conserved region of the genome at increasing distances away from the anchor sequence. For each distance (e.g., 1, 10, 33, 60, 100, 150, 200 kb) separating the mile markers and the anchor sequence calculate the percentage of linked molecules and fit these data with an exponential regression line to learn the theoretical limit in kilobases at which a successful linkage measurement can be made. This information can then be used to assess whether linkage status between other markers of interest can be determined with this DNA prep.
2. There may be an upper limit to the length of DNA strands that are compatible with reliable droplet formation from Bio-Rad's QX200 digital PCR system. The QX200 generates droplets that are ~ 0.85 nL in size; and we believe the upper limit for DNA strand length in these droplets is ~500 kb. Longer DNA strands become entangled and cause localized viscosity changes, which increase droplet polydispersity with a bias toward larger droplet formation. Excessive DNA input levels (e.g., >66 ng/well) of highly intact DNA, such as Promega's human DNA (Cat # G1471 or G1521), will completely prevent droplet formation. In light of this, when using undigested DNA, always confirm average droplet counts are in excess of 15,000/well if using QuantaSoft 1.7 or in excess of 12,000/well if using an older version of QuantaSoft. If working with

DNA that regularly fails to form droplets, mix the sample more vigorously than would otherwise be recommended to slightly shear the DNA before going into droplet generation.

3. After the DNA is extracted, immediately perform the linkage tests, as storing the DNA for extended periods will cause additional DNA fragmentation. If the experiment is to be performed within a week's time, leave the sample at 4 °C. If the sample is likely not to be processed for more than a week, freezing is recommended, but avoid repetitive freeze-thaw cycles. Always store DNA for linkage analysis in a buffered solution to minimize hydrolysis (e.g., Teknova's DNA suspension buffer: 10 mM Tris, 0.1 mM EDTA, pH 8.0).
4. In evaluating different sample preparation chemistries for their ability to purify intact DNA, as of this writing, Promega's human DNA (<https://www.promega.com/products/biochemicals-and-labware/nucleic-acids/genomic-dna/>) may serve as a good control DNA. Promega DNA is highly intact and accordingly very viscous and enables linkage of two markers separated by at least 200 Kb to be detected by the procedure described in this chapter. In the absence of restriction enzyme digestion or aggressive pipetting, it does not allow droplet formation when included in ddPCR reactions at >66 ng/well (>1 copy/droplet).
5. DNA extracted from plasma (cfDNA) or FFPE is generally too fragmented to justify a linkage experiment that includes two different assays. Instead, a single primer-pair assay (<500 bp in length) that contains FAM- and HEX- labeled probes for the markers of interest can often be used to establish linkage using the protocol described in this chapter.
6. A temptation in linkage experiments is to include a restriction enzyme in the control wells, where the restriction enzyme digests the DNA linking the two markers. However, digested DNA mixes more easily to homogeneity than undigested DNA where the DNA is entangled. To avoid incorrectly genotyping samples due to artificial linkage, when possible, use undigested DNA for the controls in linkage experiments. This requires varying the assays used between the test and control wells. An exception to this recommendation is for haploid genomes and for triplex inversion genotyping experiments [5] where the test and control assays are included in the same well.
7. When performing a linkage experiment, the amount of DNA put into the reaction to minimize statistical error is different from the amount of DNA (120 ng/well or ~1.6 copies/droplet (cpd)) needed for the most precise quantification given Bio-Rad's QX200 partition volume of 0.85 nL. Instead, for linkage applications, it is a balance between loading sufficient

DNA to maximize the number of linked molecules scored in a well (appearing as double-positive droplets), while minimizing the number of double-positive droplets due to randomly colocalized targets. Our modeling suggests that the optimal DNA load is ~0.2 target copies per droplet (cpd). For human genomic DNA, we typically add ~22.5 ng of highly intact human gDNA into a 22.5  $\mu$ L reaction, allowing for as much as 50% overestimation of the amplifiable DNA amount by UV measurement. If performing a phasing experiment, where only a single chromosome is targeted, this number should be doubled to ~40 ng/reaction. If the sample DNA is of generally smaller size distribution but still has some molecules large enough to contain both target sequences, it may be necessary to increase the number of wells run to achieve statistical confidence in a linkage result for that sample.

8. Undermixing the reaction mixtures can result in significant “artificial” linkage measurements. To protect against this risk, for the technical replicates, individually add and mix the DNA into wells which already contain Supermix and assay. This approach is preferable to splitting a large reaction mixture already containing DNA between replicate wells, as this may mask incomplete mixing and give inaccurate linkage measurements.
9. Insufficiently mixed samples result in underquantification of the sample, as the DNA is not randomly distributed into the droplets, resulting in a deficit of droplets containing the target DNA. Another way to think about this is that undermixed have more double-negative and more double-positive droplets than in a properly mixed sample. The higher number of double-negative droplets translates into a lower concentration measurement for both markers, while the higher number of double-positive droplets translates into a statistically significant linkage measurement, even though the markers may not be linked (i.e., artificial linkage through entanglement of DNA molecules). Unpublished data suggests that when using 80% stroke volume, it is necessary to pipet up and down 10 times for digested and 20 times for undigested samples to mix them to homogeneity. These same data found that wide-bore pipette tips require even more pipette strokes to achieve homogeneity. As a result, wide-bore pipette tips are not recommended unless the sample is pipetted up and down at least 40 times to produce sufficient mixing of DNA in the reaction; failure to mix sufficiently can lead to artificial linkage measurements as high as 40%.
10. Standard assays for phasing contain 900 nM primers, 250 nM each fluorescent probe, and 250 nM dark competitor probe. Dark competitor probes are designed to bind to the nontargeted allele to reduce probe cross-reactivity (e.g., dark

competitor is specific to Marker  $\alpha$  where the fluorescent probe is specific to Marker  $A$ ). Dark competitor probes have no 5' fluorophore but instead a 3' quencher which should match the quencher being used for the corresponding fluorescent probe (e.g., an Iowa Black quencher (IABkFQ) from IDT (Coralville, Iowa, USA)). The melting temperature of the dark competitor probe should be the same as the  $T_m$  of the fluorescent probe with which it is competing.

## Acknowledgments

Steven McCarroll, Nolan Kamitaki, Tina Legler, Samuel Maars, Mario Caceres, and Svilen Tzonev, Niels Klitgord, and Samantha Cooper contributed to the knowledge presented in these protocols.

## References

- Wolk DM, Struelens MJ, Pancholi P, Davis T, Della-Latta P, Fuller D, Picton E, Dickenson R, Denis O, Johnson D, Chapin K (2009) Rapid detection of *Staphylococcus Aureus* and methicillin-resistant *S. Aureus* (MRSA) in wound specimens and blood cultures: multicenter preclinical evaluation of the Cepheid Xpert MRSA/SA skin and soft tissue and blood culture assays. *J Clin Microbiol* 47(3):823–826. <https://doi.org/10.1128/JCM.01884-08>
- Muniesa M, Hammerl JA, Hertwig S, Appel B, Brussow H (2012) Shiga toxin-producing *Escherichia Coli* O104:H4: a new challenge for microbiology. *Appl Environ Microbiol* 78(12):4065–4073. <https://doi.org/10.1128/AEM.00217-12>
- Rosell R, Karachaliou N (2016) Lung cancer: using ctDNA to track EGFR and KRAS mutations in advanced-stage disease. *Nat Rev Clin Oncol* 13(7):401–402. <https://doi.org/10.1038/nrclinonc.2016.83>
- Regan JF, Kamitaki N, Legler T, Cooper S, Klitgord N, Karlin-Neumann G, Wong C, Hodges S, Koehler R, Tzonev S, McCarroll SA (2015) A rapid molecular approach for chromosomal phasing. *PLoS One* 10(3):e0118270. <https://doi.org/10.1371/journal.pone.0118270>
- Puig M, Pacheco S, Izquierdo D, Regan JF, Karlin-Neumann G, Cáceres M (2016) Droplet digital PCR (ddPCR)-based validation and genotyping of human inversions. *Eur J Hum Genet* 24(E Suppl. 1):353
- Didelot A, Kotsopoulos SK, Lupo A, Pekin D, Li X, Atochin I, Srinivasan P, Zhong Q, Olson J, Link D, Laurent-Puig P, Blons H, Hutchison JB, Taly V (2013) Multiplex picoliter-droplet digital PCR for quantitative assessment of DNA integrity in clinical samples. *Clin Chem* 59(5):815–823. <https://doi.org/10.1373/clinchem.2012.193409>
- Chen N, Schrijver I (2011) Allelic discrimination of cis-trans relationships by digital polymerase chain reaction: GJB2 (p.V27I/p.E114G) and CFTR (p.R117H/5T). *Genetics in medicine: official journal of the American College of Medical Genetics* 13(12):1025–1031. <https://doi.org/10.1097/GIM.0b013e3182272e0b>
- Paul P, Apgar J (2005) Single-molecule dilution and multiple displacement amplification for molecular haplotyping. *BioTechniques* 38(4):553–554, 556, 558–559
- Dear PH, Cook PR (1993) Happy mapping: linkage mapping using a physical analogue of meiosis. *Nucleic Acids Res* 21(1):13–20
- McDonald OG, Krynetski EY, Evans WE (2002) Molecular haplotyping of genomic DNA for multiple single-nucleotide polymorphisms located kilobases apart using long-range polymerase chain reaction and intramolecular ligation. *Pharmacogenetics* 12(2):93–99
- Wetmur JG, Kumar M, Zhang L, Palomeque C, Wallenstein S, Chen J (2005) Molecular haplotyping by linking emulsion PCR: analysis of paraoxonase 1 haplotypes and phenotypes. *Nucleic Acids Res* 33

- (8):2615–2619. <https://doi.org/10.1093/nar/gki556>
12. Wetmur JG, Chen J (2008) An emulsion polymerase chain reaction-based method for molecular haplotyping. *Methods Mol Biol* 410:351–361
13. Kitzman JO, Mackenzie AP, Adey A, Hiatt JB, Patwardhan RP, Sudmant PH, Ng SB, Alkan C, Qiu R, Eichler EE, Shendure J (2011) Haplotype-resolved genome sequencing of a Gujarati Indian individual. *Nat Biotechnol* 29 (1):59–63. <https://doi.org/10.1038/nbt.1740>



# Chapter 29

## ddTRAP: A Method for Sensitive and Precise Quantification of Telomerase Activity

Andrew T. Ludlow, Dawne Sheldon, Woodring E. Wright,  
and Jerry W. Shay

### Abstract

Telomerase is a cellular RNA template-dependent reverse transcriptase that adds telomere repeats to the 3' ends of chromosomes. Telomerase is expressed almost universally in tumor cells (>85%) to maintain telomere length, thus providing the ability of tumor cells to avoid senescence and to have unlimited replication ability, one of the key hallmarks of cancer. ddTRAP (droplet digital Telomere Repeat Amplification Protocol) is a two-step assay with whole cell lysates that utilizes a telomerase-mediated primer extension followed by droplet digital PCR (ddPCR) detection of extended products. The adoption of the TRAP assay to ddPCR has resulted in improved throughput, increased sensitivity and better repeatability of the TRAP assay. The protocol described below details our procedures for ddTRAP.

**Key words** Telomerase activity, Droplet digital PCR

---

### 1 Introduction

Telomere maintenance and telomerase enzyme activity are almost universal features of transformed cancer cells [1–4]. Telomeres are repeat DNA sequences (5'-TTAGGG<sub>n</sub>-3') found at the ends of chromosomes [5, 6]. Telomeres function the in completion DNA replication each cell cycle and also provide a "capping" function by preventing recombination, end-to-end fusions, and degradation at the chromosome ends [7, 8]. Telomeres, therefore prevent unwanted DNA damage signaling at chromosome ends and may be a potent initial tumor suppressor mechanism when telomeres become critically short and cells enter senescence [9, 10]. In normal diploid cells telomeres shorten with each cell division due to the end replication problem (i.e., due to the placement of the terminal RNA primer, the inability of DNA polymerase to fully replicate the lagging strand of DNA, and the lack of the enzyme telomerase there is a loss of about 60 nucleotides per cell division/DNA replication cycle). Almost all tumor cells adapt a telomere

maintenance mechanism that utilizes telomerase to maintain telomeres [1]. Telomerase is a ribonucleoprotein complex that utilizes a protein component with reverse transcriptase activity (hTERT) and a RNA template (*hTR* or *hTERC*) to add telomere repeats to chromosome ends [11]. The two major molecular properties that are commonly assayed are telomerase activity (ability of the enzyme to extend a substrate by adding nucleotides) and repeat addition processivity (the number of TTAGGG repeats added to a substrate) [12]. Since the vast majority of cancer cells utilize telomerase to maintain telomeres, telomerase activity an important transformation biomarker, and thus accurate detection is critical. The discovery that telomerase synthesizes telomeric repeats onto 3' ends of chromosomes led to the development of an assay for the detection and measurement of its activity in cells and tissues and was established over 20 years ago [1].

To measure the activity of telomerase the most common assay is the telomere-repeat amplification procedure or TRAP, a two-step procedure involving telomerase mediated primer extension and PCR-based detection of extended products. Briefly, cells are lysed in a buffer containing a detergent, then a portion of the lysate is mixed with an extension reaction containing the telomerase substrate (a DNA oligonucleotide) and dNTPs which are used by telomerase (if present in the sample) to add hexameric telomere repeats. Finally, the extended substrates are PCR amplified and detected (See Fig. 1 for a graphical representation of the ddTRAP assay). We have recently adapted the original TRAP assay that provided only relative quantification (quantitated to an internal standard DNA) to be detected with intercalating DNA dye (i.e., Evagreen<sup>®</sup>) on the Bio-Rad droplet digital PCR QX150/200 Eva-green<sup>®</sup> compatible droplet reader, we call this assay the “ddTRAP” [13]. The adaptation of the TRAP assay to droplet digital PCR improved the sensitivity, repeatability, and throughput of the TRAP assay [13]. These properties of ddTRAP have enabled single cell telomerase assays and the screening of genes and small molecules aimed at manipulating telomerase activity. The detailed procedure below describes how to perform a ddTRAP to assay and quantitate telomerase activity in adherent or suspension transformed cells and primary human cells in culture.

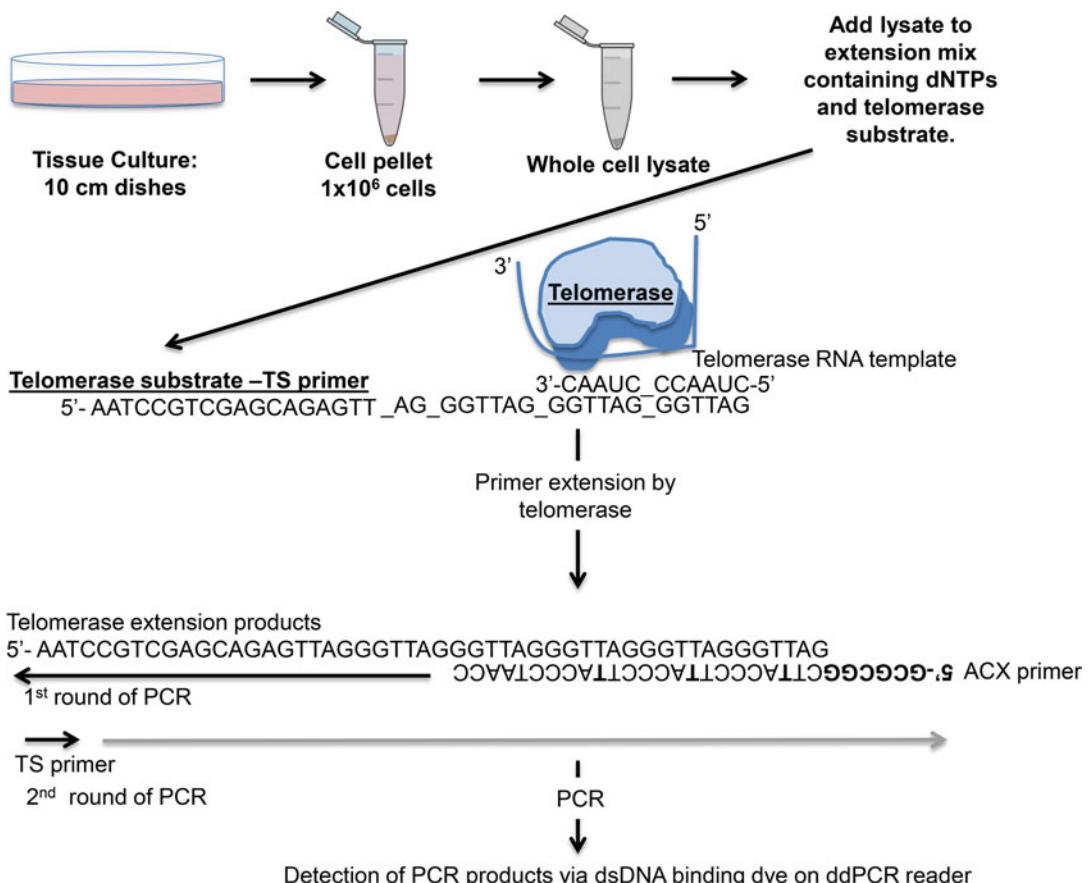
---

## 2 Materials

Prepare all solutions and set up all reactions using typical PCR precautions *see Note 1*.

### 2.1 Whole Cell Lysate Preparation

1. Use of a well-characterized telomerase positive cell line such as HeLa, HEK293, H1299, MDA-231, HCT116, or HT1080 as a positive control.



**Fig. 1** TRAP assay theory. Tissue culture cells are plated on a 10 cm culture dish and grown to about 90% confluence. Cells are then counted and pelleted in 1 million cell aliquots. Whole cell lysates are prepared and then added to the extension mix containing the telomerase substrate primer (TS primer) and dNTPs. The lysate containing active telomerase enzymes will extend the TS primer using the RNA template of the enzyme. The extension reaction is heat-inactivated, and the products are then PCR amplified in the presence of the reverse primer ACX and the forward primer TS to amplify the telomerase-extended substrates. The PCR products are double stranded and can be detected using Evagreen® dsDNA binding dye in the QX100/200 droplet digital PCR reader. Note—the bolded nucleotides in the ACX primer are mismatches with the telomerase-extended products

2. A cell line lacking telomerase activity such as U2OS (human osteosarcoma) or primary (nontransformed) BJ fibroblasts should be used as a negative control.
3. Scale up the cell line you plan to use as positive and negative controls and freeze aliquots for future use to avoid unintentional changes that could occur in long term cultured cell lines.
4. NP-40 lysis buffer (RNase/DNase-free): 10 mM Tris-HCl, pH 8.0; 1 mM MgCl<sub>2</sub>; 1 mM EDTA; 1% (vol/vol) NP-40; 0.25 mM sodium deoxycholate; 10% (vol/vol) glycerol; 150 mM NaCl; 5 mM β-mercaptoethanol; 0.1 mM AEBSF

(4-(2-Aminoethyl)benzenesulfonyl fluoride hydrochloride). (See Note 2 for cell extract preparation.)

5.  $\text{Ca}^{2+}$ - and  $\text{Mg}^{2+}$ -free PBS: 137 mM sodium chloride, 1.5 mM potassium phosphate, 7.2 mM sodium phosphate, 2.7 mM potassium chloride, pH 7.4.
6. RNase/DNase-free DEPC-treated dH<sub>2</sub>O (Ambion).
7. Cell counting device such as a hematocytometer, Coulter® counter, or automated cell counter (e.g., TC20, Bio-Rad). Alternatively, protein quantification can be determined prior to extension reaction with a BCA protein assay for cell extracts (see Note 2).
8. Bicinchoninic Acid protein assay kit—such as BCA protein assay kit (23225, ThermoFisher).
9. Spectrophotometer that can measure absorbance at 562 nm (for BCA assay).

## **2.2 Telomerase Substrate Primer Extension Reaction**

1. Ultrapure BSA (50 mg/mL, Ambion).
2. 50× dNTP mix (2.5 mM of each dATP, dCTP, dGTP, dTTP).
3. 10  $\mu\text{M}$  telomerase substrate, “TS” primer (telomerase substrate/primer for extension; 5'-AATCCGTCGAGCAGAGTT HPLC purified).
4. 10× TRAP buffer (RNase/DNase-free): 200 mM Tris-HCl, pH 8.3; 15 mM MgCl<sub>2</sub>; 630 mM KCl; 0.5% Tween 20; 10 mM EGTA.
5. Thin walled (250  $\mu\text{L}$  volume) PCR grade tubes/stripes/plates.
6. Thermocycler.
7. RNase/DNase-free DEPC-treated dH<sub>2</sub>O (Ambion).

## **2.3 ddPCR Detection of Telomerase Extended Substrates**

1. 10  $\mu\text{M}$  ACX primer (reverse amplification primer for detection—5'-GCGCGGCTTACCCCTTACCCCTTACCCATAACC HPLC purified—see Note 9 about primers).
2. 10  $\mu\text{M}$  telomerase substrate, “TS” primer (telomerase substrate for extension, also forward primer for detection—5'-AATCCGTCGAGCAGAGTT HPLC purified).
3. 2× Evagreen® super mix for ddPCR (Bio-Rad).
4. RNase/DNase-free DEPC-treated dH<sub>2</sub>O (Ambion).
5. Twin-Tec 96-well plate—Eppendorf 951020362 (Fisher).
6. Pierceable foil heat seal—(Bio-Rad #1814040).
7. Droplet generator cartridges (DG8) (Bio-Rad 186-3008).
8. Droplet generator oil (Bio-Rad 186-3005).
9. Droplet cartridge gaskets (DG8) (Bio-Rad 186-3009).
10. QX200 evagreen ddPCR supermix (Bio-Rad).

11. Thermocycler capable of fitting the 96-well skirted plates and adjusting the temperature ramp rate (i.e., Bio-Rad T100).
12. Droplet reader oil.
13. Droplet reader (Bio-Rad QX150/200) capable of reading Eva-green® DNA binding dye.
14. PX1 PCR plate sealer (Bio-Rad).

---

### 3 Method/Procedures

#### **3.1 Preparation of Tissue Culture Cells for ddTRAP Assay**

1. Culture cells to a desired density, typically to a density of 90% confluence (one needs to be consistent from experiment to experiment since in some cases cell density can affect telomerase activity). We typically grow cells on 10 cm dishes for TRAP assay analysis.
2. Trypsinize and count cells *see Note 1*.
3. Pellet cells. Centrifuge cells at  $2000 \times g$  at room temperature and aspirate media (*see Note 2*). Cells can be washed with cold 1× PBS following initial pelleting to remove carry over tissue culture media but this is not necessary.
4. If cell counting is not performed, ddTRAP assays can be performed using protein concentration (*see Note 13*).
5. Pellets can be used immediately by placing cells on ice in lysis buffer.
6. Pellets can be also be stored for later use by flash freezing in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  (*see Note 2*).

#### **3.2 Lysis of Cells**

1. Generate a whole cell lysate by adding 40  $\mu\text{L}$  of NP40 lysis buffer to pelleted cells (per 1 million cells) and incubate on ice for 40 min (*see Notes 12 and 14*).
2. During the lysis period periodically vortex or mix vigorously at least twice (e.g., at 15 min and 30 min of lysis). This is essential to ensure the cells are lysed and homogenous (*see Note 14*).

#### **3.3 Telomerase Primer Extension Reaction Setup**

1. While the cells are in lysis buffer, set up the extension reaction mix as shown below in Table 1 and store on ice.
2. Set up tubes for dilution of the lysates on ice.
3. Lysates need to be diluted to a cell equivalent of 1250 cells per microliter (2  $\mu\text{L}$  of lysate in 18  $\mu\text{L}$  of NP40 lysis buffer) on ice.
4. Prepare all diluted lysates before mixing with the extension reaction. Please *see Note 3* below for details on extension reactions.

**Table 1**  
**Extension reaction setup**

Reagent	X1 (volumes in $\mu\text{L}$ )	X10 (volumes in $\mu\text{L}$ )
10× TRAP buffer	5	50
50 mg/mL BSA	0.4	4
10 $\mu\text{M}$ TS primer	1	10
2.5 mM dNTPs	1	10
1250 cell equivalents lysate	1	—
Water ( $\text{dH}_2\text{O}$ )	40.6	406

**Table 2**  
**PCR setup**

Reagent	Volume ( $\mu\text{L}$ ) 1×	Volume ( $\mu\text{L}$ ) 10×	Volume ( $\mu\text{L}$ ) 20×
2× Evagreen ddPCR super mix	11	110	220
10 $\mu\text{M}$ TS	0.11	1.1	2.2
10 $\mu\text{M}$ ACX	0.11	1.1	2.2
25 cells/ $\mu\text{L}$ extension product	2	—	
$\text{dH}_2\text{O}$	8.8	88	176

5. In an ice cold tube containing 49  $\mu\text{L}$  of extension reaction mix add 1  $\mu\text{L}$  of diluted lysate to thin walled (250  $\mu\text{L}$ ) PCR tubes. Reactions should be mixed by pipetting.
6. When all lysates are added to the extension reaction mix, the PCR plate/tubes should be immediately moved to a thermocycler set with the following reaction conditions: 25 °C for 40 min, 95 °C for 5 min, and 12 °C hold. Please *see Notes 15–17* for further information on the lysis step and storage of lysates.

### 3.4 Digital PCR Setup

1. Using the PCR precautions outlined in *see Notes 4–11* prepare a master mix containing a final concentration of; 1× Evagreen ddPCR Super Mix v2.0 (Bio-Rad), 50 nM TS primer, 50 nM ACX primer, 50 cell equivalents or less of extension product and  $\text{dH}_2\text{O}$  to 20  $\mu\text{L}$  per sample, with 10% extra. See Table 2 for example volumes.
2. Allow reaction mix to reach room temperature (*see Notes 5–8*).
3. Set up the droplet generation (DG) cartridge. Load 20  $\mu\text{L}$  of PCR mixture into the sample well in the cartridge. Then add

70 µL of droplet generation oil for Evagreen® into the oil well. Secure a gasket on the DG cartridge.

4. Place assembled droplet cartridge into the DG machine.
5. Remove cartridge from the DG machine once droplet generation cycle is completed (about 90 s).
6. Transfer droplets to PCR plate. Remove gasket gently. Using an eight-channel pipet remove ~42 µL of emulsion (droplets) from the droplet wells in the cartridge and place into a 96-well plate (*see Note 9*).
7. Seal plate. Once all samples are loaded into the 96-well PCR plate, the plate must be foil sealed to prevent evaporation and light exposure of the emulsion (droplets).
8. Run PCR. Load the 96-well plate into the thermocycler and close the lid. In these experiments, we use a Bio-Rad T100. PCR reaction conditions—All ramp rates between temperature steps must be set to 2.5 °C/sec in order to achieve even heating of the reaction mixture.

Thermocycling:

95 °C for 5 min (activation of hot-start polymerase).

40 cycles of:

95 °C for 30 s.

54 °C for 30 s.

72 °C for 30 s.

22 °C hold.

Timing = 1 h 45 min.

### **3.5 Detection of the Telomerase Extended PCR**

#### **Amplified Substrates**

1. Read the droplets. Load the 96-well plate into the droplet reader plate holder matching well “A1” in the proper orientation.
2. Set up the reader. Open® software on the desktop of the laptop computer synchronized with the droplet reader. Double click well “A1” in the plate template. This will change the upper part of the screen to display “sample name,” “experiment,” “assay name,” and assay channels (fluorescence dyes fam or Vic®/hex). *See Note 10.*
3. Define the wells. Click “Experiment” and a pull-down menu with choices for experiment types will be displayed. Choose “absolute quantification” for the experiment type. Select wells (highlight wells) and click apply or press enter. With the same wells highlighted, Select assay channels as unknown channel 1 (6FAM/Evagreen®). Click “apply” or press enter. Name the assay—As an example for ddTRAP put the word ddTRAP and indicate the extension time used (i.e., ddTRAP 40 min ext.) and press apply. Name the samples.

4. Run the plate. Click run. A screen will prompt you to save the template. Give your plate a logical name (record this in your lab note book or experiment log in Excel® etc.). Click “save.” A screen will prompt you to pick the dye types (Fam/Vic or Fam/Hex) for ddTRAP pick FAM/Vic and if you want to read the wells in columns or rows.

### 3.6 Data Analysis

1. Determine if the sample is valid for further analysis. The data generated from the droplet reader is given in several formats. The most important component that should be addressed when analyzing ddTRAP data is the number of droplets generated. This can be queried several ways in the QuantaSoft® software. By clicking “events,” the total number of droplets per well can be visualized in a histogram. Alternatively under the “table” display the column titled “accepted droplets” will give the investigator the same information. We typically consider a sample valid for further analysis with more than 10,000 accepted droplets; more stringent criteria of 12,000 accepted droplets can also be applied, but the same criteria should be applied across all samples in a given experiment and all experiments in a particular manuscript unless noted in the figure legend.
2. Set the threshold. Setting the thresholds between positive and negative droplets can be a subjective task when analyzing digital PCR data and thus certain standard criteria should be applied for each new assay. For the ddTRAP assay, positive droplets typically fall between 6000 and 10,000 fluorescent units, however since telomerase generates amplicons of various sizes and is a “GC”-rich template, longer molecules (i.e., those with more repeats added) appear lower on the amplification plot. We use the following guides to set thresholds for ddTRAP assays: (1) A “no-template lysis buffer control” (NTC-LB) sample should be analyzed first, set threshold at 2000 fluorescence units above population of negative droplets (typically around 4000 fluorescence units see Fig. 1). There may be some background in this sample, however given the quantitative nature of ddTRAP, this background can and should be subtracted from all positive wells. See Fig. 1 for an example of background in NTC-LB controls. (2) Set thresholds for all wells. We typically use the same threshold for all wells as was set for the NTC-LB. Occasionally, the population of negative droplets will vary between wells and in this case a well specific background needs to be set to analyze the data. We recommend in this case that the threshold be set at a minimum of 2000 fluorescent units above the “negative population.” (3) Dealing with “rain.” “Rain” or intermediate droplets between the negative and positive populations are common

**Table 3a**  
**Data from select columns from a QuantaSoft spreadsheet**

Sample type	Experiment	Accepted droplets	Concentration (molecules per microliter)
NTC-LB	Abs quant	14,267	1.64
Positive control cancer line	Abs quant	12,587	43.8

with the ddTRAP assay and can have an impact on the quantitative results. Users must be cautious when setting thresholds as to not skew data one way or another. To avoid potential problematic “rain” we recommend using as few cell equivalents as possible and repeating analysis with samples that produce a large amount of “rain.” When this is not possible the user must set thresholds equivalently for all samples to be compared. Once the thresholds are set quantitative information is now available (*see Notes 18 and 19*).

3. Extracting quantitative data. In terms of the digital TRAP assay we refer to the data as “the number of telomerase extended substrates/products per number of cell equivalents analyzed.” This information is obtained after correction for background signal and number of cell equivalents analyzed. From the QuantaSoft® software refer to the “concentration in molecules per microliter” column in the tabular view of the software. This table can be exported as a .csv file and analyzed in Excel®. First we generate background corrected data by subtracting the NTC-LB control concentration (molecules per microliter) from each unknown sample and positive control. This background corrected value is then converted to total telomerase-extended products by multiplying by 20 (20 µL PCR volume), this number is then divided by the number of cell equivalents added to the extension reaction (typically 50). See example data below.
4. Determining telomerase-extended products per cell. If 1,000,000 cells are harvested, lysed in 40 µL, the lysate diluted 1:20 to give 1250 cells per microliter, 1 µL of the diluted lysate added to a 50 µL telomerase extension (25 cell equivalents per microliter) and then 2 µL of the extension (50 cell equivalents) added to a 20 µL PCR the For example calculation *see Note 20* (Table 3a and 3b).

**Table 3b**  
**Example calculations for telomerase-extended products per cell**

Data quantification	NTC-LB background correction'	Total telomerase extension products	Telomerase extension products per cell
Description of step	Subtract “NTC-LB” concentration from unknowns to generate background corrected data	Multiply background corrected concentration by 20	Divide total telomerase extension products by the number of cell equivalents added to the extension reaction. In this example 50 cell equivalents.
Positive control cancer line	42.16	843.2	16.864

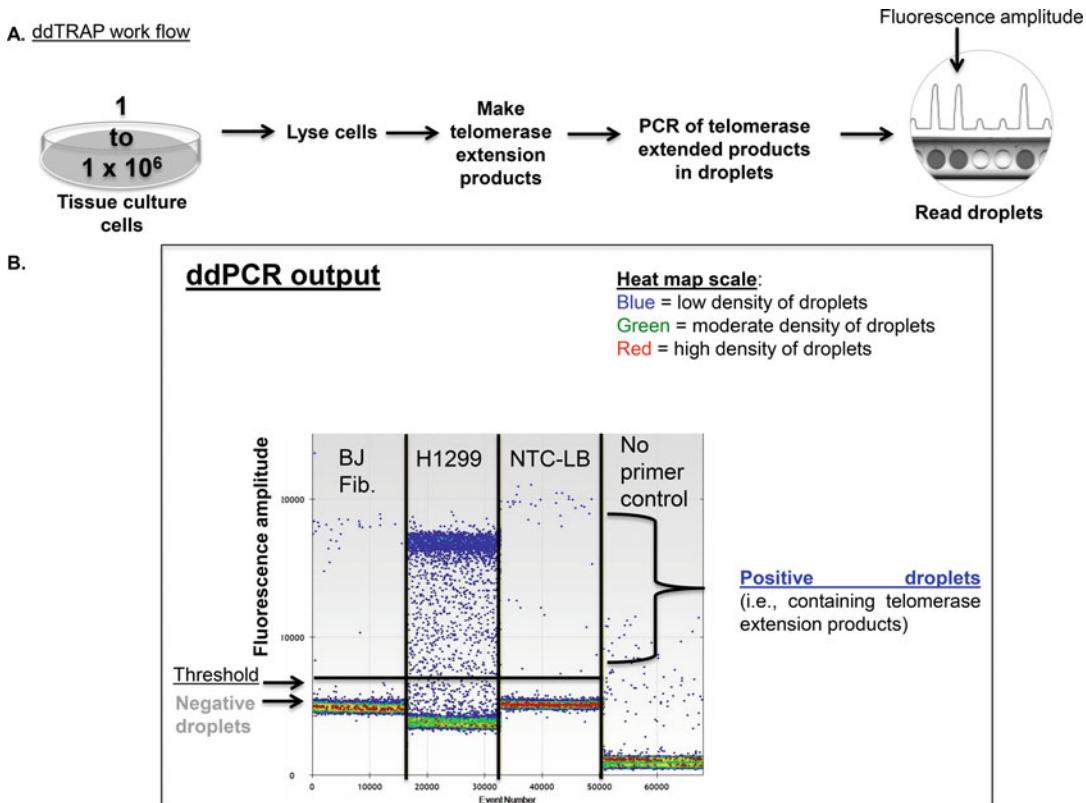
#### 4 Notes

1. Cell counts are critical to determine cell equivalents added to the extension reactions and for calculating the number of telomerase-extended products per cell.
2. Pellets are typically stable under ideal conditions for 1 year at –80 °C. If you freeze cells it is best to remove all extra liquid before placing in the –80 °C.
3. Following the extension reaction small white particulate matter may be observed in the extension reaction, to avoid carrying this material over to the ddPCR setup briefly centrifuge the PCR tubes and avoid pipetting from the bottom of the extension reactions. The amount of detergent from the lysate/extension can negatively impact droplet formation. We found that 0.008% of NP40 did not negatively impact droplet formation. We suggest testing concentrations of detergent higher than 0.008% on droplet formation prior to performing assays if higher concentration lysates are used.
4. Ten percent (10%) extra volume of each reagent should be used so that the final reaction volume is 22 µL to help prevent volume shortage when pipetting into the droplet generator cartridge.
5. You must make samples in intervals of 8.
6. The polymerase in the 2× ddPCR Evagreen® Super Mix is hot-start so all steps should be performed at RT.
7. Performing PCR set up on ice may increase solution viscosity (2× PCR master mix) which will affect droplet formation and is not recommended.

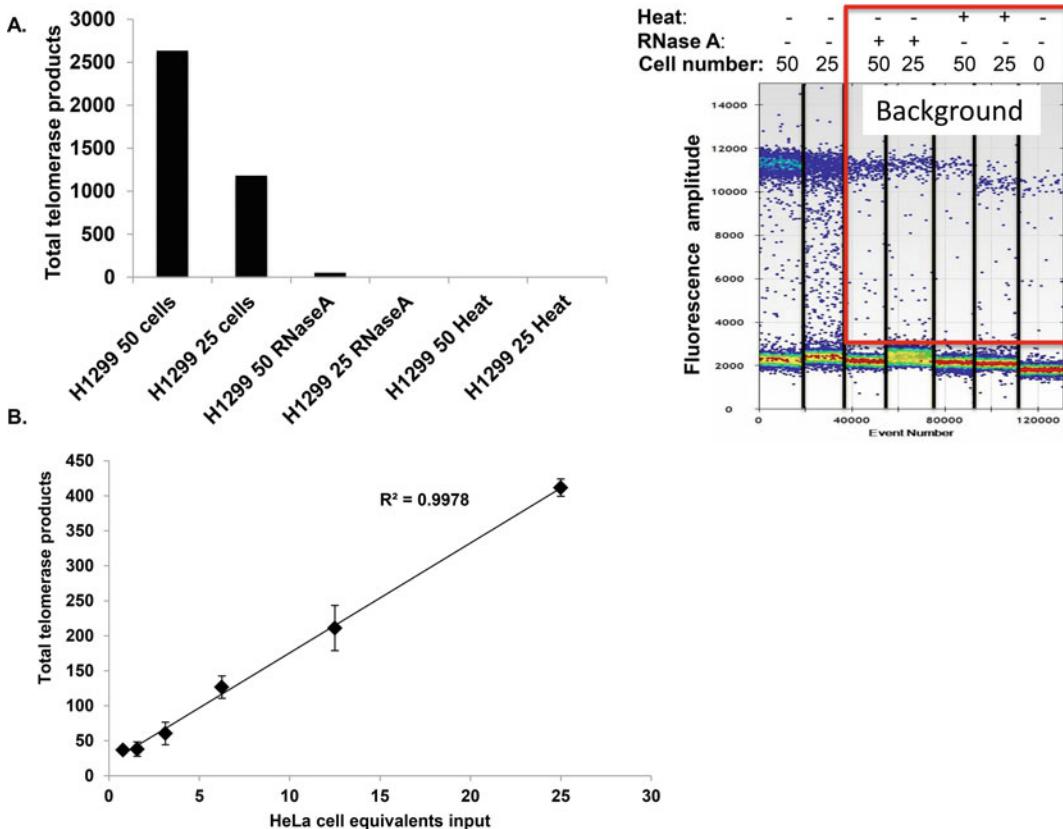
8. Order of loading the droplet generation cartridge essential. The oil is “heavy” which if loaded first causes the oil to flood the microfluidics and create a situation of poor droplet formation. Thus it is essential that you load your PCR sample first then the oil for proper droplet formation. After the PCR and oil are added to the cartridge wells attach a droplet generator cartridge gasket. The droplet generation machine works only with a full cartridge of 8 wells and a gasket properly attached to the cartridge. The machine stops when air (i.e., an empty well) is encountered or when a gasket is misplaced on the cartridge. If you only have 12 samples you must still set up 16 reactions, just four of the reactions will be blanks (mixing 11 µL 2× ddPCR supermix Evagreen® control with 11 µL of water per open well) or no template controls (NTC).
9. Transferring the droplets to the PCR plate is critical. Move slowly and consistently. Do not repeat pipet or completely aspirate the droplets from the pipet, as this will enhance droplet breaking/blending.
10. Evagreen® is read on the same channel as 6-fam.
11. Due to the sensitivity of the PCR-based TRAP assay, avoiding contamination is of utmost importance for assay success. We suggest setting up a dedicated set of pipettes and an area for ddTRAP assays only. Always use barrier tips. Do not use ART tips. Since telomerase is a reverse transcriptase with an RNA component, special precaution must be taken to ensure an RNase-free environment. All solutions should be made with DEPC-treated water, and benches, pipettes and labware decontaminated with an RNase inactivator such as RNaseZap®.
12. An alternative lysis buffer using CHAPS (0.5–1.0%) can be used in place of NP40 to generate cell extracts instead of whole cell lysates. Whole cell lysates give maximal activity detection, which is important for comparing telomerase-modulating compounds quantitatively [14, 15]. If desired, cell extracts can be made (lyse as above but spin down debris at 10,000 ×  $\varphi$  for 10 min at 4 °C), and supernatant removed for analysis. Cell extracts will underestimate the telomerase activity but may be useful in analyzing primary samples that may contain PCR inhibitors (such as primary tumor samples, [14]). CHAPS lysis buffer and cell extracts should be used when analyzing tissue samples. Mechanical homogenization (avoid heat treatment) or a mortar and pestle can be used to help disrupt the tissue. Centrifuge as above and determination of protein concentration as in note.
13. Determination of equivalent loading input for telomerase assays can be done in two ways: Cell counting or protein

concentration. We prefer cell counting when using whole cell lysates. We advise pelleting large numbers of cells so that the impact of loss of cells due to aspiration techniques is minimal. We typically use greater than 300,000 cell pellets and prefer  $1 \times 10^6$  cell pellets for optimal data. Protein concentration should be determined when cell extracts are used in ddTRAP. For ddTRAP 1–6 µg of protein is sufficient to detect telomerase activity from HeLa cells. Also, protein concentrations are needed for tissue samples but caution should be noted that telomerase positive cells will be comixed with telomerase silent stromal cells.

14. Cell pellets must be thoroughly lysed in NP40 buffer for a minimum of 40 min and a maximum of 1 h on ice. For pellets up to 1 million cells, typically 40 µL of NP40 lysis buffer is sufficient for lysis (1 million cells lysed in 40 µL of buffer results in 25,000 cell equivalents per microliter of lysate). We do not recommend using large volumes of lysis buffer to avoid dilutions; this may cause loss of telomerase activity or data that is not quantitative and repeatable. Do not lyse more than 45 min.
15. Lysates can also be stored at  $-80^{\circ}\text{C}$  but telomerase enzyme activity decreases overtime in the freezer, thus we recommend the use of fresh lysates and that lysates are aliquoted to avoid freeze–thaw cycles.
16. We have found that adding 1250 cell equivalents to a 50 µL extension reaction is the most reproducible in the ddTRAP assay (this results in a final extension reaction cell equivalent of 25 cells per microliter). Since we typically use 1 million cell count pellets, a 1:20 dilution in NP40 lysis buffer is necessary (2 µL of lysate in 18 µL of NP40 lysis buffer) and will generate a diluted lysate with 1250 cell equivalents per microliter. Once the lysates are diluted, 2 µL are added to the extension reaction mixture in thin wall PCR tubes/plates on ice. Tip—make sure that all lysates and dilutions are homogenous prior to pipetting. Important control reactions should be set up to confirm assay integrity. Controls such as treatment with RNase A to digest the RNA component of telomerase and or heat treatment ( $95^{\circ}\text{C}$  for 10 min) prior to extension can be performed to ensure specific detection of telomerase in the ddTRAP assay (Fig. 2). These controls are important for analysis of new tumor lines (i.e., lines with unknown telomere maintenance strategies) and for laboratories unfamiliar with the telomerase and TRAP assays (Fig. 3).
17. An extension time of 40 min is satisfactory for quantitative detection of telomerase (see Fig. 4), longer extension times (up to 2 h) can be used to detect maximal telomerase activity however longer extension times risk degradation of the enzyme

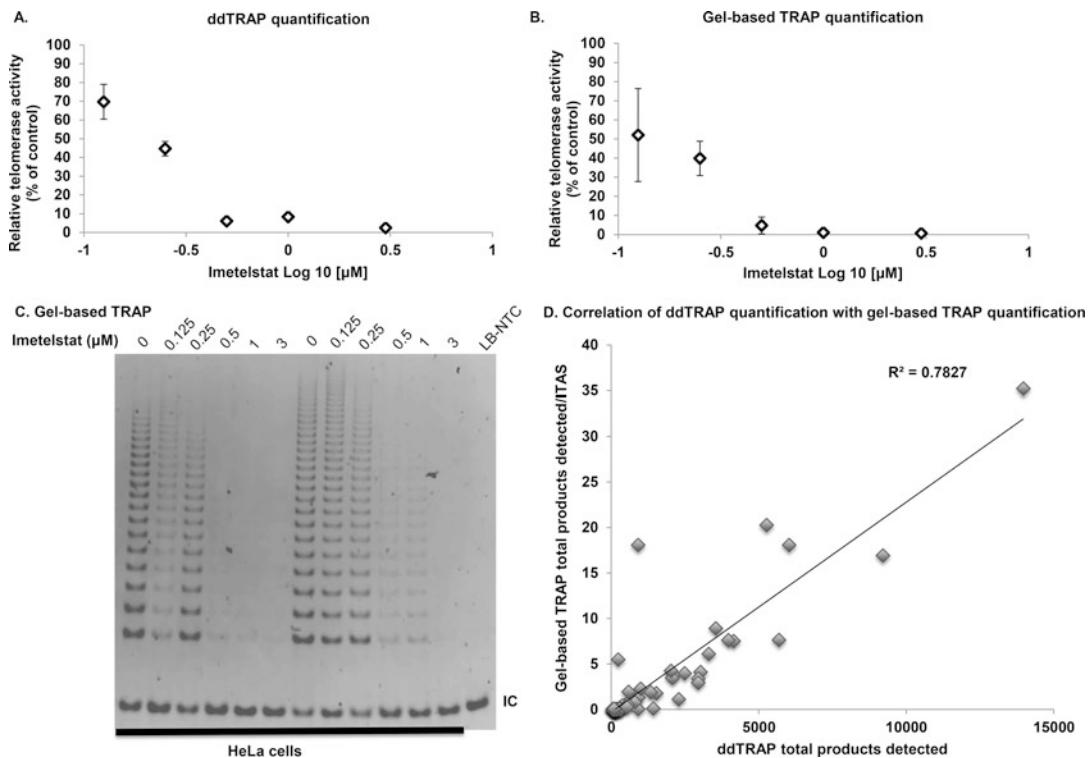


**Fig. 2** Workflow and optimization of droplet digital TRAP. (a) The ddTRAP workflow. Cells are lysed and then diluted to a concentration of 1250 cells per  $\mu\text{l}$ , telomerase extension products generated at a concentration of 25 cells/ $\mu\text{l}$ , then telomerase is heat inactivated and extension products dispersed into droplets. PCR thermocycling is done for 40 cycles and droplets analyzed for the presence or absence of fluorescence by the droplet reader (QX150/200 EvaGreen® compatible machine). (b) ddTRAP output showing BJ fibroblasts (input of 100 cell equivalents, telomerase negative), H1299 cells (input of 100 cell equivalents, telomerase positive), a lysis buffer only control, and a control with no primers and input of 100 cell equivalents of H1299 lysate to test for specificity of amplification. Only very low background signals are seen in these controls. Each well or sample of the ddPCR analyzes about 17,000 droplets. Event number at the bottom of the output represents the number of droplets counted in the wells overtime. Each dot on the ddPCR output represents a unique droplet that is either positive or negative for fluorescent signal. Fluorescence amplitude is a measure of the fluorescence detected for each droplet in the assay. Fluorescence amplitude is used to separate the positive and negative droplets. Since the droplets are detected with EvaGreen® double stranded DNA binding dye there will be inherent background fluorescence of DNA molecules not amplified during PCR. The heat map scale represents the density of droplets at given fluorescent amplitudes. NTC-LB = no-template control-lysis buffer ([13]; Figure is reproduced with permissions and slight modification from NAR and Oxford press - Ludlow, A.T., Robin, J.D., Sayed, M., Litterst, C.M., Shelton, D.N., Shay, J.W. and Wright, W.E. (2014) Quantitative telomerase enzyme activity determination using droplet digital PCR with single cell resolution. Nucleic Acids Res, 42, e104)



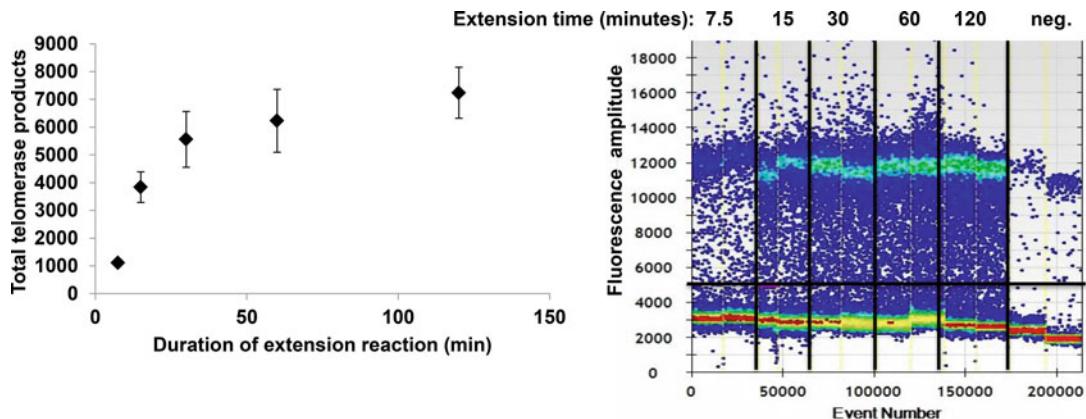
**Fig. 3** ddTRAP control reactions and linearity. **(a)** Heat and RNase A inactivation of telomerase resulted in virtually no ddTRAP signal. This permitted a small background correction to be included, and established the specific measurement of telomerase activity from lysates. Background signal is highlighted by the red box. **(b)** HeLa cell dilution series from 50 to 1 cell equivalents produced a linear relationship ( $R^2 = 0.99$ ) between input and detection of telomerase enzyme activity as indicated by total product generated in the ddPCR. Error bars are the standard deviation of the replicates. Although the extract was not made from a single cell, ddTRAP was able to reproducibly detect telomerase activity above background at the dilution equivalent to one cell input [13]

at 25 °C and thus have not been tested currently in the TRAP assay. Linear dilution analysis should be performed by inputting different amounts of cell equivalents into the extension reactions. Two types of linearity can be performed here: (1) Inputting different numbers of cell equivalents into the extension reaction (Fig. 2) or 2. Making an extension reaction and then diluting it prior to PCR to test the PCR linearity [13]. The extension products are single stranded DNA and should be used as soon as possible in the digital PCR reaction. We have found that storage of extension products at 4–12 °C overnight does not result in loss of detectable extension products (Fig. 5).



**Fig. 4** Comparison and correlation of ddTRAP to gel based TRAP. ddTRAP quantification is less variable than gel based TRAP and allowed accurate determination of the IC<sub>50</sub> of Imetelstat in HeLa (0.2 μM). HeLa cells were incubated with Imetelstat (0, 0.125, 0.25, 0.5, 1, and 3 μM) for 72 h. Cells were pelleted and triplicate extracts prepared from three separate tissue culture experiments (nine total extracts and extensions per dose). (a) ddTRAP quantification with 50 cell equivalents added to the PCR. (b) Gel quantification (representative gel image of two experiments in Fig. 3c). (c) Gel based TRAP was performed with 125 cell equivalents. Data are expressed as relative telomerase activity compared to control (untreated HeLa) and standard error of the mean. (d) Correlation analysis of gel-based TRAP to ddTRAP. The  $P < 0.0001$  positive relationship indicates that the methods are measuring the same phenomenon. IC = Internal competitive telomerase activity substrate (also known as ITAS) [13]

18. Setting digital PCR thresholds with “rainy” data. Since telomerase uses whole cell lysates and generates a “GC”-rich amplicon of various sizes there can be samples that produce “rain” or droplets of intermediate fluorescence intensity between the major populations of negative and positive droplets (see Fig. 4). This can cause problems in the analysis. For this reason controls are essential in the ddTRAP. We also always try to use 50 cell equivalents when possible. If samples seem to produce a lot of “rain” at this cell input we dilute the sample until better separation is achieved. Alternatively, the extension can be diluted and corrected for cell equivalents in the analysis.
19. We found that purchasing HPLC purified primers was necessary for the success of this assay, it will not work otherwise.



**Fig. 5** Dealing with “rainy” data. ddTRAP fluorescent droplet outputs are shown on the right, and quantified outputs graphed on the left. Lysates were incubated with TS substrate for various amounts of time prior to heat inactivation, using 100 cell equivalents from H1299 cells. Data are presented as means of background corrected total telomerase products generated  $\pm$  standard error of the mean. The horizontal black line represents the threshold used in these samples. It is essential when setting threshold for “rainy” data that controls are used as a guide and the same threshold is used for all samples [13]

We suggest diluting stock primers at a concentration of 100  $\mu$ M and making working concentration aliquots of no more than 100  $\mu$ L and avoiding freeze–thaw cycles.

20. Example calculation of cell equivalents: 1,000,000/40  $\mu$ L lysis buffer = 25,000 cells/ $\mu$ L lysate; diluted 1:20 = 1250 cells/ $\mu$ L lysate; 1  $\mu$ L of 1250 cell equivalents/50  $\mu$ L extension reaction = 25 cell equivalents/ $\mu$ L extension reaction \*2  $\mu$ L into 20  $\mu$ L ddPCR = 50 cell equivalents analyzed.

---

## Acknowledgments

We would like to acknowledge Viresh Patel and George Karlin-Neumann at Bio-Rad for materials support in the development of this technique and critical reading of the manuscript. We also acknowledge Oxford University Press for permissions to use Figure published in our article in Nucleic Acids Research. Ludlow, A.T., Robin, J.D., Sayed, M., Litterst, C.M., Shelton, D.N., Shay, J. W., and Wright, W.E. (2014) Quantitative telomerase enzyme activity determination using droplet digital PCR with single cell resolution. Nucleic Acids Res, 42, e104.

### Funding

National Institute of Health [AG01228]; National Cancer Institute [CA154805, P50CA70907, T32-CA124334-07, 5P30 CA142543-03]. Funding for open access charge: National Institute of Health Grant [AG01228]. This work was performed in

laboratories constructed with support from National Institutes of Health grant C06 RR30414.

*Conflict of interest statement*

Dawne N. Shelton is a employee of Bio-Rad Laboratories. Bio-Rad Laboratories provided materials and technical support for the development of this technique.

## References

1. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PL, Coviello GM, Wright WE, Weinrich SL, Shay JW (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science* 266:2011–2015
2. Shay JW, Wright WE (1996) Telomerase activity in human cancer. *Curr Opin Oncol* 8:66–71
3. Shay JW, Wright WE (1996) The reactivation of telomerase activity in cancer progression. *Trends Genet* 12:129–131
4. Shay JW, Wright WE (2011) Role of telomeres and telomerase in cancer. *Semin Cancer Biol* 21:349–353
5. Blackburn EH (1991) Telomeres. *Trends Biochem Sci* 16:378–381
6. Blackburn EH (2000) Telomeres and telomerase. *Keio J Med* 49:59–65
7. de Lange T (2009) How telomeres solve the end-protection problem. *Science* 326:948–952
8. de Lange T (2010) How shelterin solves the telomere end-protection problem. *Cold Spring Harb Symp Quant Biol* 75:167–177
9. Shay JW, Wright WE (2005) Senescence and immortalization: role of telomeres and telomerase. *Carcinogenesis* 26:867–874
10. Wright WE, Shay JW (1992) The two-stage mechanism controlling cellular senescence and immortalization. *Exp Gerontol* 27:383–389
11. Shay JW, Wright WE (2010) Telomeres and telomerase in normal and cancer stem cells. *FEBS Lett* 584:3819–3825
12. Nandakumar J, Cech TR (2013) Finding the end: recruitment of telomerase to telomeres. *Nat Rev Mol Cell Biol* 14:69–82
13. Ludlow AT, Robin JD, Sayed M, Litterst CM, Shelton DN, Shay JW, Wright WE (2014) Quantitative telomerase enzyme activity determination using droplet digital PCR with single cell resolution. *Nucleic Acids Res* 42:e104
14. Herbert BS, Hochreiter AE, Wright WE, Shay JW (2006) Nonradioactive detection of telomerase activity using the telomeric repeat amplification protocol. *Nat Protoc* 1:1583–1590
15. Norton JC, Holt SE, Wright WE, Shay JW (1998) Enhanced detection of human telomerase activity. *DNA Cell Biol* 17:217–219



# Chapter 30

## Highly Efficient and Reliable DNA Aptamer Selection Using the Partitioning Capabilities of ddPCR: The Hi-Fi SELEX Method

Aaron Ang, Eric Ouellet, Karen C. Cheung, and Charles Haynes

### Abstract

In addition to its growing use in detecting and quantifying genes and larger genomic events, the partitioning used in digital PCR can serve as a powerful tool for high-fidelity amplification of synthetic combinatorial libraries of single-stranded DNA. Sequence-diverse libraries of this type are used as a basis for selecting tight-binding aptamers against a specific target. Here we provide a detailed description of the Hi-Fi SELEX protocol for rapid and efficient DNA aptamer selection. As part of that methodology, we describe how Hi-Fi SELEX gains advantages over other aptamer selection methods in part through the use of the massive partitioning capability of digital PCR.

**Key words** Digital PCR, Aptamers, SELEX, Biological affinity reagents

---

### 1 Introduction

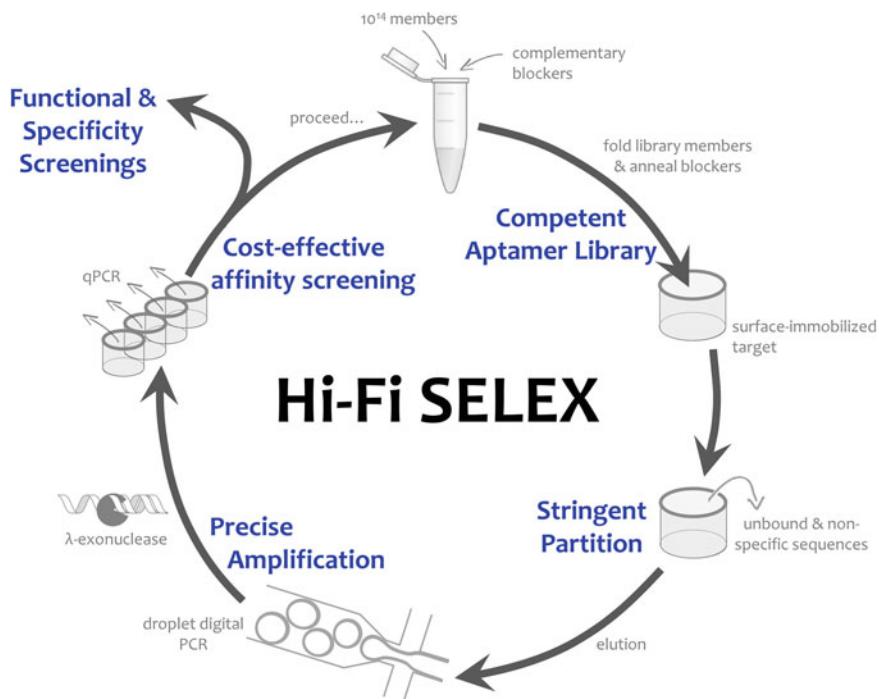
Biological reagents that recognize target molecules with both high affinity and specificity are routinely required in the life sciences and clinics, where in the latter case they often serve as capture, diagnostic, or therapeutic agents. Among the available classes of affinity reagents, antibodies currently are the gold standard. Antibodies, including monoclonal antibodies (mAbs), bind their antigen with high affinity; nM or tighter equilibrium dissociation constants ( $K_d$ s) are frequently reported [1]. Moreover, binding is generally highly selective, as demonstrated by the ability of mAbs to discriminate posttranslational modifications to proteins, as well as among more subtle protein isoforms [2, 3]. These exquisite binding properties can serve to maximize therapeutic mAb potency and minimize off-target effects, while the relatively large size of mAbs enables long circulation half-lives [4, 5].

However, antibodies are limited in certain important aspects. As large complex multisubunit proteins, they are sensitive to

environmental conditions, and can rapidly be inactivated under acidic conditions or at elevated temperatures. Though significant advances have been made toward their manufacture, mAbs remain relatively expensive to produce and purify at larger scales. Most mAbs are incapable of permeating cells efficiently, which effectively restricts their application to extracellular antigens. Moreover, though humanized-antibody technology has greatly reduced immune responses, therapeutic mAbs often do not escape immune surveillance completely, further challenging their long-term efficacy [6].

As a result, the development of mAb alternatives for use as research and diagnostic affinity reagents, as well as therapeutic agents, has gained considerable interest in recent years [7]. A number of simplified antibody forms including nanobodies,  $V_H$  and  $V_L$  antibody domains, and single-chain variable fragments have proven effective as mAb surrogates [8]. In addition, a particular class of nucleic acids, known as aptamers, has emerged as a potent option [9]. Each composed of a short single-stranded (ss) oligonucleotide, aptamers can be produced relatively inexpensively at large scale with high precision. The discovery of useful aptamers is likewise facilitated by the ability to easily synthesize large semicombinatorial libraries of ssDNA or ssRNA that can be subjected to in vitro Darwinian-type selection strategies to enrich and select sequences that exhibit high affinity and specificity for a target as a result of their unique folds. However, standard aptamer selection methods, which typically employ a recursive selection process known as systematic evolution of ligands by exponential enrichment (SELEX), are not sufficiently robust to ensure the timely and cost-effective discovery of aptamers suitable for further development [10–13].

We therefore recently reported on a new method, the high-fidelity SELEX (Hi-Fi SELEX) platform [14, 15], that greatly improves the speed and robustness of aptamer selection, in part by exploiting the partitioning capabilities of droplet digital PCR (ddPCR) to preserve library integrity during regeneration. The method requires either a commercial droplet generation system or the ability to create emulsions suitable for PCR. It does not require droplet-reading capabilities, making the method cost-effective since the remaining elements of Hi-Fi SELEX use consumables and equipment that are available in almost all standard molecular biology laboratories. Here, we describe the Hi-Fi SELEX method, for which the basic processing scheme is shown in Fig. 1, in sufficient detail to enable a user to reliably apply it to discover useful DNA aptamers against a target of interest. A number of “Notes” are also provided to give a better understanding of the methods we employ to maximize the speed and overall performance of the protocol.

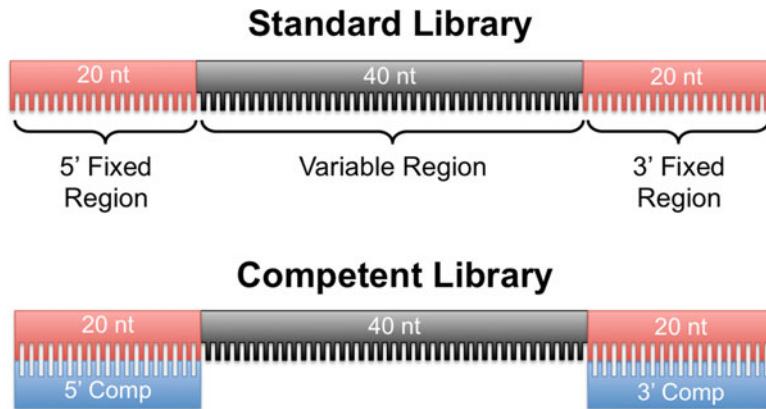


**Fig. 1** Schematic showing the sequence of operations comprising a single round of Hi-Fi SELEX. Adapted from [15] with journal permission

## 2 Materials

### 2.1 Oligonucleotides

The Hi-Fi SELEX protocol selects aptamers from a semicombinatorial library of members in which each member is composed of three oligonucleotides hybridized together. As shown in Fig. 2, these oligonucleotides include an 80-nt ssDNA “library” sequence, a 20-nt ssDNA DNA sequence (5'-Comp) complementary to the 5'-end of each 80-nt member of the library, and a 20-nt ssDNA sequence (3'-Comp) that is 5'-phosphorylated and complementary to the 3'-end of each library member. Each of the three required components (80-nt ssDNA library, 3'-Comp, 5'-Comp) is synthesized chemically by a reputable vendor (we use IDT Inc., Coralville, IA) capable of delivering it in the quantities (typically a few  $\mu$ moles), uniformities (ssDNA library and Comp sequences are independently processed by HPLC (strong anion exchange column) to eliminate truncated members and isolate a tight band of the desired molecular weight), and purities needed to satisfy the DNA quality and diversity requirements of the Hi-Fi SELEX process. Each component is then reconstituted in  $1 \times$  AF buffer to create the required stock solutions (*see Note 1*). A ssDNA library structure and corresponding set of 5'-Comp and 3'-Comp sequences appropriate for conducting Hi-Fi SELEX are provided below.



**Fig. 2** Basic structures of the semicombinatorial DNA libraries used in standard SELEX (standard library) and Hi-Fi SELEX (competent library): Hi-Fi SELEX libraries differ from standard SELEX ssDNA library through their use of novel fixed-region complements (blocking elements) to improve the functional diversity of the starting semicombinatorial library. Each member of the Hi-Fi SELEX ssDNA library is therefore composed of an 80 nt library sequence containing a 40-mer random core region ( $N_{40}$ ) flanked by a 5' universal 20-mer flanking sequence and a 3' universal 20-mer primer binding sequence, with each flanking sequence hybridized to its complement, which are hereafter denoted as 5'-Comp and 3'-Comp, respectively. By eliminating single-strand structures within the fixed regions, the competent Hi-Fi SELEX library isolates aptamer fold and function to within the variable core region of the library, while reducing artifacts that might compromise or eliminate the discovery of a tight-binding sequence within that region. Unblocked, the fixed-region sequences within a given selection library can interfere with aptamer fold and function through their potential to adopt stable secondary structures created through either (1) self-association, or (2) association with complementary nucleotides within the variable core region, the opposing fixed region, or both. These types of unwanted structures can occur either within an individual library member or between complementary regions of different members of the library [14], and we have previously shown that the net effect is to significantly reduce the total functional diversity of the library [15].

## 2.2 Hi-Fi SELEX Reagents and Consumables

1. Aptamer Folding (AF) buffer: 20 mM Tris-HCl (pH 7.4), 140 mM NaCl, 5 mM KCl, 1 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>.
2. ssDNA Library Stock Solution: 1 μmole purified 80-nt ssDNA in AF buffer (10 μl of a 100 μM stock required for complete (three rounds) Hi-Fi SELEX selection). The general structure of each library member is 5'-flanking sequence (20 nt)-N<sub>40</sub>-flanking sequence (20 nt)-3', with one suitable library sequence being library [5'-TCGCACATTCCGCTTCTACC-N<sub>40</sub>-CGTAAGTCCGTGTGCGAA-3'].
3. 5' Complementary Blocker (5'-Comp): for the library sequence shown above, 5'-GGTAGAACGGGAATGTGCGA-3' resuspended at 100 μM stock in AF buffer (10 μl each total needed for complete Hi-Fi SELEX run).
4. 3' Complementary Blocker (3'-Comp): for the library sequence shown above, 5'-TTCGCACACACGGACTTACG-3' resuspended at 100 μM stock in AF buffer (10 μl each total needed for complete Hi-Fi SELEX run).

5. Stringent Wash (SW) buffer: 20 mM Tris–HCl (pH 7.4), 1 M NaCl, 5 mM KCl, 1 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>, 0.005% Tween 20.
6. Denaturing Elution (DE) buffer: 50 mM NaOH (followed by neutralization with 50 mM HCl).
7. Tris-EDTA (TE) buffer: 10 mM Tris, 1 mM EDTA, adjusted to pH 8.0 with HCl.
8. Target Immobilization (TI) buffer: 100 µl of 100 mM sodium phosphate buffer (pH 7.5).
9. Surface Passivation (SP) buffer: AF buffer supplemented with 0.005% Tween 20.
10. λ-exonuclease; lambda exonuclease (New England Biolabs).
11. Nunc Immobilizer Amino Plates (ThermoFisher Scientific).
12. Thermomixer fitted with a plate adapter (e.g., Eppendorf Thermomixer C).
13. NanoSep® Omega 10K MWCO centrifugal filter unit (Pall Corporation).
14. Standard PCR thermal cycler (e.g., Bio-Rad C1000 or Eppendorf Mastercycler Thermal Cycler).

### **2.3 ddPCR and qPCR Reagents**

1. Forward Primer (FP): 5'-TCGCACATTCCGCTTCTACC-3' resuspended in AF at 100 µM stock concentration (for library sequence shown).
2. Phosphorylated Reverse Primer (RP): 5'-p-TTCGCACA-CACGGACTTACG-3' resuspended in AF at 100 µM stock concentration (library sequence shown).
3. SYBR Green qPCR master mix: 2× iQ SYBR Supermix (Bio-Rad).
4. ddPCR master mix: ddPCR™ Supermix for probes without dUTP (Bio-Rad).
5. Thermal Cycler capable of real-time detection (e.g., Bio-Rad CFX96).
6. DG8™ Droplet-Generation Cartridge (Bio-Rad).
7. Droplet Digital™ PCR (ddPCR™) Droplet Generator (Bio-Rad QX100™ or QX200™).

---

## **3 Methods: The Hi-Fi SELEX Protocol**

Prior to the initiation of any aptamer selection work, it is essential that proper steps be taken to prevent cross-contamination of material recovered between selection rounds. Regular decontamination of all work surfaces and regular changes of gloves during the

process is necessary to prevent such occurrences. Moreover, stock reagents used for PCR amplification must be kept separate from the initial library as well as from any recovered pools of library members during selection. Aliquots of each stock reagent should be made in designated nucleic acid template-free areas. If aptamer selection is to become routine in the lab, a dedicated separate environment for conducting Hi-Fi SELEX must be created, with regular maintenance and decontamination of all surfaces and instruments prior to and at the completion of each round of selection.

### 3.1 Library Design and Synthesis

In its standard format [10–12], SELEX enriches a subset of short single-stranded RNA or DNA sequences from a synthetic library composed of a semicombinatorial population whose total diversity is similar to the body's own antibody repertoire. The process is typically performed *in vitro*, creating the capacity to modulate the nature and stringency of the screening conditions to more efficiently select library members showing high affinity or specificity to a target.

SELEX can operate on either a ssDNA or ssRNA library, with the choice typically made by carefully examining the intended application of the aptamer. RNA-based aptamers have a more flexible backbone that allows them to adopt a wider range of secondary and tertiary structures. Though built on a less flexible backbone, ssDNA aptamers have higher chemical stability, and screening of DNA libraries requires fewer processing steps. In this chapter, the technological advances offered by Hi-Fi SELEX, including its use of the partitioning step of ddPCR, are specifically applied to ssDNA aptamer selections.

DNA libraries used for SELEX typically contain  $\sim 10^{14}$  unique members of equivalent length, each composed of a random oligonucleotide sequence within a variable core region that is flanked by universal fixed sequences at the 5' and 3' ends. Library diversity is largely encoded in the variable core region. This region can be created utilizing different randomization strategies and nucleotide chemistries. Most often, chemical syntheses that equally weight the frequency of each naturally occurring nucleotide are employed, but variable core regions composed of partially randomized sequences [16], genomic DNA inserts [17, 18], and various chemically modified nucleotides [19–21] have also been used with success.

Hi-Fi SELEX typically employs a variable core region that is 40 nt in length (Fig. 2) so as to generate reasonable diversity while keeping the mass of the starting library suitable for the screening process. However, smaller or longer randomizations can be used. The sequences of the universal fixed regions (typically 20 nt each) flanking the 5' and 3' ends of each variable core sequence are engineered to achieve high-fidelity amplification of retained library members by eliminating, or at least minimizing, self-association and primer-dimer pairing reactions that can promote formation of

unwanted by-products during PCR. In our lab, this is typically achieved by designing both the flanking sequences and the associated forward (FP) and reverse (RP) primers using Primer3 software ([http://biotools.umassmed.edu/bioapps/primer3\\_www.cgi](http://biotools.umassmed.edu/bioapps/primer3_www.cgi)), and then scoring them for self-complementary and primer-dimer formation using the OligoAnalyzer tool (<https://www.exiqon.com/ls/Pages/ExiqonOligoOptimizerTool.aspx>) of Exiqon, Inc. Several pairs of 20-mer flanking sequences designed using the same or a similar strategy have been reported and successfully used [22–24].

The random core region of the 80-mer ssDNA library is generally created combinatorially by dosing the required nucleotides at an A:C:G:T molar ratio of 3:3:2:2.4 in order to achieve equal probability incorporation of each nucleotide within the variable core region [25]. The 5'-phosphorylated 3'-Comp sequence is used as the RP for PCR amplification.

### **3.2 Target Immobilization**

Levels of nonspecifically retained library members are known to dictate the degree of enrichment of useful aptamers achieved in a given round of selection [26]. Practitioners of SELEX have attempted to quantify the overall quality of each round of selection through a parameter *PE* known as the “partition efficiency”:

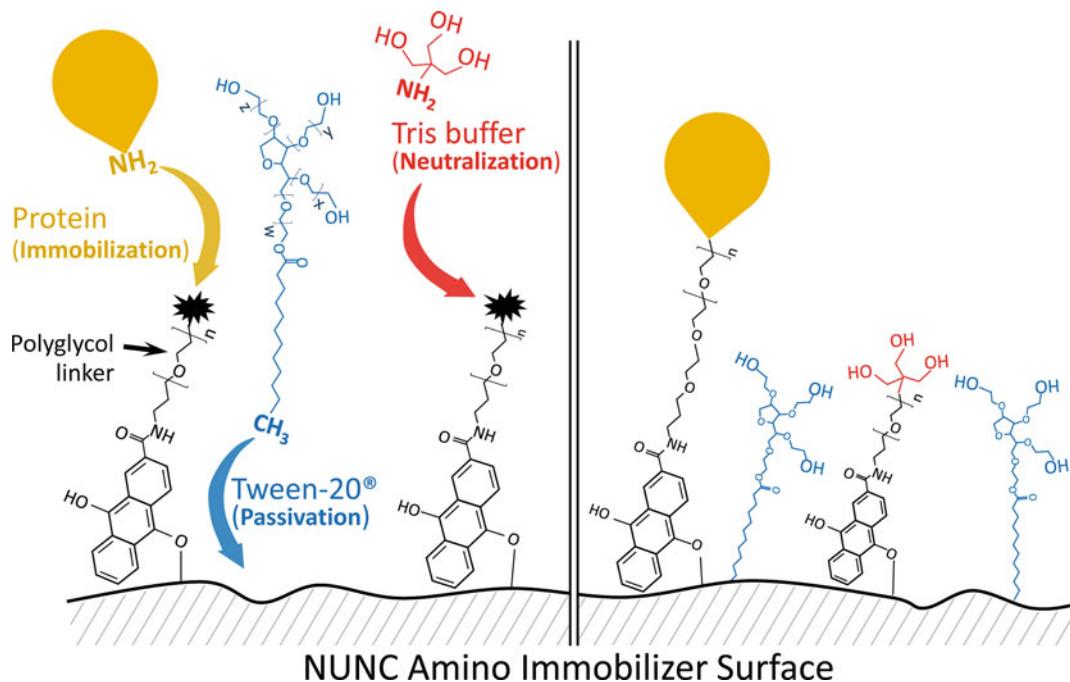
$$PE = \frac{[A]}{[AP]} \quad (1)$$

where

$$[A] = \sum_{i=1}^n [A_i] \quad (2)$$

Here,  $[A]$  is the total molar concentration of all unbound library members, given by the sum of the concentration  $[A_i]$  of each unbound library sequence  $i$ ; likewise  $[AP]$  is the molar concentration of all bound library members recovered in the fraction eluted in DE buffer. In SELEX, per round *PE* values generally range from 10 to 1000, due in part to nonspecific retention of library members. Reducing the starting population of  $\sim 10^{14}$  members to a manageable number of candidate aptamers, say  $\sim 10^4$ , then requires a relatively large number of successful selection rounds (generally identified by an improvement in the mean  $K_d$  of the retained pool).

In contrast, Hi-Fi SELEX generates per round *PE* values greater than  $10^5$  (usually close to  $10^6$ ) by displaying the target on a tailored substrate and in conditioned solvents that together greatly inhibit nonspecific retention. Nunc Immobilizer plates, which present on the surface of each well a layer of end-grafted hyperbranched polyglycol chains at densities near or above those needed for soft brush formation, display good passivation against



**Fig. 3** Schematic of chemistries used to immobilize target protein, neutralize reactive groups, and passivate nonspecific binding sites on the surface of a Nunc Immobilizer Amino plate

nonspecific adsorption of proteins [27] or oligonucleotides [28, 29], even in cases where the zeta potentials of the underlying base surface and sorbate are opposite in sign [27, 30]. This capability is integral to the Target Immobilization procedure itemized below:

1. Immobilize the target protein (50–100 nM target in TI buffer) to the Nunc Immobilizer Amino C8 strips by incubating the filled well overnight at 4 °C (*see Fig. 3 and Notes 2 and 3 for information on the coupling chemistry*).
2. Wash the wells three times with 300  $\mu\text{l}$  of SP buffer (*see Note 4*).
3. Fill the wells once more with 300  $\mu\text{l}$  of SP buffer.
4. Incubate for 1 h under gentle agitation (we set a thermomixer fitted with a plate adapter at 500 rpm) at room temperature to complete the neutralization reaction.
5. Aspirate the SP buffer out of each well at the end of the neutralization/passivation process.
6. Perform a final wash ( $3 \times$  of 300  $\mu\text{l}$  of SP buffer) before sealing the wells with optical film.
7. Store the processed and sealed C8 strip at 4 °C until further use.

### 3.3 Partitioning and Retained Fraction Recovery

Prior to Hi-Fi SELEX screening, the library is subjected to a thermal denaturation–renaturation cycle designed to promote annealing of the 5'-Comp and 3'-Comp strands to each 80-mer library member, as well as folding of each member into its thermodynamically favourable conformation at screening conditions [31] (*see Note 5*). Note that the success of Hi-Fi SELEX in reliably and efficiently discovering useful aptamers requires strict adherence to this protocol during each selection round. The combined use of an aptamer-to-target ratio of 100:1 to 1000:1 (*see Note 6*), conditioned solvents, and high-salt wash steps creates the stringent partitioning conditions (*see Note 7*) required to reduce the sequence diversity 10<sup>5</sup>–10<sup>6</sup> orders of magnitude in each of the first few selection rounds. Approximately 10<sup>8</sup> unique library sequences are therefore retained and recovered in the first partitioning step—a number optimal for the remaining steps of the selection cycle.

The library partitioning and retained fraction recovery protocol is as follow:

1. Mix 1 nmole (~10<sup>14</sup> sequences) of the ssDNA library from the prepared stock solution with equimolar amounts of 5'-Comp and 3'-Comp in 100 µl of SP buffer.
2. Heat that mixture to 95 °C for 5 min (denaturation) before slowly cooling it down to 25 °C at a rate of 0.5 °C min<sup>-1</sup> (renaturation) in a standard thermal cycler.
3. Add the resulting 10 µM of competent library to a neutralized and passivated Nunc Immobilizer well displaying the target.
4. Equilibrate that system for 1 h at 25 °C under gentle agitation (500 rpm) in a thermomixer equipped with a plate reader (*see Note 6*).
5. Remove unbound and weakly bound members by washing three times with 300 µl of SP buffer.
6. Apply a second more stringent pair of 3× washes with 300 µl of SW buffer. These washing steps do not require agitation (*see Notes 7 and 8*).
7. Elute the retained pool of specifically bound library members from the Nunc wells by incubating with 100 µl of DE buffer for 10 min at 70 °C and an agitation rate of 600 rpm.
8. After agitation, neutralize the pH by adding 100 µl of 50 mM HCl.
9. Repeat **steps 7 and 8**.
10. Combine the eluted aptamer pools.
11. Desalt that pool by exchange into AF buffer and concentrating to 25 µl (final) using a centrifugal filter unit operating at 14,000 × g for 15 min (*see Note 9*). Begin that centrifugation

without addition of AF buffer so as to remove as much DE buffer as possible. Then add AF buffer and begin buffer exchange.

### **3.4 Retained Pool Amplification by ddPCR**

Following each selection round, the polymerase chain reaction (PCR) is generally used to amplify the pool of retained library sequences. In SELEX, this step is conducted as a conventional bulk PCR reaction using a standard thermal cycler and universal forward and reverse primers targeting the fixed flanking sequences of each library member.

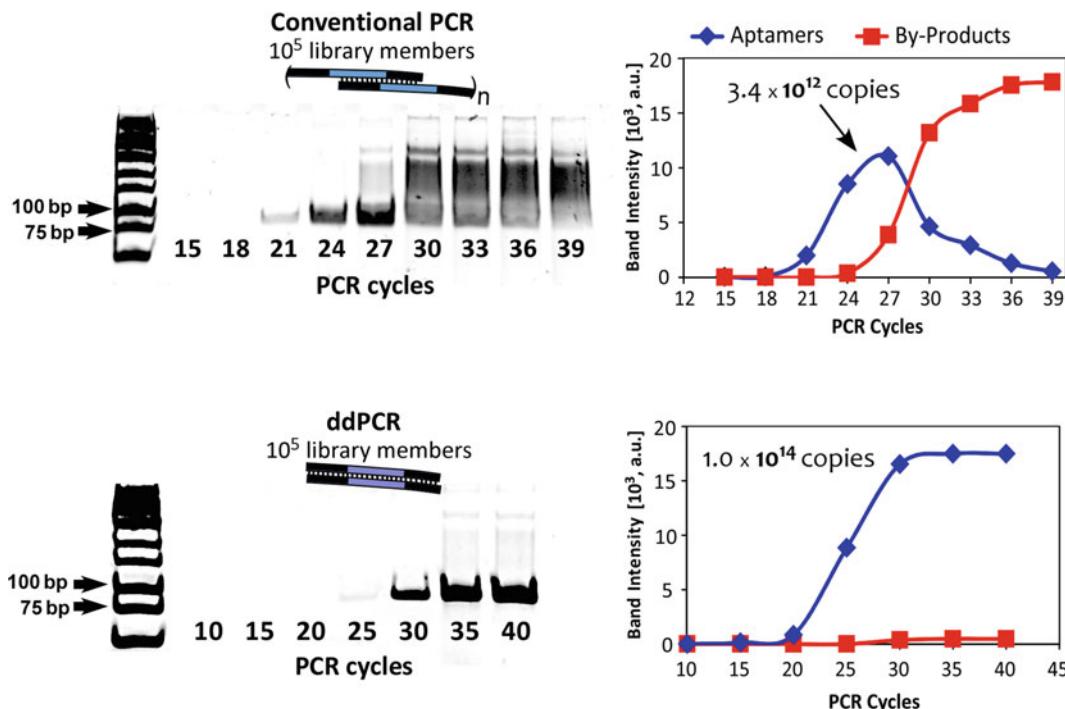
This approach tends to be problematic, as illustrated in Fig. 4 for a relatively simple pool of  $10^5$  unique library members. Though the desired 80-bp dsDNA amplicon is created, a maximum in its total abundance is typically reached after a limited number of cycles. Beyond that cycle, various artifacts, including formation of 80-mer heteroduplexes hybridized together through only their common flanking sequences, oligonucleotide stretches within the universal primer regions mispriming certain variable core region sequences, and improperly extended products acting as spurious primers on heterologous sequences, promote conversion of the library to increasingly aberrant high molecular weight (HMW) by-products.

To avoid these complications, Hi-Fi SELEX replaces traditional PCR with the partitioning capabilities of ddPCR. When as few as  $10^5$  retained library members are recovered in a given Hi-Fi SELEX round and then amplified by the ddPCR-based protocol described, an 80-bp amplicon concentration of greater than  $1 \mu\text{M}$  is generally realized in the final  $\sim 25 \mu\text{l}$  sample. Moreover, all amplicons produced are in their fully complementary (homoduplexed) dsDNA state. As a result, a regenerated ssDNA library can be created from as little as  $10^5$  retained members in quantities sufficient to not only proceed to the next selection round, but also to determine the mean binding affinity of the enriched pool after each selection round, providing a metric of how the overall selection is proceeding.

The amplification protocol used in Hi-Fi SELEX is conducted in two steps as follows:

*Determine the concentration of the retained library  $C_{\text{Library}}$ :*

1. Dilute  $0.5 \mu\text{l}$  of the desalted concentrated pool of retained 80-nt library members in  $9.5 \mu\text{l}$  nanopure water.
2. To a qPCR well, add  $5 \mu\text{l}$  of the solution created in **step 1**, along with SYBR green mastermix (to  $1 \times$  final), FP and RP ( $250 \text{ nM}$  final each). Top up the final volume in the well to  $20 \mu\text{l}$  by adding nanopure water.
3. Begin cycling with an initial activation step at  $95^\circ\text{C}$  for 3 min followed by 39 cycles of amplification, each composed of denaturation at  $95^\circ\text{C}$  for 30 s and annealing/extension at  $60^\circ\text{C}$  for



**Fig. 4** Comparison of amplification of a retained pool ( $10^5$ ) of library members by conventional PCR (standard SELEX; upper panel) and by ddPCR (Hi-Fi SELEX; lower panel): Although the desired 80-bp dsDNA amplicon is created in conventional PCR, a maximum in its total abundance is typically reached after a limited number of cycles. Beyond that cycle, various artifacts, including formation of 80-mer heteroduplexes hybridized together through only their common flanking sequences, oligonucleotide stretches within the universal primer regions mispriming certain variable core region sequences, and improperly extended products acting as spurious primers on heterologous sequences, promote conversion of the library to increasingly aberrant high molecular weight (HMW) by-products. A small-scale “pilot” PCR reaction [32, 33] is therefore generally performed to determine the maximum number of PCR cycles that can be conducted before accumulating unacceptable amounts of by-products, which are known to adversely affect selection and must therefore be removed by gel electrophoresis or other means [34]. That pilot reaction typically shows that standard PCR amplification of the retained pool must be stopped at ca. 22–25 cycles since HMW by-products generally start accumulating when the amplicon concentration reaches ca. 20–50 nM. Termination of amplification at this relatively low cycle number generally yields  $\sim 10^{10}$ – $10^{12}$  80-bp amplicons, or  $\leq 1\%$  of that needed to initiate the next selection round. As a result, in SELEX, the PCR step must be multiplexed across 100 or more parallel reactions, each amplifying between  $\sim 10^5$ – $10^6$  library members to create  $10^{11}$ – $10^{12}$  amplicons per well. The products of the parallel reactions are then pooled and concentrated to reach the concentration (i.e.,  $10^{14}$  amplicons in 100  $\mu$ L) required for downstream processing and the next round of SELEX. The use of emulsions in ddPCR to isolate and amplify single templates by PCR is well established, and it is known that the resulting partitioning of single templates into individual droplets reduces formation of unwanted by-products when coamplifying mixtures of templates (e.g., multiple genes) [35, 36]. Spurious priming events are greatly reduced within each droplet, in part because competition between different templates and biases resulting from differences in amplification efficiencies are avoided [37]. Moreover, post amplification, the emulsions can be broken to recover the full set of amplicons in an aqueous phase suitable for downstream processing. In Hi-Fi SELEX, the pool ( $\sim 10^8$ ) of competent 80-nt ssDNA library members retained after a selection round is therefore partitioned among a similar number of nL-sized droplets. ddPCR partitions ca. 20,000 droplets per well, which means 100 wells

30 s. Set the heating and cooling rates at  $2.5\text{ }^{\circ}\text{C s}^{-1}$  for even heat distribution in the well.

4. Determine  $C_{\text{Library}}$  by comparing the recorded quantitation cycle  $C_q$  to corresponding  $C_q$  data from a standard curve derived from a dilution series of the initial 100  $\mu\text{M}$  library stock with sequence diversity from  $10^8$  down to  $10^2$  unique library members.
5. From  $C_{\text{Library}}$  and the retained library volume ( $\sim 25\text{ }\mu\text{l}$ ) the number of ddPCR wells required to conduct the retained library amplification is computed by assuming 17,000 readable drops having a CPD of 50 are created per ddPCR well. If, for example, the retained library contained  $10^8$  members (typical for 1st and 2nd rounds of Hi-Fi SELEX), 120 wells are required ( $120\text{ wells} \times 17000\text{ droplets} \times 50\text{ copies/droplet} = 10^8$ ).

*Droplet based amplification of the retained pool (see Note 10)*

6. From the remaining 24–25  $\mu\text{l}$  of desalting retained library, prepare an appropriate volume ( $= 20\text{ }\mu\text{l} \times \# \text{ of required wells}$  ( $= \sim 2\text{--}2.4\text{ ml}$ )) of ddPCR sample mixture by adding ddPCR master mix (1× final), 900 nM each of FP and phosphorylated RP (final) and nanopure water.
7. Load a 20  $\mu\text{l}$  aliquot of this sample mixture into each well of a droplet-generation cartridge.
8. Add 70  $\mu\text{l}$  of fluorinated oil to each of the corresponding oil wells of the cartridge.
9. Insert and process the cartridge in the droplet generator.
10. Transfer the stable emulsion (containing ca. 17,000 readable droplets) formed in each well into a well of a standard 96-well PCR plate for thermal cycling.
11. Repeat steps 7 to 10 until the sample is fully processed.
12. Begin cycling with an initial activation step at  $95\text{ }^{\circ}\text{C}$  for 5 min, followed by 35 amplification cycles (note that this is fewer cycles than typically used for ddPCR quantitation), each composed of denaturation at  $95\text{ }^{\circ}\text{C}$  for 30 s then annealing/extension at  $64\text{ }^{\circ}\text{C}$  for 30 s. Set the heating and cooling rates at  $2.5\text{ }^{\circ}\text{C s}^{-1}$  to ensure even heating of all droplets in each well.

---

**Fig. 4** (continued) are required to accommodate 50 templates into each droplet (CPD = 50). As a result of the low sequence heterogeneity per droplet, minimal HMW by-products formation is observed over 40 or more ddPCR cycles, permitting high-fidelity end-point amplification of all retained library members into more than  $10^{14}$  total copies of the desired 80-bp dsDNA amplicon products. Adapted from [15] with journal permission

13. Immediately after amplification, pool all of the wells and spin at  $5000 \times g$  to separate the reacted droplets from the continuous oil (bottom) phase.
14. Discard the continuous oil phase.
15. Recover the double-stranded DNA by subjecting the droplet phase to a freeze ( $-80^{\circ}\text{C}$  for 15 min)/thaw cycle.
16. Immediately spin down the frozen droplets at  $14,000 \times g$  for 5 min to create sufficient force to burst them.
17. Repeat **steps 15** and **16** two more times to generate and recover a clear aqueous (top) phase containing the soluble amplified material. Set aside 5  $\mu\text{l}$  of the clarified aqueous phase.
18. Process the remaining clarified aqueous phase in a centrifugal filter unit to concentrate the amplified library to  $\sim 25 \mu\text{l}$ .
19. Analyze the 5  $\mu\text{l}$  aliquot from **step 18** on a 1.5% agarose gel alongside a properly sized molecular weight ladder to confirm proper amplification of the library.

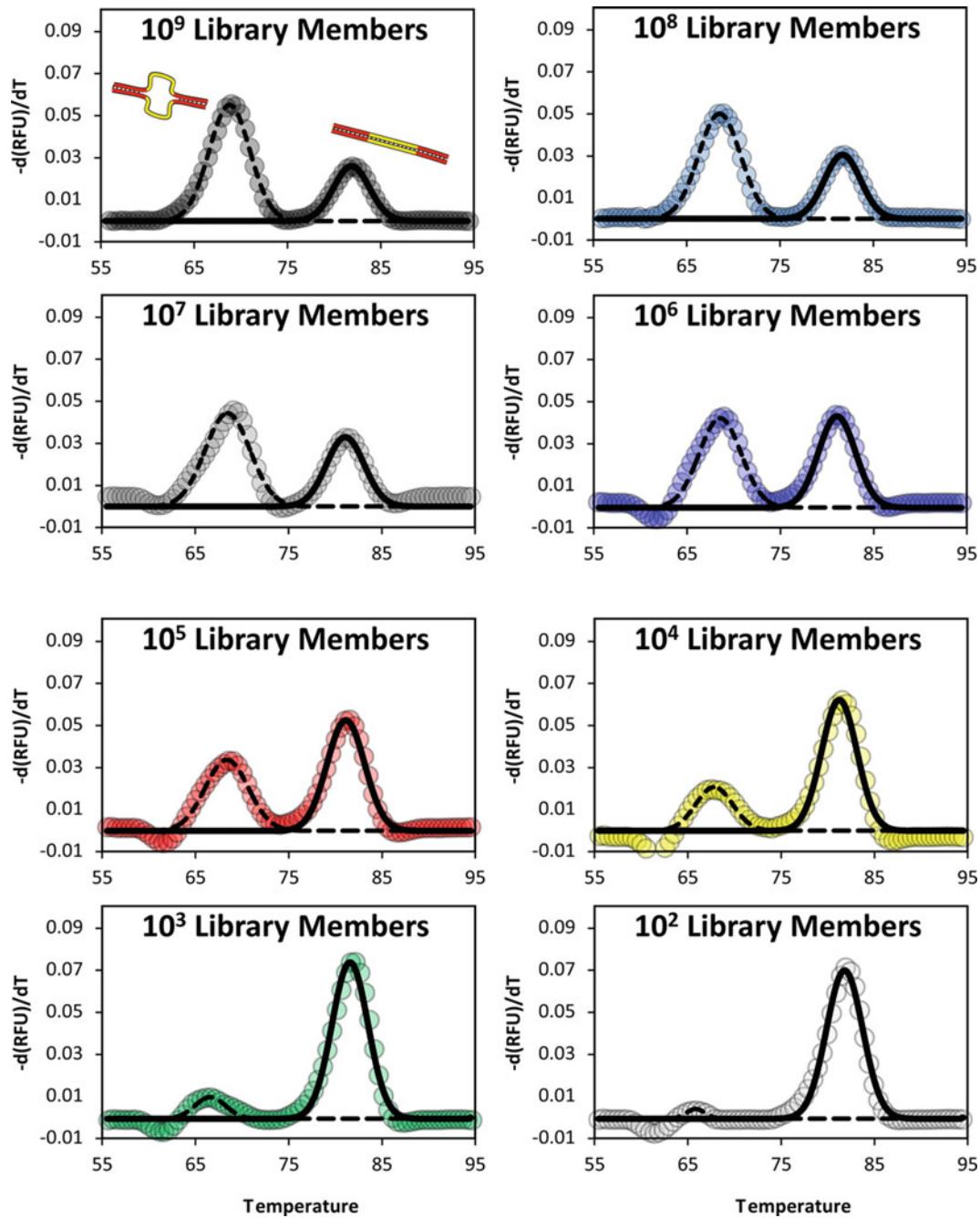
### **3.5 Measuring the Sequence Diversity of the Amplified Retained Pool**

In general, a proper balancing of retained library sequence diversity and mean binding affinity must be maintained across selection rounds for a Hi-Fi SELEX screening to prove successful in discovering a useful set of candidate aptamers. Maintaining that balance requires methods to quantify both the mean  $K_d$  and the total sequence diversity of the retained pool after each selection round. Retained library diversities can be determined using next-generation sequencing, but that approach is neither fast nor inexpensive [38, 39].

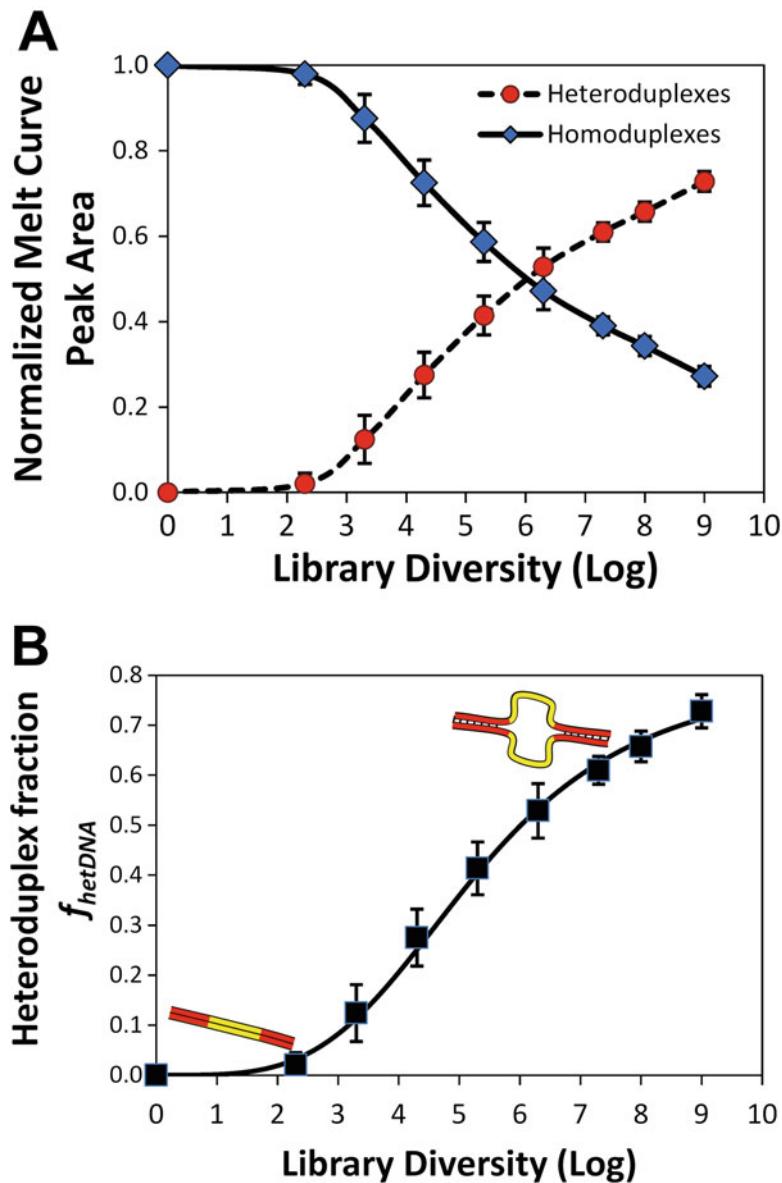
In Hi-Fi SELEX, the sequence diversity of a retained pool of library members is estimated using a novel qPCR assay that is simple and inexpensive, requiring equipment common to all molecular biology labs. The area of the melting peak for each population is recorded and the two peak areas are used to compute the fraction  $f_{\text{hetDNA}}$  of amplicons in the heteroduplexed state:

$$f_{\text{hetDNA}} = \frac{A_{67^{\circ}\text{C}}}{(A_{67^{\circ}\text{C}} + A_{81^{\circ}\text{C}})} \quad (3)$$

This melt analysis is repeated for serial reductions in the sequence diversity of the dsDNA-amplicon representations of the starting library (see above) to generate data as shown in Fig. 5. The corresponding peak areas for the set of serial reductions are used to create a standard curve relating  $f_{\text{hetDNA}}$  to the known sequence diversity (Fig. 6). For those standards, the area of the  $T_m = 67^{\circ}\text{C}$  melting peak decreases with decreasing sequence diversity, reflecting the higher probability of forming fully complementary homoduplexes (which melt at  $T_m = 81^{\circ}\text{C}$ ) in such systems. The value of  $f_{\text{hetDNA}}$  measured for a retained pool thereby permits estimation of its sequence diversity (see Note 11).



**Fig. 5** Fluorescence (SYBR green) based melt analysis for serial reductions in the sequence diversity of dsDNA-amplicon representations of an 80-nt Hi-Fi SELEX library: The double stranded amplicons of retained library members are homogeneous in terms of their two flanking sequences, while presenting a highly diverse ensemble of variable core-region sequences. Denaturation of the amplified library followed by cooling to 55 °C therefore results in two distinct dsDNA populations: fully homoduplexed amplicons characterized by a Gaussian melting envelope centered at a relatively high melting temperature ( $T_m \sim 81$  °C), and heteroduplexes exhibiting only partial complementarity (typically through only their common flanking sequences). The heteroduplexed pool of amplicons collectively exhibits a Gaussian melting peak characterized by a much lower  $T_m$  ( $\sim 67$  °C).



**Fig. 6** Standard curve for qPCR-based sequence diversity determination: normalized melt peak areas (**a**) and  $f_{hetDNA}$  values (**b**) as a function library diversity

The protocol used to measure sequence diversity is as follows:

1. Amplify the remaining 5  $\mu$ l volume of the diluted working aliquot prepared in **step 1** from Section 3.4 by qPCR using SYBR Green master mix (1× final) and qPCR thermal cycling conditions composed of initial activation at 95 °C for 5 min followed by 13–15 amplification cycles: denaturation at 95 °C for 30 s, annealing at 64 °C for 30 s and extension at 72 °C for 30 s, with the heating and cooling ramp rate set at 3 °C  $s^{-1}$  (*see Note 12*).

2. Following cycling, cool the SYBR dye-bearing amplicons to 55 °C.
3. Then, melt by heating from 55 °C to 95 °C in 0.5 °C increments in a real-time PCR thermal cycler and record the required  $A_{67^{\circ}\text{C}}$  and  $A_{81^{\circ}\text{C}}$  values.
4. Determine the sequence diversity from the standard curve relating  $f_{\text{hetDNA}}$  to the known sequence diversity (e.g., Fig. 6).

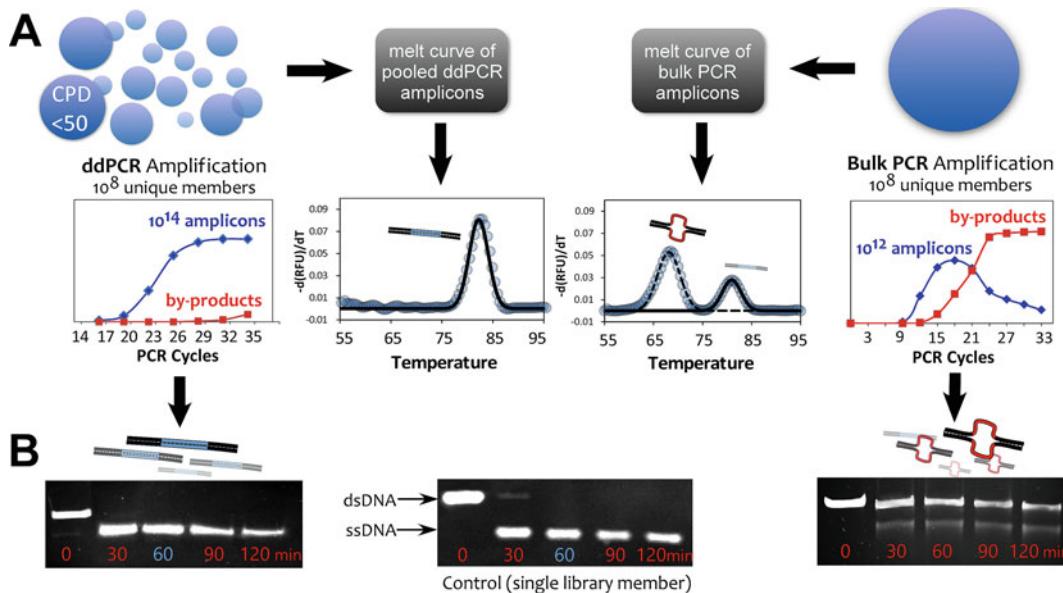
### 3.6 Enzymatic Regeneration of the Single-Stranded Library

The ddPCR amplification step generates 80-bp dsDNA products from which the sense strands must be recovered in order to continue the Hi-Fi SELEX process. A number of purification methods have been established for stoichiometric removal of the antisense strand from the sense (aptamer library) strand of amplicons, including alkaline denaturation followed by streptavidin capture of biotinylated antisense strands (generated by chemically modifying the RP) [40] or electrophoretic separation of poly-T-labeled antisense strands [41]. The relative performance of these various methods has been studied by Civit et al. [42] and others [43, 44], who have collectively shown that enzymatic digestion of 5'-phosphorylated ( $\text{PO}_4$ ) antisense strands with  $\lambda$ -exonuclease generally works best.

However,  $\lambda$ -exonuclease hydrolysis activity is appreciably lower on ssDNA than on complementary dsDNA [45]. As a result, we have shown that enzymatic regeneration of a ssDNA library stalls when acting on heteroduplexed amplicons produced in the bulk PCR amplification step employed in standard SELEX [15]. A partially regenerated library composed of a mixture of desired 80-nt ssDNA and partially processed dsDNA material is therefore created (Fig. 7, right half), compromising the next round of selection.

In Hi-Fi SELEX, in addition to mitigating formation of unwanted HMW products, the ddPCR based amplification of retained members described above provides a means to eliminate formation of heteroduplexes during amplification and thereby yield a pool of fully homoduplexed amplicons (as evidenced by melt analyses of those pools (Fig. 7)). Complete stoichiometric regeneration of the required 80-nt ssDNA library by  $\lambda$ -exonuclease processing can then be achieved (Fig. 7, left half) using the following protocol:

1. Generate the sense-strand library from the 25  $\mu\text{l}$  pool of concentrated amplicons prepared in step 18 from Section 3.4 by reacting with 5 U of  $\lambda$ -exonuclease at 37 °C for 1 h, followed by heat inactivation at 75 °C for 10 min.
2. Purify the digested product using standard phenol chloroform extraction by adding to the sense-strand library a 2× volume (i.e., ~50  $\mu\text{l}$ ) of phenol:chloroform:isoamyl alcohol (25:24:1).
3. Vortex to create an emulsion and centrifuge at 14,000  $\times g$  for 5 min at room temperature.



**Fig. 7** Efficient stoichiometric regeneration of the ssDNA library in Hi-Fi SELEX: (a) ddPCR amplification of retained members results in formation of fully complimentary homoduplexed amplicons, while conventional bulk PCR used in SELEX yields a mixture of homoduplexed and heteroduplexed amplicons; (b)  $\lambda$ -exonuclease processing of the homoduplexed amplicons produced by Hi-Fi SELEX results in stoichiometric recovery of the ssDNA library in 30 min, while relatively little ssDNA product is recovered from the bulk PCR product. Adapted from [15] with journal permission

4. Recover the top aqueous phase and repeat **steps 2** and **3** once more.
5. Mix the recovered top aqueous phase with 2.5 volumes of ice-cold 100% ethanol and incubate at  $-80\text{ }^\circ\text{C}$  for 2 h to precipitate the nucleic acids.
6. Following this incubation, centrifuge the tube at  $14,000 \times g$  for 30 min in a refrigerated centrifuge set at  $4\text{ }^\circ\text{C}$ .
7. Wash the resulting pellet, which should be clearly visible, twice with 500  $\mu\text{l}$  of 70% ethanol by centrifugation at  $14,000 \times g$  and  $4\text{ }^\circ\text{C}$  for 2 min.
8. Air-dry the recovered nucleic acid pellet.
9. Resuspend the pellet in 25  $\mu\text{l}$  AF buffer. Following regeneration, the retained library is typically at a concentration of ca. 1  $\mu\text{M}$ , which is suitable for the next round of selection.

### 3.7 Measuring the Mean $K_d$ of the Retained and Regenerated Library

In conventional SELEX, the selection process is typically monitored by measuring the mean  $K_d$  of retained pools after blindly conducting a few selection rounds to enrich sufficient amounts of high-affinity binders to make  $K_d$  determination possible (typically by surface plasmon resonance (SPR) or fluorescence spectroscopy).

Parallel PCR amplifications followed by antisense strand removal are required to produce the required quantities of regenerated ssDNA library. If the mean  $K_d$  is to be determined by spectroscopy, that process must include fluorescent labeling or biotinylation of the regenerated library; no modification is required if SPR is to be used. Either process is time consuming, fails to provide information in the critical early rounds of selection, and is costly if SPR is used to measure mean  $K_d$  values. Moreover, a risk of altered binding properties that bias selection is created if modification of library members is required.

To reduce costs and eliminate need for specialized SPR or fluorescence spectroscopy equipment, Hi-Fi SELEX incorporates a novel qPCR-based method to quantify the mean  $K_d$ . The method is based on creating a standard relation to quantify the concentration of library  $C_{\text{Library}}$  in a given sample before and after incubation with an immobilized target, and then using that knowledge to measure the adsorption isotherm for binding of the regenerated library to protein target presented on Nunc Immobilizer Amino plates as described above. This simple qPCR method has been used to show that Hi-Fi SELEX typically delivers a retained pool offering a mean  $K_d \leq \mu\text{M}$  after the first round of selection, and a mean  $K_d$  of order nM after three rounds of selection [15]. It may be used to monitor the mean  $K_d$  after each round of selection and to terminate selection when the mean bulk affinity of retained members either remains unchanged for consecutive rounds or has reached a suitable value (often a mean  $K_d$  of order nM) for a particular application. It requires accurate determination of  $C_{\text{Library}}$  values, necessitating a statistically more robust measurement protocol.

The protocol is as follow:

1.  $2\times$  dilute in SP the retained 25  $\mu\text{l}$  of 80-nt ssDNA library.
2. Take 4  $\mu\text{l}$  of that diluted library and dilute it 1000-fold in SP.
3. From **step 2**, prepare a serial dilution set ( $2\times, 4\times, 8\times, 16\times, 32\times, 64\times, 128\times, 256\times$ ) to a final volume of 500  $\mu\text{l}$  each.
4. Subject two 5  $\mu\text{l}$  aliquots of each dilution to qPCR-based determination using **steps 2** and **3** from Section 3.4.
5. Record the mean  $C_q$  value (for a given RFU threshold value  $T$ ) and standard error computed from the duplicates for each dilution.
6. In a base-10 semi-logarithmic graph, plot the threshold cycle versus the dilution factor and fit the data to a straight line and record the slope and error in it.
7. Compute the amplification efficiency  $E$  from that slope as  $10^{-1/\text{slope}} - 1$ .
8. Determine  $C_{\text{Library}}$  from the  $C_q$  value for each dilution using the standard relation (see Note 13)

$$\log(C_{\text{Library}}) = \log(T) - C_q \log(E) \quad (4)$$

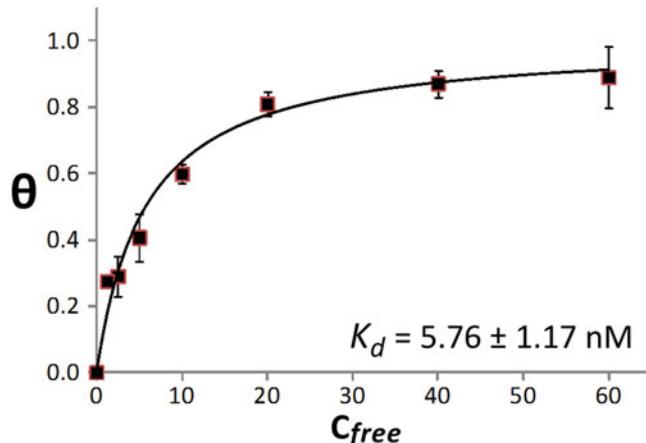
9. In duplicate, incubate each of the 8 serial dilutions (200  $\mu$ l each) in a target-presenting Nunc well for 1 h at 25 °C with gentle mixing at 300 rpm.
10. Remove the solution and then successively wash the retained pools in the equilibrated wells 3× with 300  $\mu$ l of AF.
11. Remove the final wash from each well.
12. Elute the retained members in 100  $\mu$ l of 50 mM NaOH at 70 °C.
13. Immediately neutralize the eluted pool from each well with 100  $\mu$ l of 50 mM HCl.
14. Add 100  $\mu$ l of 10 mM TE buffer.
15. Dilute 5  $\mu$ l of each neutralized eluted pool by 50-fold to make the library concentration appropriate for analysis.
16. In duplicate, quantify the  $C_{\text{Library}}$  in each neutralized eluent by qPCR according to steps 2 and 3 from Section 3.4.
17. For each dilution, determine the fraction  $\theta$  of the library that is bound to the target as the ratio of the dilution-corrected values of  $C_{\text{Library}}$  for the eluted and initial samples. The concentration of library remaining in the solution phase ( $C_{\text{free}}$ ) after equilibration with the immobilized target is given by the difference in dilution-corrected  $C_{\text{Library}}$  values for the initial and eluted libraries, respectively.
18. From the  $\theta$  and  $C_{\text{free}}$  values, construct the binding isotherm as illustrated in Fig. 8. Nonlinear fitting of the Langmuir isotherm relation

$$\theta = \frac{C_{\text{free}}}{K_d + C_{\text{free}}} \quad (5)$$

to the binding isotherm data yields a value for the mean  $K_d$ .

## 4 Notes

1. The initial library must be synthesized at a 1  $\mu$ mole scale in order to generate enough material for conducting Hi-Fi SELEX. Following synthesis, the library is purified as described and then lyophilized. When required, the lyophilized product is reconstituted at a 100  $\mu$ M concentration in 1× AF buffer and stored at –20 °C away from all other reagents used in the Hi-Fi-SELEX protocol.
2. Covalent immobilization of the target onto the well surface proceeds through reaction of amino groups (or other nucleophiles) on the protein with the electrophilic coupling agents



**Fig. 8** Representative adsorption isotherm and mean  $K_d$  value determined using the qPCR based binding analysis method: data shown are for a retained pool of tight binding library members with fixed regions against human  $\alpha$ -thrombin

displayed on polyglycol chain ends of the Nunc Immobilizer well. The manufacturer recommends conducting this reaction at pH 9.6, but the use of a basic pH raises possibilities for chemical modifications (e.g., de-amidation) that alter target protein structure, chemistry and/or activity. Hi-Fi SELEX therefore uses the alternate but equally efficient coupling conditions described in steps 1–7 of the protocol. Under these milder conditions, the reaction proceeds more slowly and, in coordination with selection of an appropriate target solution concentration, permits facile control of the surface density of immobilized target.

3. The hydrodynamic diameter of a random 80-mer ssDNA aptamer is ca. 8 nm. We therefore set the solution concentration (50–100 nM) of the target protein in the coupling reaction so as to achieve a mean distance of separation of immobilized target molecules that is a bit greater than 8 nm (i.e., 9–10 nm). For thrombin (MW = 37 kDa), this is achieved using a solution concentration of 80 nM. A lower concentration would be required for targets of higher molecular weight, and vice versa. To determine the optimal solution concentration, the bound moles of target are measured as a function of solution concentration using a mass balance and  $A_{280\text{ nm}}$  values for the starting and final solution. The contact area of the well ( $0.95\text{ cm}^2$ ) is then used to compute the average distance of separation of immobilized target as a function of the starting solution concentration used.
4. During this wash sequence, the amines (Tris (hydroxyl-methyl-aminomethane)) present in SP buffer neutralize unreacted

electrophiles displayed on the end-grafted polyglycol surface. Supplementing the SP buffer used in both the 3X wash and 1X incubation steps with 0.005% Tween 20 *is absolutely essential* to fully passivate the surface against nonspecific retention.

5. In conventional SELEX, one round of screening in the absence of the molecular target is often first performed in an attempt to reduce the number of members of the library retained through mechanisms unrelated to the target. This step is not required in Hi-Fi SELEX, as the unique surface passivation methods used have been shown to reduce nonspecific retention to undetectable levels [15].
6. The competent library concentration used (10  $\mu$ M), in combination with setting the surface density of immobilized target to achieve a mean separation distance of ~9 nm, results in library screening within the theoretically preferred 100:1 to 1000:1 aptamer-to-target range [46, 47], while also eliminating the possibility of bridging of bound library members between proximal targets, an effect that has been shown to confound aptamer selection [48].
7. The polyanionic form of aptamers naturally creates the unwanted potential to over-select for sequences whose binding is dominated by coulombic (electrostatic) interactions. The second set of high-salt washes employed in Hi-Fi SELEX are designed to remove library members that bind the target through ion-exchange type mechanisms that generally lack sufficient specificity for the target.
8. It is important to ensure that the immobilized target and aptamers are always kept hydrated. When done properly, the passivation step performed after target immobilization will significantly increase the hydrophilicity of the Nunc Immobilizer Amino well surface, in part due to the presence of Tween-20. The surface of the wells thereby remains hydrated during various required solution exchanges provided vacuum aspirators are not used. A single or multichannel pipette should instead be used to add and remove solutions.
9. The molecular weight of the 80-nt aptamer is 24 kDa. We therefore use a 10 kDa molecular weight cut-off membrane for this step to ensure good aptamer recovery yields. Buffer exchange should proceed until the final concentration of salt in the aptamer pool falls to within the mM range.
10. The number of reaction wells required (and thus total droplets generated) is defined by setting the average copies per droplet ( $CPD = -\ln(\text{empty droplets}/\text{total read droplets})$ ) at 50 library members. We have found that at a CPD of 50 (or lower), end-point amplification of the library is achieved with minimal accumulation of HMW by-products [15]. However, ddPCR at

a CPD > 50 results in by-product formation and a concomitant loss of desired 80-bp amplicon product (and product quality) in a manner similar to that observed in standard PCR of retained library members. The CPD limit applies to libraries retained in the first two rounds of selection, where each of the 50 templates within each droplet is almost certain to harbour a unique sequence within its variable core region. In later rounds of selection, duplicate library sequences will increasingly partition into any given drop, diminishing template heterogeneity per drop and the probability of forming by-products during amplification. Higher CPDs may then be used, up to a CPD of no greater than 1000.

11. As indicated in Fig. 6, the qPCR-based diversity assay provides reliable estimates of sequence diversity for retained libraries comprising  $\geq 1000$  unique variable-region sequences. Hi-Fi SELEX screening is typically stopped before the retained library diversity falls to that limit, as the target mean  $K_d$  of the retained population has generally been met. But in a case where characterization of less diverse libraries is required, one may employ an alternative PCR-based method for characterizing pool diversities that relies on fluorometric quantitation of dsDNA after amplification. The protocol for that method is provided in a previous publication [15].
12. The number of cycles is set so as to generate sufficient amplified material from the amount of retained library present in the previously quantified starting aliquot. Typically, ca. 13 to 15 cycles is sufficient to generate the  $\geq 10$  ng of 80-bp amplicon required for the diversity assay. Note that cycling should be stopped ( $\sim 16$ –20 cycles) before production of HMW amplification by-products is observed in gel-based visualization of the reaction product. The pilot PCR reaction is performed first to determine the optimal cycle number, where a small amount of the pool is amplified using standard 50  $\mu$ l PCR. A sample is taken out every 3 cycles and a gel-based visualization is performed similar to that in Fig. 4. The optimal cycle number should yield enough material to proceed to the next step. The amount of product can be quantified using NanoDrop<sup>TM</sup> ND2000 spectrophotometer.
13. Determination of library concentrations can be done using other methods. In our experience, the oft-used NanoDrop<sup>TM</sup> ND2000 spectrophotometer is not sufficiently accurate in measuring dsDNA at concentrations below ca. 50  $\mu$ g/ml, typically overestimating the true concentration. The Qubit<sup>TM</sup> Fluorescence Monitor is generally accurate at these concentrations, and one can use it as a secondary check of  $C_{Library}$  values.

## References

1. Edwards BM, Barash SC, Main SH, Choi GH, Minter R, Ullrich S, Williams E, Du Fou L, Wilton J, Albert VR, Ruben SM, Vaughan TJ (2003) The remarkable flexibility of the human antibody repertoire–isolation of over one thousand different antibodies to a single protein, BLyS. *J Mol Biol* 334:103–118
2. Xing PX, Russell S, Prenzoska J, McKenzie JF (1994) Discrimination between alternatively spliced STP-A and -B isoforms of CD46. *Immunology* 83:122–127
3. Raska M, Czernekova L, Moldoveanu Z, Zachova K, Elliott M, Novak Z, Hall S, Hoelscher M, Maboko L, Brown R, Smith P, Mestecky J, Novak J (2014) Differential glycosylation of envelope gp120 is associated with differential recognition of HIV-1 by virus-specific antibodies and cell infection. *AIDS Res Ther* 11:1–16
4. Mould DR, Sweeney KR (2007) The pharmacokinetics and pharmacodynamics of monoclonal antibodies–mechanistic modeling applied to drug development. *Curr Opin Drug Discov Devel* 10:84–96
5. Chapman AP, Antoniw P, Spitali M, West S, Stephens S, King DJ (1999) Therapeutic antibody fragments with prolonged in vivo half-lives. *Nat Biotech* 17:780–783
6. Chames P, Van Regenmortel M, Weiss E, Baty D (2009) Therapeutic antibodies: successes, limitations and hopes for the future. *Br J Pharmacol* 157:220–233
7. Ruigrok VJ, Levisson M, Eppink MH, Smidt H, van der Oost J (2011) Alternative affinity tools: more attractive than antibodies? *Biochem J* 436:1–13
8. Skerra A (2007) Alternative non-antibody scaffolds for molecular recognition. *Curr Opin Biotechnol* 18:295–304
9. Jayasena SD (1999) Aptamers: an emerging class of molecules that rival antibodies in diagnostics. *Clin Chem* 45:1628–1650
10. Tuerk C, Gold L (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249(4968):505–510
11. Ellington AD, Szostak JW (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature* 346(6287):818–822
12. Robertson DL, Joyce GF (1990) Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature* 344 (6265):467–468
13. Ozer A, Pagano JM, Lis JT (2014) New technologies provide quantum changes in the scale, speed, and success of SELEX methods and aptamer characterization. *Mol Ther Nucleic Acids* 3:e183
14. Ouellet E, Lagally ET, Cheung KC, Haynes C (2014) A simple method for eliminating fixed-region interference of aptamer binding during SELEX. *Biotechnol Bioeng* 111 (11):2265–2279
15. Ouellet E, Foley JH, Conway EM, Haynes C (2015) Hi-fi SELEX: a high-fidelity digital PCR based therapeutic aptamer discovery platform. *Biotechnol Bioeng* 112(8):1506–1522
16. Bartel DP, Zapp ML, Green MR, Szostak JW (1991) HIV-1 rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* 67:529–536
17. Lorenz C, von Pelchrzim F, Schroeder R (2006) Genomic systematic evolution of ligands by exponential enrichment (genomic SELEX) for the identification of protein-binding RNAs independent of their expression levels. *Nat Protoc* 1:2204–2212
18. Zimmermann B, Bilusic I, Lorenz C, Schroeder R (2010) Genomic SELEX: a discovery tool for genomic aptamers. *Methods* 52:125–132
19. Keefe AD, Cload ST (2008) SELEX with modified nucleotides. *Curr Opin Chem Biol* 12:448–456
20. Kuwahara M, Obika S (2013) In vitro selection of BNA (LNA) aptamers. *Artif DNA PNA XNA* 4:39–48
21. Klussmann S, Nolte A, Bald R, Erdmann VA, Furste JP (1996) Mirror-image RNA that binds D-adenosine. *Nat Biotech* 14:1112–1115
22. Hall B, Micheletti JM, Satya P, Ogle K, Pollard J, Ellington AD (2001) Design, synthesis, and amplification of DNA pools for in vitro selection. In: Current protocols in molecular biology. John Wiley & Sons, Inc., Hoboken, NJ, pp 24.2.1–24.2.27
23. Jiménez E, Sefah K, López-Colón D, Van Simaeys D, Chen HW, Tockman MS, Tan W (2012) Generation of lung adenocarcinoma DNA aptamers for cancer studies. *PLoS One* 7(10):e46222
24. Sefah K, Shangguan D, Xiong X, O'Donoghue MB, Tan W (2010) Development of DNA aptamers using cell-SELEX. *Nat Protoc* 5:1169–1185
25. Gopinath SC, Sakamaki Y, Kawasaki K, Kumar PK (2006) An efficient RNA aptamer against human influenza B virus hemagglutinin. *J Biochem* 139:837–846

26. Weng C-H, Huang C-J, Lee G-B (2012) Screening of aptamers on microfluidic systems for clinical applications. *Sensors* 12 (7):9514–9529
27. Gong P, Grainger DW (2007) Nonfouling surfaces: a review of principles and applications for microarray capture assay designs. *Methods Mol Biol* 381:59–92
28. Cattani-Scholz A, Pedone D, Blobner F, Abstreiter G, Schwartz J, Tornow M, Andruzz L (2009) PNA-PEG modified silicon platforms as functional bio-interfaces for applications in DNA microarrays and biosensors. *Biomacromolecules* 10:489–496
29. Schlapak R, Pammer P, Armitage D, Zhu R, Hinterdorfer P, Vaupel M, Fröhwirth T, Howorka S (2005) Glass surfaces grafted with high-density poly(ethylene glycol) as substrates for DNA oligonucleotide microarrays. *Langmuir* 22:277–285
30. Poncin-Epaillard F, Vrlinic T, Debarnot D, Mozetic M, Coudreuse A, Legeay G, El Moualij B, Zorzi W (2012) Surface treatment of polymeric materials controlling the adhesion of biomolecules. *J Funct Biomater* 3:528–543
31. SantaLucia Jr J, Hicks D (2004) The thermodynamics of DNA structural motifs. *Annu Rev Biophys Biomol Struct* 33:415–440
32. Lou X, Qian J, Xiao Y, Viel L, Gerdon AE, Lagally ET, Atzberger P, Tarasow TM, Heeger AJ, Soh HT (2009) Micromagnetic selection of aptamers in microfluidic channels. *Proc Natl Acad Sci* 106:2989–2994
33. Nieuwlandt D (2000) In vitro selection of functional nucleic acid sequences. *Curr Issues Mol Biol* 2:9–16
34. Musheev MU, Krylov SN (2006) Selection of aptamers by systematic evolution of ligands by exponential enrichment: addressing the polymerase chain reaction issue. *Anal Chim Acta* 564:91–96
35. Nakano M, Komatsu J, Matsura S-i, Takashima K, Katsura S, Mizuno A (2003) Single-molecule PCR using water-in-oil emulsion. *J Biotechnol* 102:117–124
36. Williams R, Peisajovich SG, Miller OJ, Magdassi S, Tawfik DS, Griffiths AD (2006) Amplification of complex gene libraries by emulsion PCR. *Nat Methods* 3:545–550
37. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen Y-J, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer MLI, Jarvie TP, Jirage KB, Kim J-B, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376–380
38. Schütze T, Wilhelm B, Greiner N, Braun H, Peter F, Mörl M, Erdmann VA, Lehrach H, Konthur Z, Menger M, Arndt PF, Glöckler J (2011) Probing the SELEX process with next-generation sequencing. *PLoS One* 6:e29604
39. Hoon S, Zhou B, Janda KD, Brenner S, Scolnick J (2011) Aptamer selection by high-throughput sequencing and informatic analysis. *BioTechniques* 51:413–416
40. Kai E, Sumikura K, Ikebukuro K, Karube I (1998) Purification of single stranded DNA from asymmetric PCR product using the SMART system. *Biotechnol Tech* 12:935–939
41. Pagratis NC (1996) Rapid preparation of single stranded DNA from PCR products by streptavidin induced electrophoretic mobility shift. *Nucleic Acids Res* 24:3645–3646
42. Civit L, Fragoso A, O'Sullivan CK (2012) Evaluation of techniques for generation of single-stranded DNA for quantitative detection. *Anal Biochem* 431:132–138
43. Svobodova M, Pinto A, Nadal P, O'Sullivan CK (2012) Comparison of different methods for generation of single-stranded DNA for SELEX processes. *Anal Bioanal Chem* 404:835–842
44. Avci-Adali M, Paul A, Wilhelm N, Ziemer G, Wendel HP (2009) Upgrading SELEX technology by using lambda exonuclease digestion for single-stranded DNA generation. *Molecules* 15:1–11
45. Lee G, Yoo J, Leslie BJ, Ha T (2011) Single-molecule analysis reveals three phases of DNA degradation by an exonuclease. *Nat Chem Biol* 7:367–374
46. Irvine D, Tuerk C, Gold L (2001) Selexion: systematic evolution of ligands by exponential enrichment with integrated optimization by non-linear analysis. *J Mol Biol* 222 (3):739–761
47. Vant-Hull B, Payano-Baez A, Davis RH, Gold L (1998) The mathematics of SELEX against complex targets. *J Mol Biol* 278(3):579–597
48. Ozer A, White BS, Lis JT, Shalloway D (2013) Density-dependent cooperative non-specific binding in solid-phase SELEX affinity selection. *Nucleic Acids Res* 41:7167–7175

# INDEX

## A

- Absolute quantification ..... 4, 6, 18, 19, 42, 70, 99, 113, 120, 168, 230, 258, 265, 276, 283, 316, 323, 350, 364, 373, 377–380, 398, 425, 446, 453, 459–470, 478, 519  
Accuracy ..... 7, 15, 16, 18, 20, 21, 78, 115, 119, 123, 138, 144, 174, 183, 230, 258, 296, 375, 380, 389, 397, 414, 446, 478  
Allele burden ..... 5, 257, 267–269  
Allele-specific expression ..... 407, 414, 420  
Allelic imbalance ..... 408–414, 418, 420, 490  
Allelic skew ..... 402, 404, 406, 415  
Aptamers ..... 7, 531–552  
Assay design ..... v, 4, 5, 7, 32, 109, 144, 146–148, 151, 152, 174, 194, 195, 201, 242, 277, 279–280, 289, 290, 312, 313, 351, 352, 354, 357–359, 367, 369, 374, 403, 405–410, 425, 441, 494, 497, 499  
Assay multiplexing ..... 79, 283–286, 327, 428

## B

- BCL1/IGH ..... 230–237, 240–244, 250, 251  
BCL2/IGH ..... 241, 251  
Biofluids ..... 465  
Biological affinity reagents ..... 531  
Biomarkers ..... 4, 12, 41, 111–125, 303, 304, 320, 335, 364, 388, 445, 459, 514  
Blood ..... 5, 26, 45, 100, 111, 127, 162, 178, 194, 210, 232, 259, 279, 303, 323, 337, 366, 432, 445, 460, 493  
Breast cancer ..... 4, 5, 169, 170, 209–226, 234, 304, 315, 318, 321, 389, 423, 460

## C

- Calibration ..... 12, 112, 123, 426, 432  
Cancer ..... 4, 12, 143, 161, 174, 193, 209, 234, 275, 303, 323, 363, 388, 460, 513  
Cell-free circulating tumor DNA ..... 275, 303–321

- Cell-free DNA (cfDNA) ..... 5, 46–50, 52–62, 161, 164, 166, 167, 194, 196–199, 202, 211, 214, 223, 278, 279, 298, 303–305, 310, 312, 318, 319, 335–337, 340–343, 345–347, 365, 442, 509  
Cell-sorting ..... 424, 426  
Cerebrospinal fluid (CSF) ..... 4, 101, 102, 107, 111, 114, 116, 120, 445, 460, 465  
Circulating cell-free DNA (ccfDNA) ..... 45, 324–327  
Clonal evolution ..... 275  
Co-localization ..... 393, 490, 491  
Colorectal cancer (CRC) ..... 13, 363  
Complex alleles ..... 490  
Compound heterozygote ..... 489  
Control ..... 7, 32, 46, 74, 100, 115, 136, 144, 162, 178, 195, 213, 234, 261, 278, 313, 325, 343, 351, 366, 402, 431, 446, 480, 492, 514, 550  
Control materials ..... 19, 20, 45  
Copy number variants (CNVs) ..... 174, 176, 177  
Copy number variation (CNV) ..... 4, 16, 45, 143–159, 326  
Counting ..... 8, 19, 25–43, 332, 335, 352, 402, 470, 488, 516, 517, 523  
CRISPR/Cas9 ..... 350

## D

- Deoxyribo nucleic acid (DNA) ..... 3, 12, 45, 70, 100, 112, 144, 161, 173, 194, 210, 231, 258, 275, 323, 335, 349, 363, 365, 387, 423, 446, 490, 513, 532  
Diagnostics ..... 12, 13, 21, 22, 41, 45, 46, 48, 70, 111, 112, 128, 235, 237, 245, 248, 257, 258, 304, 531  
Digital PCR (dPCR) ..... 3, 11, 25, 46, 70, 162, 423  
DNA extraction ..... 128  
DNA integrity ..... 74, 92, 137, 377  
DNA methylation ..... 296, 363–383  
Droplet ..... 5, 14, 61, 70, 99, 122, 128, 144, 163, 174, 198, 210, 230, 279, 304, 324, 336, 350, 364, 389, 402, 425, 446, 460, 478, 490, 514, 532

- Droplet digital PCR (ddPCR) ..... 5, 51, 70, 99, 113, 128, 144, 163, 173, 194, 210, 230, 277, 304, 323, 336, 350, 363, 389, 402, 425, 446, 460, 479, 492, 515, 532  
Duplex digital PCR ..... 128, 276

## E

- Efficiency ..... 4, 12, 16, 19–21, 46, 47, 52, 54–56, 58, 59, 61, 70, 92, 123, 155, 239, 249, 253, 280, 319, 374, 381, 389, 391, 395, 398, 402, 440, 441, 446, 459, 470, 537, 548  
*EGFR* ..... 193–196, 199–202, 489  
*Enterococcus* ..... 127, 128, 130, 133, 136, 139  
Extraction ..... 7, 45, 74, 101, 112, 128, 162, 211, 232, 305, 324, 337, 391, 446, 491, 545

## F

- False negative ..... 20, 21, 26, 31, 32, 38–40, 79, 92, 139, 315  
False positive ..... 8, 21, 26, 31, 32, 35, 36, 38–41, 62, 70, 79, 94, 95, 119, 202, 205, 297, 309, 315, 320, 321, 453  
Fecal contamination ..... 4, 127–139  
Follicular lymphoma (FL) ..... 230, 235, 236, 241, 244, 252

## G

- Gene expression ..... 6, 143, 363, 397, 401, 423, 424, 432  
Genetic testing ..... 6, 275, 335  
Genetically modified organism (GMO) ..... v, 4, 69, 70, 74, 77, 78  
Genome editing ..... v, 12, 349  
Genomic structural variation ..... v  
Genotyping ..... 193, 194, 202, 205, 267, 276, 323, 337, 343, 344, 490, 509  
Genotyping assay design ..... 194, 267, 326, 327  
Genotyping inversions ..... 490, 494  
GM Maize ..... 69, 71, 72, 75, 76, 80, 83, 85, 87, 88, 90, 91, 95

## H

- Haplotype ..... 144, 416, 421, 490  
*HER2* ..... 4, 161–172  
HF183 ..... 128, 130, 133, 139  
HHV\_6A ..... 99, 103, 104  
HHV\_6B ..... 99, 103, 104  
Homology-directed repair (HDR) ..... 6, 349–361  
Human herpesvirus 6 ..... 99–109  
Human telomerase reverse transcriptase (hTERT) ..... 388–396, 398, 514

## I

- Immunoglobulin heavy-chain gene (IGH) ..... 230–240, 242–244, 248, 251

- Inhibitor tolerance ..... 111

## J

- JAK2 V617F ..... 5, 257–272

## L

- Large DNA isolation ..... 240, 367  
Library balancing ..... 478, 487  
Library loading ..... 305, 307, 478  
Limit of blank (LOB) ..... 21, 34, 35  
Limit of detection (LOD) ..... 7, 21, 27, 33, 34, 36, 58, 117, 355, 414, 460  
Linkage ..... v, 6, 401, 404, 407, 490–499, 503, 505–510  
Liquid biopsy ..... 4, 5, 26, 210, 426

## M

- Mantle cell lymphoma (MCL) ..... 230, 235–238, 240, 245, 246, 248, 250  
MethyLight ..... 364, 365, 369, 370, 373, 374, 376–380, 382  
MethyLight ddPCR ..... 364, 365, 368–370, 373, 374, 376–379, 382  
Microbial source tracking ..... 128  
MicroRNA (miRNA) ..... 6, 445, 454, 459  
Minimal residual disease (MRD) ..... 229, 230, 233–236, 238, 242–246, 248, 250–252, 254  
Mitochondrial DNA (mtDNA) ..... 4, 111, 114, 116, 120, 303  
Mobile elements ..... 174  
mRNA alternative splicing ..... 387–400  
mRNA expression ..... 408, 409  
Multiple myeloma (MM) ..... 230, 235, 236, 238, 239, 246  
Multiplex ..... 5, 16, 48, 69–95, 100, 113, 117, 119, 128, 277, 294, 327, 377, 392, 425, 431, 436, 438  
Mutation detection ..... 258, 292  
Myeloproliferative neoplasms (MPNs) ..... 257

## N

- Nanofluidic chip ..... 14  
Next generation sequencing (NGS) ..... 3, 5, 6, 230, 275, 298, 423, 425, 477–479, 482, 488  
Nonhomologous end joining (NHEJ) ..... 6, 349, 351, 352, 356–360  
Noninvasive diagnosis ..... 303  
Noninvasive genetic profiling ..... 275

Non-small cell lung cancer (NSCLC)..... 5, 193, 194, 202  
Nucleic acids quantification..... 4, 6, 13, 16, 18, 479

**P**

Partitioning ..... v, 6, 15, 16, 19, 25, 28–31, 41, 99, 113, 144, 175, 258, 402, 459, 490, 531, 541  
Pathogen detection ..... 275  
Performance characteristics ..... 18, 37, 38  
Personalized medicine ..... 489  
Phasing ..... 489–511  
PIK3CA ..... 210  
Plasma ..... 5, 45, 107, 194, 210, 279, 303, 323, 335, 365, 446, 460  
Poisson distribution ..... 27–30, 35, 41, 70, 169, 222, 276, 332, 470  
Polymerase chain reaction (PCR) ..... 74, 173, 230, 258, 540  
Precision ..... v, 4, 6, 7, 11, 14–16, 18, 20, 56, 58, 61, 78, 113, 144, 157, 159, 174, 186, 187, 195, 216, 261, 265–269, 272, 287, 297, 364, 377, 407, 420, 459, 460, 532

**Q**

Quantification ..... 3, 12, 25, 46, 69, 99, 112, 128, 162, 174, 230, 245, 276, 303, 323, 335, 350, 364, 388, 425, 445, 459, 478, 514  
Quantitative PCR (qPCR) ..... 4, 12, 16, 19, 20, 69, 70, 113–115, 128, 144, 148, 175, 230, 245, 251, 254, 276, 280, 343, 344, 389, 398, 410, 446, 470, 473, 478, 535, 540, 542, 545, 548–550, 552

**R**

Radial multiplexing ..... 6, 425, 426, 429, 442  
Rare allele detection ..... 5, 265, 277  
Rare mutation detection ..... 292  
Real-time PCR ..... 3, 74, 78, 92, 144, 336, 337, 341–343, 346, 347, 388, 402, 459  
Rejection monitoring ..... 335  
Reproducibility ..... v, 4, 7, 13, 20, 22, 112, 113, 123, 230, 258, 449, 460, 478  
Resistance mutations ..... 5, 205, 304  
Ribo nucleic acid (RNA) ..... 3, 74, 129, 164, 304, 342, 388, 401–421, 424, 446, 461, 490, 513  
Risk biomarkers ..... 364  
RNA quantification ..... 6

**S**

Sensitivity ..... v, 5, 8, 14, 33–36, 38–40, 42, 56, 69, 78, 94, 102, 107, 109, 112, 113, 115, 119, 128, 161, 166, 174, 176, 177, 183, 202, 210, 213, 214, 220, 229, 230, 245, 254, 258, 259, 261, 275, 279, 287, 290, 297–300, 305, 307, 309, 324, 350, 361, 364, 377, 389, 398, 402, 425, 439, 460, 470, 488, 514, 523  
Sequencing library qualification ..... 347, 478  
Sequencing library quantification ..... 478, 488  
Serum ..... 6, 49, 101, 102, 107, 210, 303, 319, 365, 432, 446–448, 452, 453, 459

Single cell ..... 423–442  
Single color ..... 323–333  
Single nucleotide variations (SNVs) ..... 174–176, 181, 186, 326  
Somatic mosaicism ..... v, 4, 174  
Specificity ..... v, 8, 20, 33–35, 38–40, 61, 62, 78, 94, 104, 109, 113, 128, 175, 179, 186, 202, 205, 210, 213, 220, 244, 251, 275, 276, 290, 297, 299, 326, 379, 392, 395, 470, 525, 531, 532, 536, 551

Standardization ..... 46, 111, 230, 245, 258

Statistics ..... 15, 16, 18, 19, 25, 27, 28, 30, 31, 34, 94, 157, 158, 203, 225, 258, 264, 315, 380, 412

Subsampling ..... 16, 26, 27, 30, 31, 36, 40–42

Systematic evolution of ligands by exponential enrichment (SELEX) ..... 7, 531, 533, 534, 541, 544, 547, 551

**T**

t(11;14) translocation ..... 230  
t(14;18) translocation ..... 230  
TALEN ..... 350, 360, 361  
Telomerase activity ..... 6, 388, 389, 398, 513–528  
Transcripts ..... 6, 387, 393, 394, 399, 402, 420, 423–425, 429, 438, 441, 442  
Transplantation ..... 100, 335–347

**V**

Validation ..... 11–22, 38, 78, 111, 144, 174, 177, 186, 245, 283–286, 310, 313, 316, 354, 355, 445, 446  
Viral detection ..... 100

**W**

Water quality ..... 4, 11, 128, 129  
Whole-genome sequencing ..... 176, 304, 310