

Methods in
Molecular Biology 2065

Springer Protocols

Roberto Biassoni
Alessandro Raso *Editors*

Quantitative Real-Time PCR

Methods and Protocols
Second Edition

 Humana Press

METHODS IN MOLECULAR BIOLOGY

Series Editor

John M. Walker

School of Life and Medical Sciences

University of Hertfordshire

Hatfield, Hertfordshire, UK

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

For over 35 years, biological scientists have come to rely on the research protocols and methodologies in the critically acclaimed *Methods in Molecular Biology* series. The series was the first to introduce the step-by-step protocols approach that has become the standard in all biomedical protocol publishing. Each protocol is provided in readily-reproducible step-by-step fashion, opening with an introductory overview, a list of the materials and reagents needed to complete the experiment, and followed by a detailed procedure that is supported with a helpful notes section offering tips and tricks of the trade as well as troubleshooting advice. These hallmark features were introduced by series editor Dr. John Walker and constitute the key ingredient in each and every volume of the *Methods in Molecular Biology* series. Tested and trusted, comprehensive and reliable, all protocols from the series are indexed in PubMed.

Quantitative Real-Time PCR

Methods and Protocols

Second Edition

Edited by

Roberto Biassoni

Molecular Diagnostics, IRCCS, Istituto Giannina Gaslini, Genova, Italy

Alessandro Raso

ASL3 Sistema Sanitario Regione Liguria, S.C. Laboratorio d'Analisi, Genova, Italy



Editors

Roberto Biassoni
Molecular Diagnostics
IRCCS, Istituto Giannina Gaslini
Genova, Italy

Alessandro Raso
ASL3 Sistema Sanitario Regione Liguria
S.C. Laboratorio d'Analisi
Genova, Italy

ISSN 1064-3745

Methods in Molecular Biology

ISBN 978-1-4939-9832-6

<https://doi.org/10.1007/978-1-4939-9833-3>

ISSN 1940-6029 (electronic)

ISBN 978-1-4939-9833-3 (eBook)

© Springer Science+Business Media, LLC, part of Springer Nature 2020

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.

This Humana 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.

Dedication

With great sadness, we have learned that Kary Mullis, the inventor of the PCR technique and the Chemistry Nobel Prize winner for this discovery, passed away on August 7th during the editing phase of this volume. His contribution to the molecular biology field, thanks to this technique and its modifications, has obviously been immense not only as a research technique but later on for countless clinical diagnostics, forensic, and industrial applications. He will be missed.

Preface

Polymerase chain reaction (PCR) and, later on, its quantitative evolution (qPCR) have been initially used in scientific research only, but qPCR's flexibility, sensitivity, and velocity made it a suitable technique to be used in a wide range of applications.

From the first description of qPCR 25 years ago, nonprofit researchers and private companies have developed a huge number of its applications in different disciplines, leading to significant improvements even in the diagnostics field. Under this point of view, qPCR has needed a change in the instrumentation that was possible, thanks to the technological implementation and the development of new reagents in order to also fulfill ethical and legal issues. So, qPCR is now an up-to-date technology widely used in research and clinical diagnostics.

Five years ago, we edited the first volume about qPCR in the Methods in Molecular Biology series that was not monothematic but regarded different fields and applications. In the time since its publication, the first edition of the book seems to have inspired a fairly good amount of interest, so we followed the same idea in this second edition of the qPCR book. Thanks to the contribution of experts in the field, we tried as much as possible to integrate new methodological applications of qPCR developed since the first edition.

We would like to thank all the colleagues who have contributed to the book and have helped us during the editing process. Moreover, we hope to arouse interest in the qPCR community once again.

Genova, Italy

Roberto Biassoni

Alessandro Raso

Contents

Preface	v
Contributors	ix
1 A Quarter Century of PCR-Applied Techniques and Their Still-Increasing Fields of Use	1
<i>Alessandro Raso and Roberto Biassoni</i>	
2 Parameters for Successful PCR Primer Design	5
<i>Stephen A. Bustin, Reinhold Mueller, and Tania Nolan</i>	
3 MIQE-Compliant Validation of MicroRNA Biomarker Signatures Established by Small RNA Sequencing	23
<i>Veronika Mussack, Stefanie Hermann, Dominik Buschmann, Benedikt Kirchner, and Michael W. Pfaffl</i>	
4 Enhanced Probe-Based RT-qPCR Quantification of MicroRNAs Using Poly(A) Tailing and 5' Adaptor Ligation	39
<i>Valentin Vautrot and Isabelle Behm-Ansmant</i>	
5 A Novel System to Discriminate HLA-C mir148a Binding Site by Allele-Specific Quantitative PC R	55
<i>Priscilla Biswas, EddiDi Marco, and Mauro S. Malnati</i>	
6 Detection of Yellow Fever Virus by Quantitative Real-Time PCR (qPCR)	65
<i>Gisela Freitas Trindade, Sheila Maria Barbosa de Lima, Constança Britto, and Alice Gomes Fernandes-Monteiro</i>	
7 Evaluation of the Abundance of Fungi in Wastewater Treatment Plants Using Quantitative PCR (qPCR)	79
<i>Paula Maza-Márquez, Elisabet Aranda, Jesús González-López, and Belén Rodelas</i>	
8 Early Detection of Fungal Plant Pathogens by Real-Time Quantitative PCR: The Case of <i>Diplodia sapinea</i> on Pine	95
<i>Nicola Luchi, Alberto Santini, Francesca Salvianti, and Pamela Pinzani</i>	
9 A General Protocol for Accurate Gene Expression Analysis in Plants	105
<i>Ellen De Keyser, Laurence Desmet, Magali Losschaert, and Jan De Riek</i>	
10 Gene Expression Analysis in Bacteria by RT-qPCR	119
<i>Danilo J. P. G. Rocha, Thiago L. P. Castro, Eric R. G. R. Aguiar, and Luis G. C. Pacheco</i>	
11 Detection and Characterization of Circulating Tumor Cells by Quantitative Real-Time PCR	139
<i>Francesca Salvianti, Filomena Costanza, Gemma Sonnati, and Pamela Pinzani</i>	

12	Molecular Monitoring of Chronic Myeloid Leukemia	153
	<i>Katherine Dominy, Katya Mokretar, Alistair G. Reid, and Jamshid S. Khorashad</i>	
13	Normalization in Human Glioma Tissue.	175
	<i>Ana Paula Santin Bertoni, Isabele Cristiana Iser, Rafael Paschoal de Campos, and Márcia Rosangela Wink</i>	
14	qPCR Applications for the Determination of the Biological Age	191
	<i>Mauro Castagnetta, Ulrich Pfeffer, Aldo Chiesa, Elena Gennaro, Massimiliano Ceconni, Domenico Covello, and Nicoletta Sacchi</i>	
15	QuantStudio™ 12K Flex OpenArray® System as a Tool for High-Throughput Genotyping and Gene Expression Analysis	199
	<i>Chiara Broccanello, Letizia Gerace, and Piergiorgio Stevanato</i>	
16	Digital PCR and the QuantStudio™ 3D Digital PCR System	209
	<i>Marion Laig, Christie Fekete, and Nivedita Majumdar</i>	
	<i>Index</i>	233

Contributors

ERIC R. G. R. AGUIAR • *Post-Graduate Program in Biotechnology, Institute of Health Sciences, Federal University of Bahia, Salvador, BA, Brazil*

ELISABET ARANDA • *Department of Microbiology, Institute of Water Research, University of Granada, Granada, Spain*

SHEILA MARIA BARBOSA DE LIMA • *Laboratório de Biologia Molecular e Doenças Endêmicas (LABIMDOE), Fundação Oswaldo Cruz (Fiocruz)/Instituto Oswaldo Cruz (IOC), Manguinhos, RJ, Brazil*

ISABELLE BEHM-ANSMANT • *UMR 7365 IMoPA, Université de Lorraine, Vandoeuvre-lès-Nancy, France; UMR 7365 IMoPA, CNRS, Vandoeuvre-lès-Nancy, France*

ROBERTO BIASSONI • *Molecular Diagnostics, IRCCS, Istituto Giannina Gaslini, Genova, Italy*

PRISCILLA BISWAS • *Unit of Human Virology, Division of Immunology, Transplantation and Infectious Diseases, IRCCS Ospedale San Raffaele, Milan, Italy*

CONSTANÇA BRITTO • *Laboratório de Biologia Molecular e Doenças Endêmicas (LABIMDOE), Fundação Oswaldo Cruz (Fiocruz)/Instituto Oswaldo Cruz (IOC), Manguinhos, RJ, Brazil*

CHIARA BROCCANELLO • *DAFNAE, Università Degli Studi di Padova, Legnaro, (PD), Italy*

DOMINIK BUSCHMANN • *Animal Physiology and Immunology, School of Life Sciences Weihenstephan, Technical University of Munich, Munich, Germany*

STEPHEN A. BUSTIN • *Faculty of Health, Education, Medicine, and Social Care, Anglia Ruskin University, Chelmsford, Essex, UK*

MAURO CASTAGNETTA • *Ente Ospedaliero Ospedali Galliera, Genoa, Italy*

THIAGO L. P. CASTRO • *Post-Graduate Program in Biotechnology, Institute of Health Sciences, Federal University of Bahia, Salvador, BA, Brazil*

MASSIMILIANO CECCONI • *IRCCS Istituto Giannina Gaslini, Genoa, Italy*

ALDO CHIESA • *Ente Ospedaliero Ospedali Galliera, Genoa, Italy*

FILOMENA COSTANZA • *Department of Experimental and Clinical Biomedical Sciences "Mario Serio", University of Florence, Florence, Italy*

DOMENICO COVIELLO • *IRCCS Istituto Giannina Gaslini, Genoa, Italy*

RAFAEL PASCHOAL DE CAMPOS • *Laboratório de Biologia Celular, Departamento de Ciências Básicas da Saúde (DCBS), Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Porto Alegre, RS, Brazil; Laboratório de Sinalização e Plasticidade Celular, Departamento de Biofísica, Instituto de Biociências, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil*

ELLEN DE KEYSER • *Plant Sciences Unit, Fisheries and Food (ILVO), Flanders Research Institute for Agriculture, Melle, Belgium*

JAN DE RIEK • *Plant Sciences Unit, Fisheries and Food (ILVO), Flanders Research Institute for Agriculture, Melle, Belgium*

LAURENCE DESMET • *Plant Sciences Unit, Fisheries and Food (ILVO), Flanders Research Institute for Agriculture, Melle, Belgium*

EDDI DI MARCO • *IRCCS Istituto Giannina Gaslini, Genoa, Italy*

KATHERINE DOMINY • *Imperial Molecular Pathology, Hammersmith Hospital, Imperial College Healthcare, London, UK*

- CHRISTIE FEKETE • *Thermo Fisher Scientific, South San Francisco, CA, USA*
- ALICE GOMES FERNANDES-MONTEIRO • *Laboratório de Biologia Molecular e Doenças Endêmicas (LABIMDOE), Fundação Oswaldo Cruz (Fiocruz)/Instituto Oswaldo Cruz (IOC), Manguinhos, RJ, Brazil*
- ELENA GENNARO • *IRCCS Istituto Giannina Gaslini, Genoa, Italy*
- LETIZIA GERACE • *Thermo Fisher Scientific, Life Sciences Solutions, Monza (MB), Italy*
- JESÚS GONZÁLEZ-LÓPEZ • *Department of Microbiology, Institute of Water Research, University of Granada, Granada, Spain*
- STEFANIE HERMANN • *Animal Physiology and Immunology, School of Life Sciences Weihenstephan, Technical University of Munich, Munich, Germany*
- ISABELE CRISTIANA ISER • *Laboratório de Biologia Celular, Departamento de Ciências Básicas da Saúde (DCBS), Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Porto Alegre, RS, Brazil*
- JAMSHID S. KHORASHAD • *Imperial Molecular Pathology, Hammersmith Hospital, Imperial College Healthcare, London, UK; Centre for Haematology, Hammersmith Hospital, Imperial College London, London, UK*
- BENEDIKT KIRCHNER • *Animal Physiology and Immunology, School of Life Sciences Weihenstephan, Technical University of Munich, Munich, Germany*
- MARION LAIG • *Thermo Fisher Scientific, South San Francisco, CA, USA*
- MAGALI LOSSCHAERT • *Plant Sciences Unit, Fisheries and Food (ILVO), Flanders Research Institute for Agriculture, Melle, Belgium*
- NICOLA LUCHI • *Institute for Sustainable Plant Protection, National Research Council (IPSP-CNR), Florence, Italy*
- NIVEDITA MAJUMDAR • *Thermo Fisher Scientific, South San Francisco, CA, USA*
- MAURO S. MALNATI • *Unit of Human Virology, Division of Immunology, Transplantation and Infectious Diseases, IRCCS Ospedale San Raffaele, Milan, Italy*
- PAULA MAZA-MÁRQUEZ • *Department of Microbiology, Institute of Water Research, University of Granada, Granada, Spain; Environmental Microbiology Group, Department of Microbiology, Faculty of Pharmacy, University of Granada, Granada, Spain*
- KATYA MOKRETAR • *Imperial Molecular Pathology, Hammersmith Hospital, Imperial College Healthcare, London, UK*
- REINHOLD MUELLER • *RM Consulting, San Diego, CA, USA*
- VERONIKA MUSSACK • *Animal Physiology and Immunology, School of Life Sciences Weihenstephan, Technical University of Munich, Munich, Germany*
- TANIA NOLAN • *Faculty of Medical and Human Sciences, Institute of Population Health, University of Manchester, Manchester, UK*
- LUIS G. C. PACHECO • *Post-Graduate Program in Biotechnology, Institute of Health Sciences, Federal University of Bahia, Salvador, BA, Brazil*
- MICHAEL W. PFAFFL • *Animal Physiology and Immunology, School of Life Sciences Weihenstephan, Technical University of Munich, Munich, Germany*
- ULRICH PFEFFER • *IRCCS Ospedale Policlinico San Martino, Genoa, Italy*
- PAMELA PINZANI • *Molecular and Clinical Biochemistry Unit, Department of Experimental and Clinical Biomedical Sciences “Mario Serio”, Careggi Hospital, University of Florence, Florence, Italy*
- ALESSANDRO RASO • *ASL3 Sistema Sanitario Regione Liguria, S.C. Laboratorio d’Analisi, Genoa, Italy*
- ALISTAIR G. REID • *Imperial Molecular Pathology, Hammersmith Hospital, Imperial College Healthcare, London, UK*

DANILO J. P. G. ROCHA • *Post-Graduate Program in Biotechnology, Institute of Health Sciences, Federal University of Bahia, Salvador, BA, Brazil*

BELÉN RODELAS • *Department of Microbiology, Institute of Water Research, University of Granada, Granada, Spain*

NICOLETTA SACCHI • *Ente Ospedaliero Ospedali Galliera, Genoa, Italy*

FRANCESCA SALVIANTI • *Molecular and Clinical Biochemistry Unit, Department of Experimental and Clinical Biomedical Sciences “Mario Serio”, Careggi Hospital, University of Florence, Florence, Italy*

ANA PAULA SANTIN BERTONI • *Laboratório de Biologia Celular, Departamento de Ciências Básicas da Saúde (DCBS), Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Porto Alegre, RS, Brazil*

ALBERTO SANTINI • *Institute for Sustainable Plant Protection, National Research Council (IPSP-CNR), Florence, Italy*

GEMMA SONNATI • *Department of Experimental and Clinical Biomedical Sciences “Mario Serio”, University of Florence, Florence, Italy*

PIERGIORGIO STEVANATO • *DAFNAE, Università Degli Studi di Padova, Legnaro (PD), Italy*

GISELA FREITAS TRINDADE • *Laboratório de Biologia Molecular e Doenças Endêmicas (LABIMDOE), Fundação Oswaldo Cruz (Fiocruz)/Instituto Oswaldo Cruz (IOC), Manguinhos, RJ, Brazil*

VALENTIN VAUTROT • *UMR 7365 IMoPA, Université de Lorraine, Vandoeuvre-lès-Nancy, France; UMR 7365 IMoPA, CNRS, Vandoeuvre-lès-Nancy, France*

MÁRCIA ROSANGELA WINK • *Laboratório de Biologia Celular, Departamento de Ciências Básicas da Saúde (DCBS), Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Porto Alegre, RS, Brazil*



Chapter 1

A Quarter Century of PCR-Applied Techniques and Their Still-Increasing Fields of Use

Alessandro Raso and Roberto Biassoni

Abstract

Quantitative polymerase chain reaction (PCR) is the basis of a variety of scientific applications and publications in a broad range of interests. It also plays a fundamental role in nucleic acid sequencing applications, including Next Generation Sequencing (NGS)-based ones. The potential of PCR diagnostics is enormous, particularly for the early diagnosis of life-threatening infections. Some other fields of applications that use PCR on a regular basis include oncology, genetics, microbiology, biochemistry, immunogenetics, NGS, ecology, comparative genome evolution, ancestry DNA, pharmacogenomics, personalized medicine, and even general medicine.

Key words quantitative Polymerase Chain Reaction (qPCR), High Resolution Melting (HRM), digital Polymerase Chain Reaction (dPCR), Dye-labeled probe, Intercalating dye

Polymerase chain reaction (PCR) is a simple-to-use technology that consists of adding selected polymerizing nucleotides called primers (the synthesis of which is now fairly inexpensive) based on the target sequence; a master mix containing DNA-polymerase enzymes, buffered salt solution, and magnesium solution; and finally the target DNA. Specific amplicons of the target sequence are then obtained after appropriate thermal cycling. In reality, however, this description is simplistic. For example, the primers may bind not specifically or even not begin the polymerization at all. In such a case, an experiment may produce meaningless data without any solid biological significance [1, 2].

From the chemical point of view, PCR is an enzymatic-based reaction that reaches dynamic equilibrium among reactants to result in a product. The primers and the probe (in quantitative PCR [qPCR]) need to hybridize with the target sequence in order to produce a specific amplicon, without the possibility of obtaining false amplifications or non-specific products—not only on the selected target sequence, but also on all genomes and transcriptomes. Thus, oligodeoxynucleotide primers and probes

(if needed) require a specific design based on the DNA or cDNA sequence, the thermodynamic properties of the hybridization reaction, and the template folding to reach an equilibrium to produce a final specific amplicon [3].

The length and the deoxynucleotide base assortment of the selected primers determine their melting temperature (at which 50% of them dissociate from the target sequence) and their thermodynamic behavior patterns, such as the concentration of use and their specificity. In addition, the neighbor sequence helps to determine the folding of the DNA region in which the primers reside, thus indicating which would be good primers [4]. Moreover, the reaction is governed by a decreasing amount of reactants and a complex mix of specific and secondary products, such as dimerized primers, as well as deterioration of the polymerase enzymatic properties, which change the efficiency of the reaction during the PCR reaction itself.

PCR occurs at a defined temperature, which is dependent on the primer design, in the presence of a sodium-buffered solution and appropriate magnesium salt concentration. All of these parameters, together with the thermal cycling conditions, need to be selected based on the thermostable polymerase in use.

Clearly, the preparation of a PCR system requires careful in silico design and extensive empirical optimization. Validation of the experiment has to fulfill the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines [5, 6]. Any analytical system that neglect such rules could negatively affect the interpretation of biological data and thus the results published in the literature, giving rise to weak and often misinterpreted data [2].

The thermodynamic and folding characteristics of the primers requires careful in silico design combined with validation and experimental optimization, which are fundamental to achieve the required specificity of the diagnostic system. In addition, the efficiency of a PCR reaction is directly dependent on good design, which is fundamental when precise quantitation is not only necessary but is a pre-requisite of a quantitative PCR reaction.

The master mix is another critical component of the assay because the enzymatic properties are fundamental achieving an amplification. In addition, even if the same primers are used on the same template with identical thermodynamic protocol, different reaction efficiencies may be achieved, which result in threshold cycle (C_t) differing by 4 cycles, among others. In many situations, this is ideal. However, for some critical determinations such as allele discrimination assays or in cases where specificity could be a problem, there may be a risk of false-positive identification when a master mix is more efficient than the one used originally to calibrate the assay. When applied correctly, all of the above-mentioned issues

helped researchers to approach quantitative PCR assays in the right way.

When qPCR or PCR are performed for clinical purposes, some methods are needed to control the different phases of the diagnostic procedure. In particular, all pre-analytical factors related to specimen handling and identification should be controlled. In addition, the analytic phase and analytical data interpretation should be checked carefully when a particular sample needs to be clearly linked to a diagnostic pathology report. Clinical diagnostic laboratories should be certified for all needed quality assessments, using all appropriate internal positive/negative and quantitative controls for each test performed. In addition, a certified diagnostic laboratory needs to routinely perform external quality assessment (EQA) assays for each type of determination used [7, 8]. Such external validation tests allow the laboratory to control all pre-analytical steps.

It is possible to obtain information on incorrect sample handling during preparation, such as the shipping of the biological material/reagents to the laboratory, storage before the analysis, the use of expired materials, or the use of reagent batches with manufacturing problems. A specific analytical phase can give information regarding instrument calibration or reagent defects, staff competency, incorrect analysis methods, and finally (but not less important), the use of biological material that could result in the matrix-dependent inhibition of the reaction. The latter issue is important because biological samples (e.g., blood, serum, urine, feces, swabs of different types, and other fluids) may contain contaminants that are known to deeply interfere with the assay if not properly handled. Furthermore, they may cause report misinterpretation and transcription errors, which is especially important for a high-throughput diagnostic laboratory.

Because these assays are based on running blind patient-like samples, it is possible to compare the obtained data with the results from different diagnostic laboratories in order to retrospectively monitor the accuracy of produced data reports. More importantly, they enable a comparison of the performance of procedures between different laboratories that use the same detection qPCR system but different DNA/RNA extraction instrumentation protocols. Thus, EQA provides an early warning for systematic errors in a diagnostic procedure, thus indicating areas for quality improvement. EQA also shows where personnel work properly, or where it is important to review staff training and check for reagent problems.

References

1. Raso A, Biassoni R (2014) Twenty years of qPCR: a mature technology? *Methods Mol Biol* 1160:1–3
2. Bustin S (2017) The continuing problem of poor transparency of reporting and use of inappropriate methods for RT-qPCR. *Biomol Detect Quantif* 12:7–9
3. Bustin SA (2005) Real-time, fluorescence-based quantitative PCR: a snapshot of current procedures and preferences. *Expert Rev Mol Diagn* 5:493–498
4. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31:3406–3415
5. Bustin SA, Wittwer CT (2017) MIQE: a step toward more robust and reproducible quantitative PCR. *Clin Chem* 63(9):1537–1538
6. Bustin SA, Benes V, Garson J, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley G, Wittwer CT, Schjerling P, Day PJ, Abreu M, Aguado B, Beaulieu JF, Beckers A, Bogaert S, Browne JA, Carrasco-Ramiro F, Ceelen L, Ciborowski K, Cornillie P, Coulon S, Cuypers A, De Brouwer S, De Ceuninck L, De Craene J, De Naeyer H, De Spiegelaere W, Deckers K, Dheedene A, Durinck K, Ferreira-Teixeira M, Fieuw A, Gallup JM, Gonzalo-Flores-S, Goossens K, Heindryckx F, Herring E, Hoenicka H, Icardi L, Jaggi R, Javad F, Karampalias M, Kibenge F, Kibenge M, Kumps C, Lambertz I, Lammens T, Markey A, Messiaen P, Mets E, Morais S, Mudarra-Rubio A, Nakiwala J, Nelis H, Olsvik PA, Pérez-Novo C, Plusquin M, Remans T, Rihani A, Rodrigues-Santos P, Rondou P, Sanders R, Schmidt-Bleek K, Skovgaard K, Smeets K, Tabera L, Toegel S, Van Acker T, Van den Broeck W, Van der Meulen J, Van Gele M, Van Peer G, Van Poucke M, Van Roy N, Vergult S, Wauman J, Tshuikina-Wiklander M, Willems E, Zaccara S, Zeka F, Vandesompele J (2013) The need for transparency and good practices in the qPCR literature. *Nat Methods* 10:1063–1067
7. Kristensen GB, Meijer P (2017) Interpretation of EQA results and EQA-based trouble shooting. *Biochem Med* 27:49–62
8. Badrick T, Punyalack W, Graham P (2018) Commutability and traceability in EQA programs. *Clin Biochem* 56:102–104



Chapter 2

Parameters for Successful PCR Primer Design

Stephen A. Bustin, Reinhold Mueller, and Tania Nolan

Abstract

Primers are critical components of any PCR assay, as they are the main determinants of its specificity, sensitivity, and robustness. Despite the publication of numerous guidelines, the actual design of many published assays is often unsound: primers lack the claimed specificity, they may have to compete with secondary structures at their binding sites, primer dimer formation may affect the assay's sensitivity or they may bind only within a narrow temperature range. This chapter provides simple guidance to avoid these most common issues.

Key words MIQE, Oligonucleotide primers, qPCR, Assay design

1 Introduction

The success of any PCR-based experiment depends, to a large extent, on the diligence and attention paid to the design of the primers. This process must pass through a series of defined procedures, conveniently separated into *in silico* and empirical, that is, wet lab steps [1–3]. *In silico* analysis of assay designs allows predictions with regards to primer characteristics and subsequent assay behavior, but in no way substitutes for empirical confirmation of those expectations. Hence an essential part of any PCR assay design is that every *in silico* projection must be followed by an extensive empirical optimization and validation [4]. This takes time, effort, and expense—but is an integral part of performing a PCR assay. The more thorough the empirical validation, the more likely it is that any subsequent results will be biologically relevant. Focus of this chapter is the wet lab workflow, which is summarized in Fig. 1 and involves identifying the processes that deliver the optimal primer combinations and assay conditions so as to generate assays that are reliable, accurate and reproducible. It is worth emphasizing that optimization data are mastermix-dependent; that is, primer design, concentration, and annealing conditions optimal for one mastermix may not be carried over if a different one is used [5].

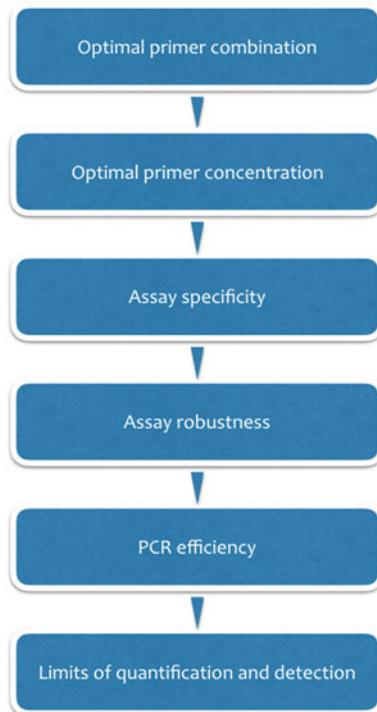


Fig. 1 Wetlab workflow for primer optimization

In general terms, most PCR primers are characterized by the following parameters:

- Primers must be sufficiently complex to minimize the likelihood of annealing to sequences other than the chosen target. There is a 1:256 chance of finding a specific four nucleotide sequence in any given DNA sequence. Hence, an 18-base sequence will statistically be present only once in every 6.8×10^{10} bases, around 20 times the size of the human genome. Each additional nucleotide makes a primer four times more specific, making the ideal length for a primer between 18 and 24 nucleotides. However, the longer an oligonucleotide, the more likely it is to form secondary structures that have a negative impact on PCR efficiency. Longer primers are only needed when amplifying targets with an expectation of target heterogeneity, when using mutagenic primers or when adding extra sequence information (e.g., RNA polymerase binding sites).
- Annealing temperatures (T_a) should be between 60 and 65 °C. However, some applications may require modifications to both primer length and T_a . For example, primers targeting AT-rich bacterial genomic DNA may need to be longer, those targeting GC-rich fungal genomic DNA may need to be shorter. As an alternative to elongating AT-rich sequences, modified bases may

be included (e.g., LNA to increase the effective T_a at the given location or along the primer) [6–8].

- While primer pairs should have a GC content of around 50%, this is not an unqualified requirement. Too high a GC content can lead to mispriming, since such primers anneal more stably with nontarget templates. Even if mispriming is only transient, it can be sufficient to initiate priming by the *Taq* (or other DNA) polymerase. If this happens during the first few cycles of the PCR reaction, nonspecific products will be amplified and may muddle the results. If PCR reactions target GC-rich targets, annealing times must be kept short [9].
- Primers should not have any significant secondary structures, since these can significantly affect PCR efficiency [10].
- There should be no complementarity between primers, especially at their 3'-ends. Runs of four or more of same nucleotide, especially G or C, should be avoided, but as long as they do not occur at the 3'-end, they will probably not affect primer performance.

We stress that these guidelines are generally appropriate. In practice primer sets that do not follow the design rules may result in the desired PCR amplicon and with an acceptable yield. Primer optimization changes an uncontrolled process into a user-controlled process.

2 Primer Combination

Oligonucleotides have become commodity items that are inexpensive to synthesize so that it is feasible to assess multiple forward and reverse primers for each assay. We recommend three of each, resulting in nine combinations of primers. Figure 2 shows the results of an experiment where three forward and three reverse primers targeting the *Clostridium difficile* tcdB gene (KC292190.1) were subjected to an initial screen evaluating sensitivity, as determined by the quantification cycle (C_q) and potential for primer dimer formation, as determined by amplification in the no template controls (NTCs). The results show that small differences in primer sequence can make a significant difference to the sensitivity of an assay, in this case 70-fold (high to low C_q difference of 6.19).

3 Primer Concentration

Although for many assays primer concentration can vary considerably without significant impact on PCR performance, optimization of primer concentration is important if the target is expected to be

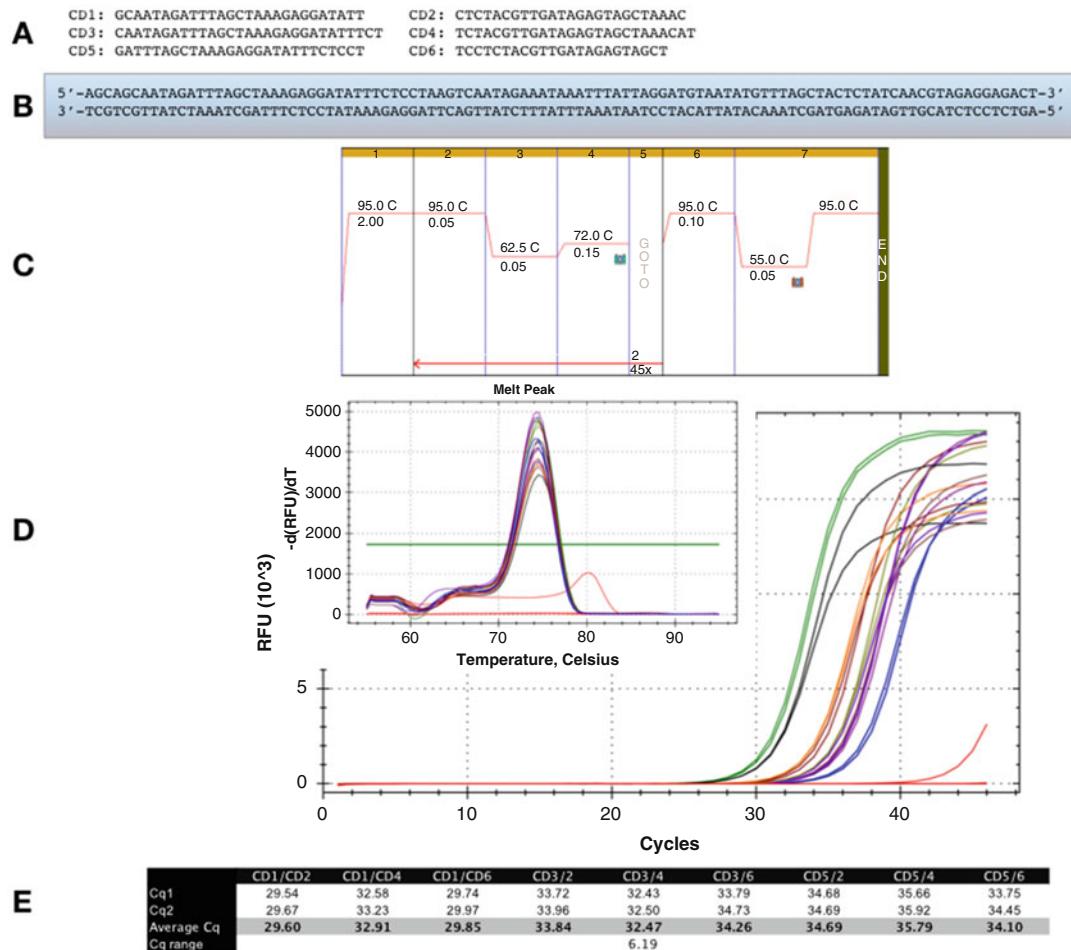


Fig. 2 Selection of primers targeting *C. difficile* tcdB gene. (a) Three forward (CD1, 3 and 5) and three reverse (CD2, 4 and 6) primers were designed using BeaconDesigner (PremierBiosoft). (b) All nine possible primer combinations were used to amplify in duplicate the same region of chromosomal DNA extracted from *C. difficile* 630 (CP010905.2), together with appropriate NTCs. (c) Amplification and melt curve conditions using Agilent's Brilliant III SYBR Green mastermix (catalogue no 600882). (d) Amplification with CD1/2 (green) and CD5/4 (blue) recorded the lowest and highest C_q s, respectively. Only one of the NTCs recorded a C_q at 45. E. The C_q range between the most and least sensitive primer combinations was 6.2. Hence the preferred primer pair is CD1/CD2. CD5/4 the least ideal primer set

of low copy number or to improve specificity. Having tested a large number of assays, we observe that the ΔC_q s between the worst and optimal primer concentrations have varied considerably, with around fivefold estimated to be a working average, although it can be considerably larger [11]. We have not seen a primer set that could not be optimised at least a little (twofold), but have seen several where optimization made a huge difference, improving sensitivity by more than 100-fold. Our advice is that it is worth carrying out this additional step as a routine part of validating any

new primer set, unless it is certain that the target will be present in vast amounts.

Nevertheless, although variations in primer concentration can have a significant influence on PCR results [11], primer concentration is not likely to be the main reason for the likely success or failure of a PCR experiment. This is demonstrated by the results shown in Fig. 3, where the difference in sensitivity is less than fourfold as long as primer concentration is kept in the usual range of around 200–400 nM.

Reducing primer concentration to below 100 nM is not advisable, as shown in Fig. 4a, where having one of the primers at 50 nM final concentration results in a significantly less sensitive assay. Interestingly, for this primer set the 100 nM final concentration does not result in a major reduction in sensitivity compared with 200 or 300 nM final concentration, in contrast to the example in Fig. 3. These results also demonstrate that choosing too low a primer concentration affects the linearity of the assay. Figure 4b shows the results obtained with the same primer concentrations when 1×10^5 times more template was used. The expected ΔC_q of 13.3 between the two experimental conditions is approximately achieved for all primer concentrations bar except the 50 nM ones, where the difference is much higher (Fig. 4d). Again, the conclusion is that results are very much primer dependent and if the aim is to combine sensitivity, with accurate quantification, then careful optimization of primer concentration will become important.

4 Primer Specificity

In most instances, the aim of primer design is to maximize the specificity of the PCR, which is determined by more or less predictable effects of numerous variables [12], an important one being the sequence at the 3'-end of the primers. Importantly, a PCR assay designed for specificity is more likely to be highly efficient over a wide dynamic range, since the assay will not generate nonspecific amplification products that will compete for PCR reagents or inhibit the main amplification reaction. There are, of course, instances when specificity is not a top priority, for example when the aim is to quantify closely related, but divergent pathogens, where special design, optimization, and validation criteria apply [13].

Melt curves are a standard way of evaluating the specificity of amplicons, at least with regards to whether a single target is being amplified [14]. However, it is important to emphasize that melt curves can be misleading, as, for example, they can be affected by a combination of suboptimal primers and low template concentration. Figure 5 shows melt curves obtained with samples where the template is present at different concentrations. Figure 5a reveals

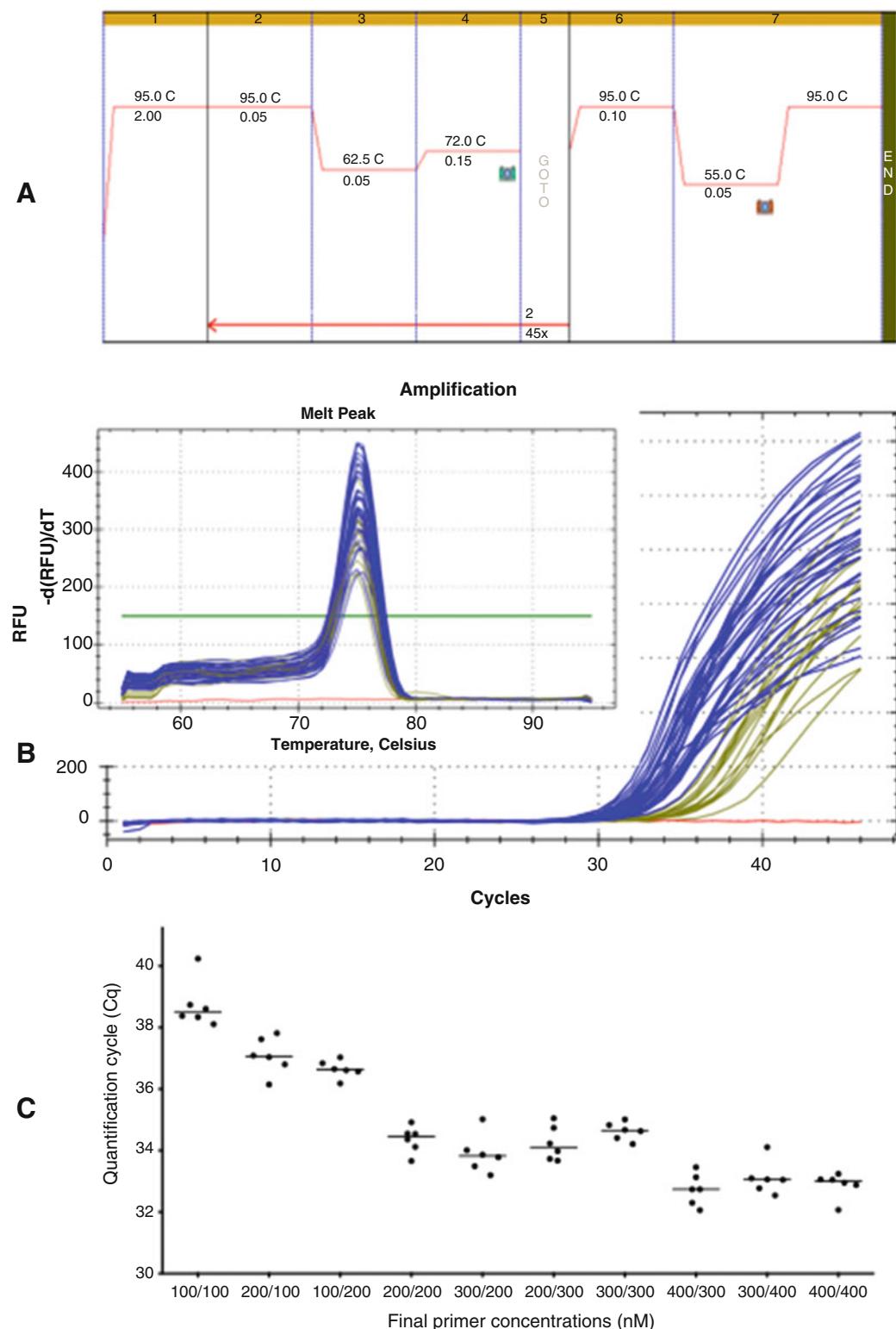


Fig. 3 Effect of primer concentration on qPCR assay. **(a)** Amplification and melt curve conditions using Agilent's Brilliant III SYBR Green mastermix (catalogue no 600882) and primers tcd-F: AGAGTTGGTA-GAAAGGTGGAATTAG and tcd-R: ACATACCACCAATTCTTTAATGC targeting the *C. difficile* tcdB toxin gene

that at the two lowest concentrations, nonspecific amplification products with lower T_m s than the specific amplicon are generated. Clearly, this assay cannot be used reliably to detect targets that are present at low concentration. Interestingly, the NTCs, that is, samples with no DNA present at all, record no (nonspecific) amplification products, suggesting that background genomic DNA can participate in nonspecific amplification/dimerization. Sometimes such background priming and nonspecific amplification cannot be remedied, but often it is possible to design assays that record no nonspecific amplification at any template concentration as well as in the NTC (Fig. 5b). Here even amplification of the target concentration recording a C_q of 35 results in a specific melt curve. Again, the NTCs show no sign of nonspecific amplification. Sometimes, assay behavior can be mastermix-dependent, with nonspecific amplification detected only with some buffer compositions, possibly associated with different Mg^{2+} concentrations.

5 Robustness

Optimization of the T_a is a useful step during the empirical validation and optimization of a qPCR assay. It provides an immediate indication of the robustness of a primer set by indicating the temperature (or temperature range) that generates the lowest C_q without amplifying the NTC. A two to fourfold difference in sensitivity might not matter to someone targeting a highly expressed mRNA, but for a diagnostic assay it can mean the difference between a positive and false negative result.

The T_a characteristics of qPCR primers can vary widely. Some assays are not very robust and fall apart very quickly if they are not performed at the optimal T_a of the primers. This is important because such assays tend to also be problematic in the real world, where samples may not be optimal in terms of purity, DNA concentration or in the presence of other DNA. Furthermore, target copy numbers might vary over a huge range and reagents, plastic ware or instruments may differ from the ones used to establish the assay. A typical result for an unsatisfactory assay is shown in Fig. 6, where the qPCR was carried out using a gradient of T_a s between 59 °C and 67 °C (Fig. 6a) using primers targeting three human

Fig. 3 (continued) (HM062503.1) from *C. difficile* 630 (CP010905.2). Primers slpA-F: ATTAGTTAT-TATTTTGTGTTCTAAATTAT and slpA-R: CCTGTTTTGCTGCAACTACT targeting the *C. difficile* slpA gene for s-layer protein (AB258978.1) were used to amplify chromosomal DNA extracted from *C. difficile* 630 (CP010905.2). (b) Amplification plots and melt curves. The green plots were obtained if either F or R primer were kept at 100 nM final concentration, indicating that at least for this primer set 100 nM is too low. (c) Plot of C_q vs primer concentration, demonstrating little difference between 200 and 400 nM final primer concentration

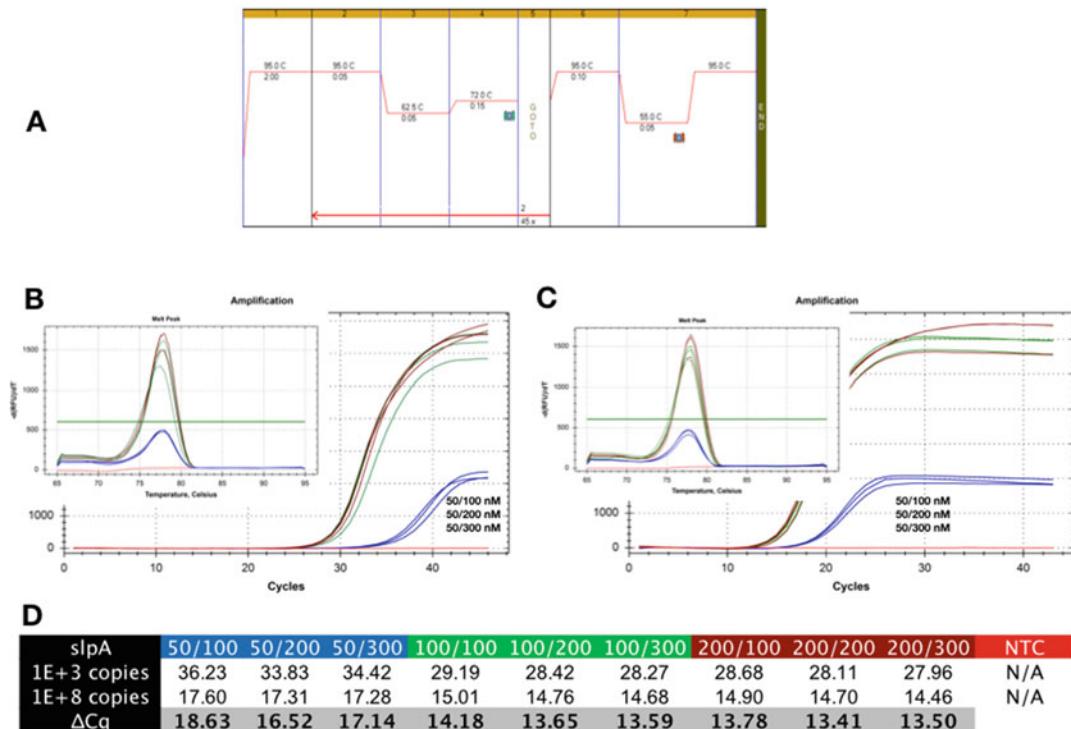


Fig. 4 Pronounced effect of primer concentration on qPCR assay. Primers slpA-F and slpA-R targeting the *C. difficile* slpA gene for s-layer protein (AB258978.1) were used to amplify chromosomal DNA extracted from *C. difficile* 630 (CP010905.2). **(a)** Amplification and melt curve conditions using Agilent's Brilliant III SYBR Green mastermix (catalogue no 600882). **(b)** Amplification plots and melt curves (inset) obtained using 1×10^3 target copies. The three blue amplification plots were obtained with a final concentration of 50 nM F primer, green and brown plots were obtained with final concentrations of 100 and 200 nM F primer, respectively. **(c)** Amplification plots and melt curves (inset) obtained using 1×10^8 target copies. **(d)** ΔC_q s between different target concentrations amplified with primers at a range of concentrations

brain-specific genes [15]. It is apparent from the amplification plot that the Opalin primers are far from ideal, as they have a very narrow optimal T_a range (Fig. 6b), that is, the C_q s are widely spread out, resulting in significantly lower C_q s compared to their optimal C_q . This assay is not robust, is likely to result in sub-optimal amplification and consequently, this primer pair should be redesigned. Furthermore, melt curve analysis (inset) suggests that there may also be problems with the specificity of this assay, as the melt curves are not the same at each T_a . The ACSBG1 assay shown in Fig. 6c is more robust than the Opalin assay above but is still far from ideal and likely could have been improved on. However, we stress that robustness and specificity are not necessarily linked, since the melt curves generated by this assay show the same peak at all T_a s (inset). Robust assays, on the other hand, are very much more tolerant and result in similar C_q s over a wide range of T_a s, as seen for the GFAP assay shown in Fig. 6d. The C_q s obtained over the

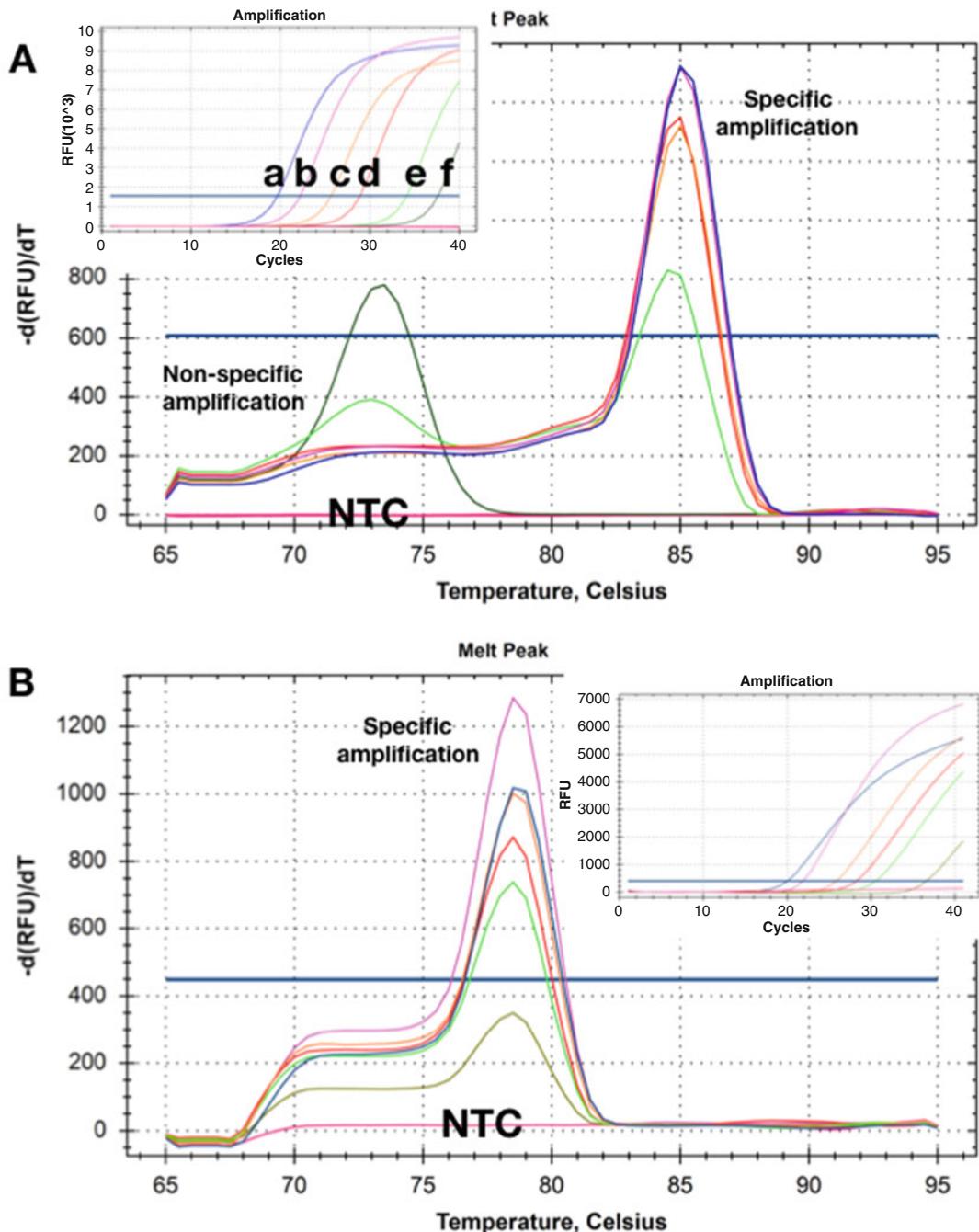


Fig. 5 Melt curves showing T_m shift in melt curves obtained from two assays targeting different amounts of two target DNA. **(a)** At the higher concentrations (*a-d*)), no primer dimer is apparent after the qPCR assay has been completed. As template concentration is reduced to 50 copies (*e*), a nonspecific product start to appear, which becomes the only product at the lowest concentration (*f*). **(b)** This assay records the same T_m s at all target concentrations, with no primer dimer apparent even at the lowest concentration (5 copies). No amplification products are detected in the NTCs when using either assay

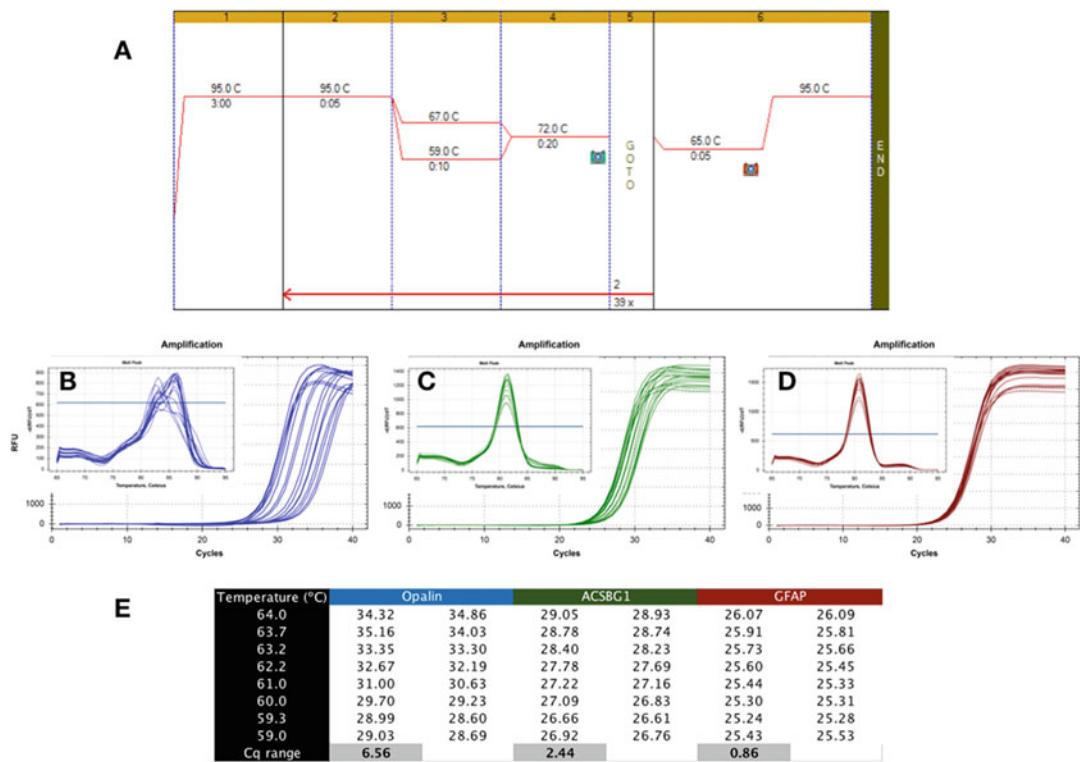


Fig. 6 Temperature gradients reveal differing robustness of PCR assays. (a) The protocol used to subject cDNA prepared from human brain RNA to PCRs using Bioline's Sensifast SYBR mastermix (catalogue no BIO-98050). (b) Amplification plots and melt curves recorded by Bio-Rad's CFX qPCR instrument obtained for opalin (NM_033207, F: GCCATGGAGGAAAGTGACAGACC, R: CTCATGTGTGGGTGATCTCCTAGG). (c) Amplification plots and melt curves obtained for ACSBG1 (NM_015162.4, F: CTACACTTCCGGCACCACTGG, R: GTCCACGTGATTGTCTTGACTCAG). (d) Amplification plots and melt curves obtained for GFAP (NM_002055.5, F: TGGAAGGAAATTGAGTCGCTGG, R: CGAACCTCCTCCTCGTGGATCTTC). (e) C_q s recorded at the different annealing temperature, showing the differences in C_q recorded across the 7 °C temperature gradient

same 8 °C range differ by less than 1 and the melt curves (inset) confirm the assay specificities over that temperature range. It is worth noting that calculated T_{as} and the actual T_a range can be substantially different [5].

There are numerous guidelines that aim to assist researchers in designing efficient primers, mostly based on long established rules [16], and a lot of attention is paid to the 3'-end of the primer. The advice frequently given is to include a G or C at the 3'-end, and to have two G or C bases (GC clamp) but no more than two in last 5 bases. In practice, these rules can guide the researcher, but are not necessarily true in every situation. For example, the results in Fig. 7 contradict the 3'-end rules, with all designs generating results that are essentially identical, with the exception of two primer combination resulting in nonspecific amplification in the NTC. However, we cannot support the effects of a GC-clamp as, in this case, having

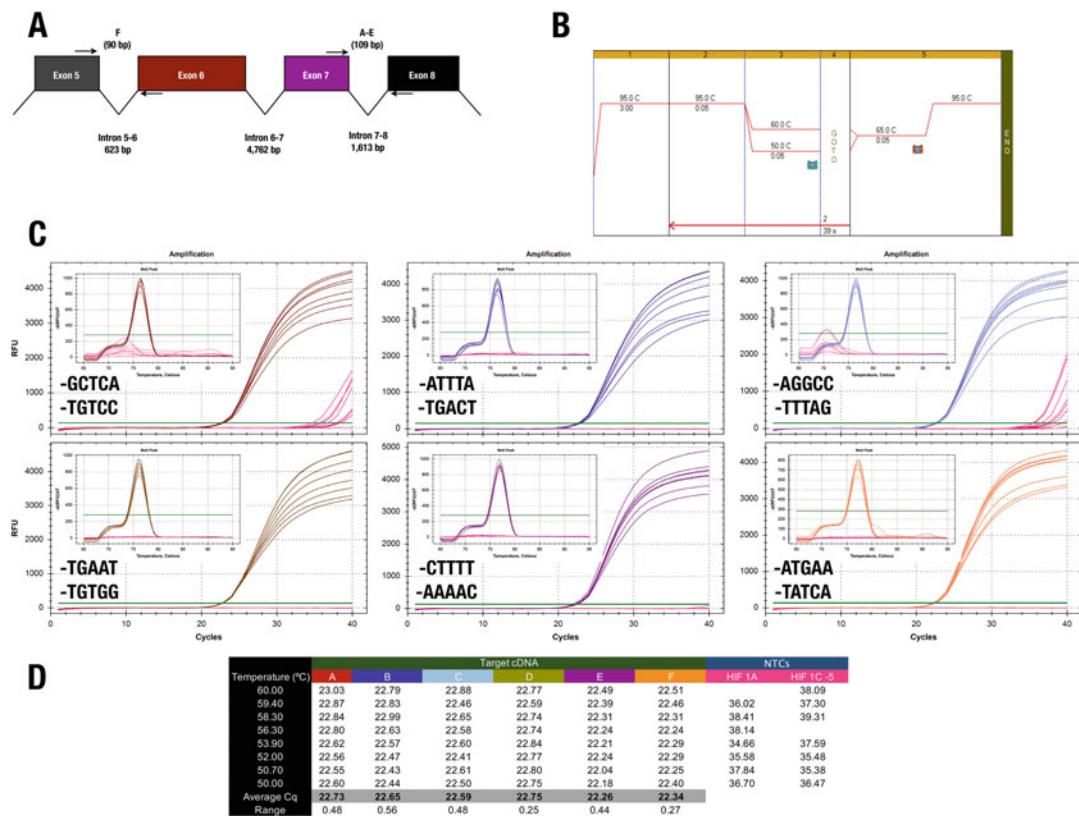


Fig. 7 Minimal effect of primer 3'-end on specificity or efficiency. **(a)** Location of primers targeting the human HIF-1 α (NM_181054.2) gene. **(b)** Protocol used to amplify the six assays using Agilent's Brilliant III SYBR Green mastermix (catalogue no 600882). **(c)** Amplification plots and melt curves recorded by Bio-Rad's CFX qPCR instrument and 3'-ends of primers. NTCs are shown in red. **(d)** C_qs recorded for each assay

A or T as the most 3'-base does not reduce specificity. Assay C, where the F primer ends with GGCC, does record C_qs in the NTCs, suggesting that one might want to avoid these sequences at the 3'-ends. We emphasize that the only way to determine the optimal 3'-end sequence of primer pairs is to perform an empirical evaluation of a number of candidate primers.

6 Efficiency

Importantly, while a nonspecific PCR assay can never become specific, amplification efficiency can be tweaked and maximized in many different ways by changing enzymes, master mixes, additives and cycling conditions. The efficiency of a PCR assay is best assessed using tenfold or fivefold serial dilutions of the target nucleic acid, that is, the “Standard Curve Method.” If PCR amplicons or synthetic DNA targets are used for generating standard

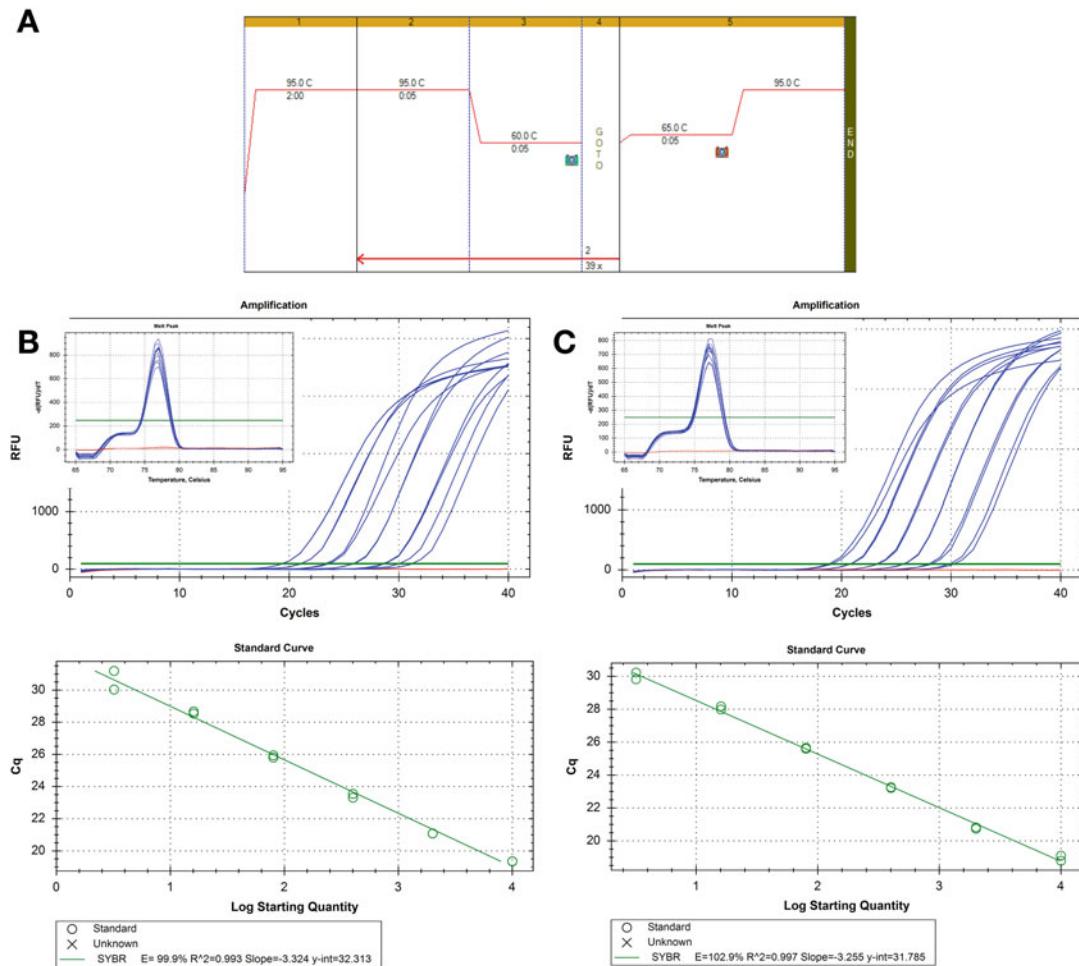


Fig. 8 Dilution curves assessing the efficiency of a PCR.. (a) PCR and melt curve condition using primers targeting HIF-1: F: AAGAACTTTAGGCCGCTCA and R: TGTCCCTGTGGTGACTTGTCC and Agilent's Brilliant III SYBR Green mastermix (catalogue no 600882). (b) 100 ng of RNA was reverse transcribed, diluted twofold and the serially diluted cDNA sample was diluted fivefold into 1 ng human genomic DNA. Melt curves are shown in the insert. (c) The RT reaction, dilution and serial dilutions were repeated for a second cDNA sample with a similar result

curves, serial dilutions of these targets should be mixed with a constant amount of background DNA (e.g., genomic DNA). Figure 8 shows two standard curves, using the same assay on two different cDNA samples results in comparable efficiencies of around 100% and similar R^2 values, that is, how well the experimental data fit the regression line or how linear the data are.

The two standard curves are comparable, but not identical and if the aim to accurately quantify targets, then it is important to note that it is not acceptable to provide copy number calculations without any indication of uncertainty. Figure 9 shows, for an optimised

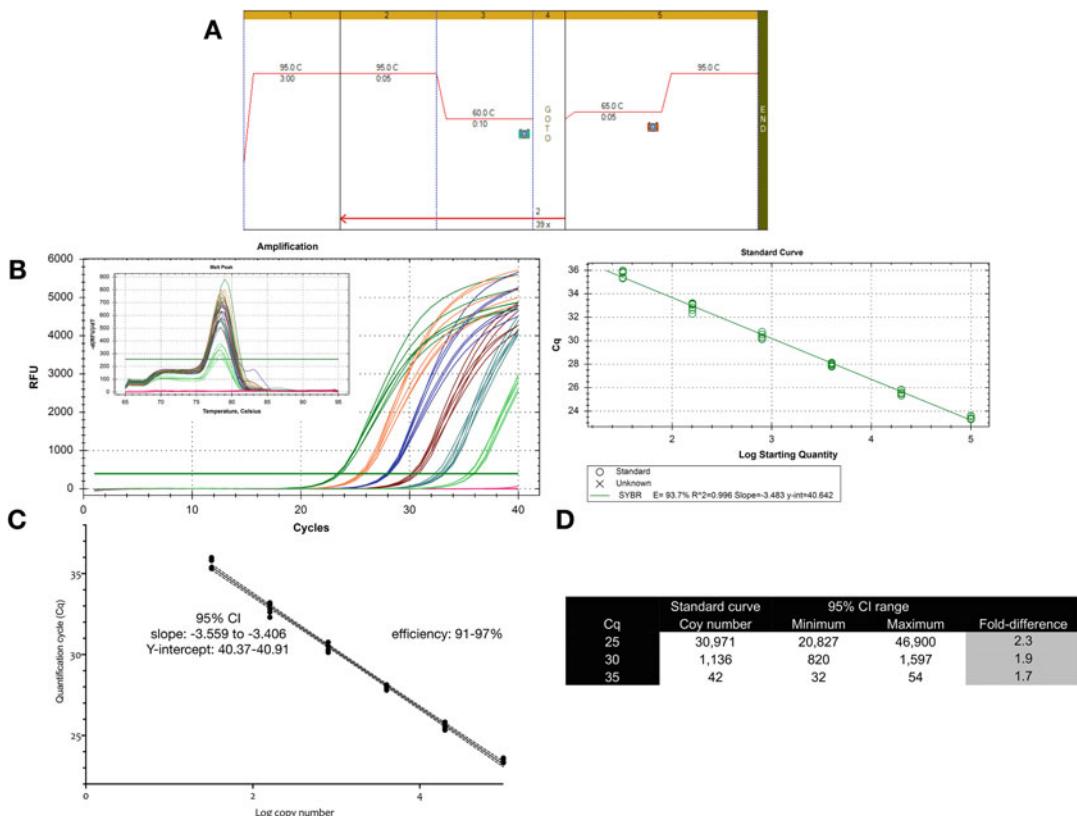


Fig. 9 Measurement uncertainty associated with quantification using a standard curve. (a) PCR and melt curve condition using primers targeting GAPDH (NM_002046): F: ACAGTTGCCATGTAGACC and R: TAACTGGTTGAG-CACAGG and Bioline's Sensifast SYBR mastermix (catalogue no BIO-98050). (b) Amplification plot, melt curves and standard curve recorded by Bio-Rad's CFX qPCR instrument. (c) Standard curve plotted with 95% Confidence interval (CI). (d) Copy numbers and 95% CI of three C_q values derived from the dilution curve in c

assay, the variability inherent in a single standard curve, which is about twofold (95% Confidence Interval, minimum to maximum), which is probably the minimum variability that can be expected.

7 Limits of Quantification and Detection

Establishing the efficiency of a primer set also allows the use of linear regression analysis to evaluate the assay's linear range, that is, the range of target concentrations over which C_q s are directly proportional to the concentration. This is essential for quantification and only samples that fall within the practical limits of assay should be quantified, although they can of course be reported as "detected." Limits of quantification (loq) is the lowest copy number that reports a predefined variability, often recorded as the coefficient of variation (CV), although we feel that a measure of

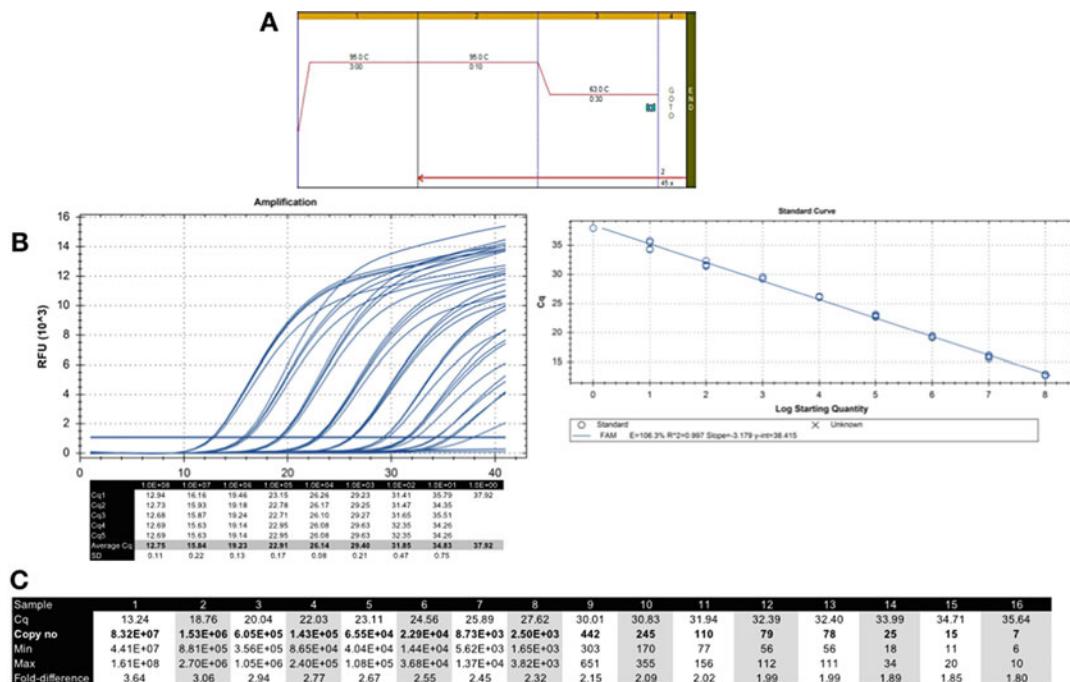


Fig. 10 Quantification of samples with unknown copy number, using a standard curve with F primer GTTTGGTGTGAGCAATACGAC and R primer CTACCTGATTGAGGTCAAAGTTG targeting the *Candida albicans* 18S rRNA gene using Agilent's Brilliant III SYBR Green mastermix (catalogue no 600882). (a) PCR conditions. (b) C_{qs}, amplification plot and standard curve recorded by Bio-Rad's CFX qPCR instrument. (c) Copy numbers derived from the dilution curve in (b). Also shown are the uncertainties associated with these copy numbers, as calculated from the 95% CI and expressed as fold-differences. Statistical analyses were carried out using Prism 8 for Mac

fold-difference, probably two to threefold, would be more appropriate for qPCR. Loqs are always higher than limits of detection, which refers to the lowest target copy number that can be reliably detected. Theoretically, the limit of detection is one, as even a single copy of a target DNA can be amplified by PCR and generate a signal above the fluorescence noise. However, at very low analyte concentrations, the solution composition is not homogeneous. The sampling error (Poisson error) determines the lowest concentration of target DNA that can be detected. In practice, 10 copies of target DNA is probably the lowest concentration that can be reliably amplified each time using optimized qPCR.

A typical example of a standard curve and the copy numbers calculated for a series of samples of unknown concentration that fall within its linear range and can therefore be quantified, in theory, as shown in Fig. 10. These data reaffirm the concept that the inherent variability within a single experiment is at least twofold, more if several replicates are run. If independent extractions of the same sample are compared, that variability can be considerably higher. It

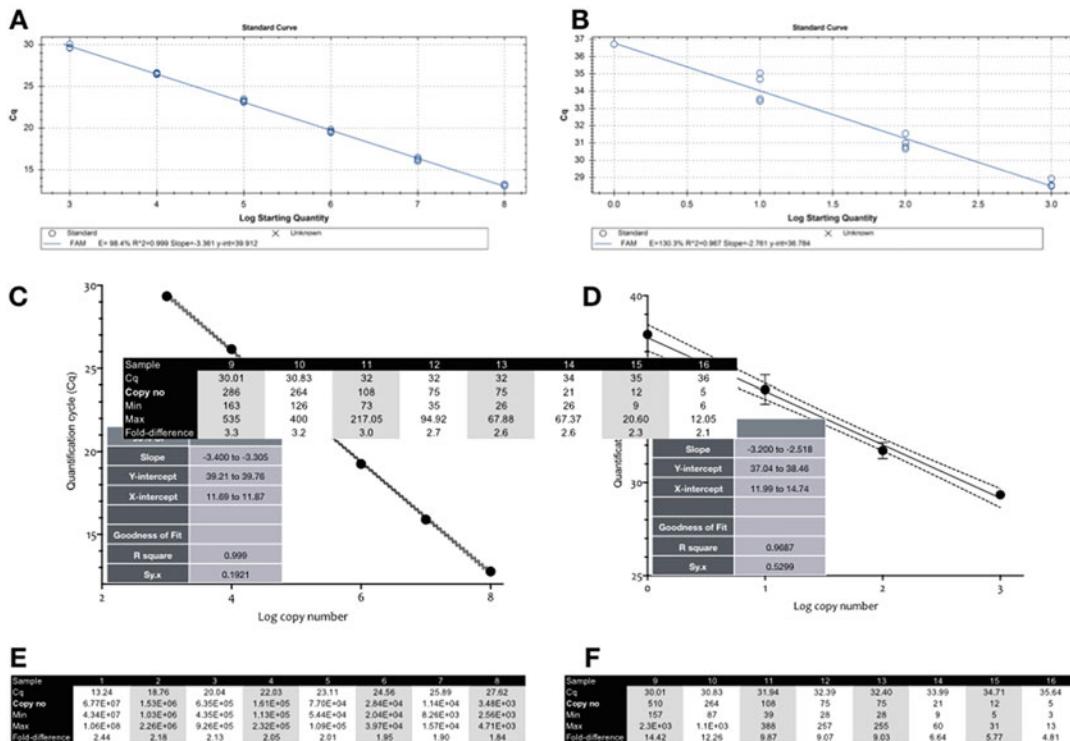
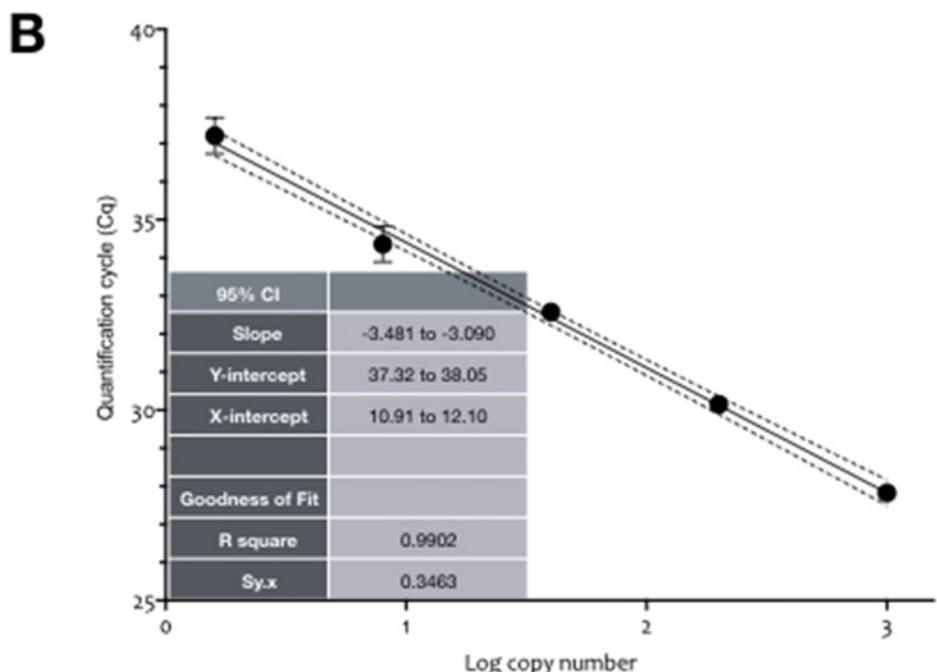
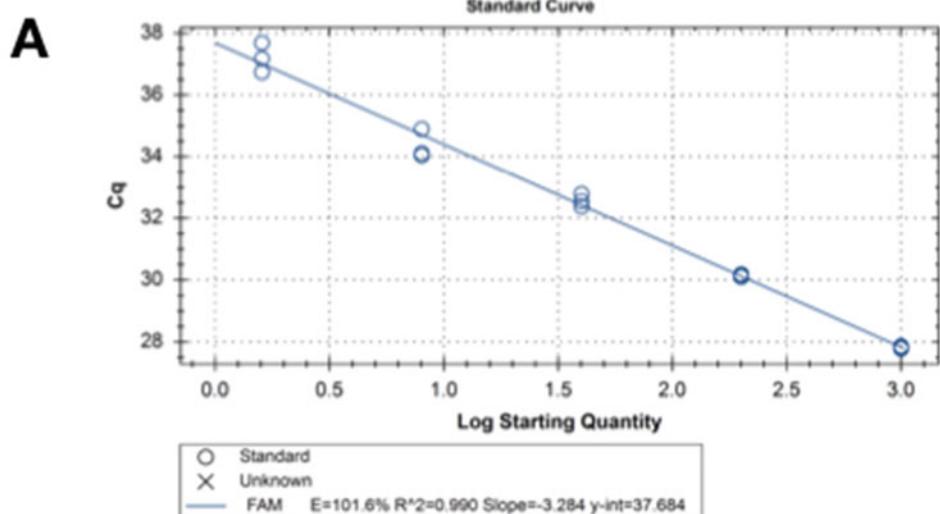


Fig. 11 Determination of LOD and LOQ. (a) Standard curve for the highest copy number standards (1×10^8 – 1×10^3). (b) curve obtained for the highest copy number standards (1×10^3 –1). (c) Standard curve from a replotted with 95% CI. (d) Standard curve from b replotted with 95% CI. (e) Copy numbers obtained from the standard curve in a. (f) Copy numbers obtained from the standard curve in a

is important to bear this in mind when reviewing publications that claim significance for smaller fold-changes.

Another point to bear in mind is that while it might look good to be able to demonstrate linearity across eight or more orders of magnitude, the use of such a standard curve masks true variability, especially when low copy number targets are being quantified. This is demonstrated if instead of plotting the standard curve shown in Fig. 10, the lower and higher dilutions are plotted separately and used to quantify samples within their respective ranges. Figure 11a, b shows the standard curves plotted by the instrument, with single values for PCR efficiencies of 98.4% and 130.3%, respectively. If the data are replotted to include the 95% Confidence intervals (Fig. 11c, d) and then used to calculate target copy numbers (Fig. 11e, f), it becomes apparent that quantification based on the standard curve obtained for higher copy numbers gives similar, or indeed better results to the data based on the standard curve shown in Fig. 10. This is because both copy numbers and the uncertainties associated with them are quite different for the lower copy number targets due to the fact that efficiency and r2 values are worse (98% vs



C

Sample	9	10	11	12	13	14	15	16
C _q	30.01	30.83	31.94	32.39	32.40	33.99	34.71	35.64
Copy no	286	264	108	75	75	21	12	5
Min	163	126	73	35	26	26	9	6
Max	535	400	217	95	68	67	21	12
Fold-difference	3.3	3.2	3.0	2.7	2.6	2.6	2.3	2.1

Fig. 12 Determination of LOQ for low copy number targets. **(a)** Fresh standards were prepared by pipetting PCR amplicons obtained by amplifying Candida DNA with the primers described in Fig. 10 into 100 ng of human chromosomal DNA and carrying out serial fivefold dilutions. **(b)** Standard curve from **a** replotted with 95% CI. **(d)** Standard curve from **b** replotted with 95% CI. **(c)** Copy numbers obtained from the standard curve, recording the fold-differences within the 95% CI

130% and 0.999 vs 0.967, respectively). Clearly, quantification of copy number below 3000 or so requires additional optimization, which at its simplest, involves more accurate dilutions and pipetting. Figure 12 shows a standard curves optimised for high copy number targets, with DNA samples diluted into 100 ng of chromosomal DNA and subjected to fivefold dilution. Both efficiency (101%) and r² value (0.990) are now comparable to the higher copy number standard curve shown in Fig. 11 and the fold difference for the samples recording higher C_q values are now around two to threefold, that is, within the loq range.

8 Conclusions

Knowledgeable and coherent primer design is at the heart of any project designed to quantify nucleic acids. Each new design must be validated in silico by dismissing primers with suspected secondary structures, such as hairpins and primer-dimers, and nonspecific binding that can reduce specificity of amplification and amplicon yield, that is, sensitivity of the assay. Furthermore, it is important to avoid SNP sites within the primer binding regions that could interfere with the efficient annealing of the primer or prevent amplification altogether. Once primer sets have been designed, they need to be verified by experimentation for specificity and optimized for amplicon yield. Finally, linearity of the optimized primer combination should be determined, which allows for determination of sensitivity and inherent variability of the assay. Finally, results of optimization and validation processes should be reported when qPCR data are published, in accordance with the MIQE guidelines [1, 3, 17, 18].

References

1. Bustin SA, Benes V, Garson JA et al (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55:611–622. <https://doi.org/10.1373/clinchem.2008.112797>
2. Taylor S, Wakem M, Dijkman G et al (2010) A practical approach to RT-qPCR-publishing data that conform to the MIQE guidelines. *Methods* 50:S1–S5. <https://doi.org/10.1016/j.ymeth.2010.01.005>
3. Huggett JF, Foy CA, Benes V et al (2013) The digital MIQE guidelines: minimum information for publication of quantitative digital PCR experiments. *Clin Chem* 59:892–902. <https://doi.org/10.1373/clinchem.2013.206375>
4. Derveaux S, Vandesompele J, Hellemans J (2010) How to do successful gene expression analysis using real-time PCR. *Methods* 50:227–230. <https://doi.org/10.1016/j.ymeth.2009.11.001>
5. Bustin S, Huggett J (2017) qPCR primer design revisited. *Biomol Detect Quantif* 14:19–28. <https://doi.org/10.1016/j.bdq.2017.11.001>
6. Latorra D, Arar K, Hurley JM (2003) Design considerations and effects of LNA in PCR primers. *Mol Cell Probes* 17:253–259
7. Fratzczak A, Kierzek R, Kierzek E (2009) LNA-modified primers drastically improve hybridization to target RNA and reverse transcription. *Biochemistry* 48:514–516. <https://doi.org/10.1021/bi8021069>

8. Malgoyre A, Banzet S, Mouret C et al (2007) Quantification of low-expressed mRNA using 5' LNA-containing real-time PCR primers. *Biochem Biophys Res Commun* 354:246–252. <https://doi.org/10.1016/j.bbrc.2006.12.194>
9. Mamédov TG, Pienaar E, Whitney SE et al (2008) A fundamental study of the PCR amplification of GC-rich DNA templates. *Comput Biol Chem* 32:452–457. <https://doi.org/10.1016/j.compbiochem.2008.07.021>
10. Singh VK, Govindarajan R, Naik S et al (2000) The effect of hairpin structure on PCR amplification efficiency. *Mol Biol Today* 1:67–69
11. Nolan T, Hands RE, Bustin SA (2006) Quantification of mRNA using real-time RT-PCR. *Nat Protoc* 1:1559–1582. <https://doi.org/10.1038/nprot.2006.236>
12. Ruiz-Villalba A, van Pelt-Verkuil E, Gunst QD et al (2017) Amplification of nonspecific products in quantitative polymerase chain reactions (qPCR). *Biomol Detect Quantif* 14:7–18. <https://doi.org/10.1016/j.bdq.2017.10.001>
13. Linhart C, Shamir R (2005) The degenerate primer design problem: theory and applications. *J Comput Biol* 12:431–456. <https://doi.org/10.1089/cmb.2005.12.431>
14. Morrison TB, Weis JJ, Wittwer CT (1998) Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *BioTechniques* 24:954–962
15. van den Berge M, Sijen T (2017) Extended specificity studies of mRNA assays used to infer human organ tissues and body fluids. *Electrophoresis* 38:3155–3160. <https://doi.org/10.1002/elps.201700241>
16. Dieffenbach CW, Lowe TM, Dveksler GS (1993) General concepts for PCR primer design. *PCR Methods Appl* 3:S30–S37
17. Bustin SA, Beaulieu JF, Huggett J et al (2010) MIQE précis: practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments. *BMC Mol Biol* 11:74. <https://doi.org/10.1186/1471-2199-11-74>
18. Bustin SA, Benes V, Garson JA et al (2011) Primer sequence disclosure: a clarification of the MIQE guidelines. *Clin Chem* 57:919–921. <https://doi.org/10.1373/clinchem.2011.162958>



Chapter 3

MIQE-Compliant Validation of MicroRNA Biomarker Signatures Established by Small RNA Sequencing

Veronika Mussack, Stefanie Hermann, Dominik Buschmann, Benedikt Kirchner, and Michael W. Pfaffl

Abstract

MicroRNAs (miRNAs), a class of small non-coding RNAs that modulate gene expression at the post-transcriptional level, are attractive targets in many academic and diagnostic applications. Among them, assessing miRNA biomarkers in minimally invasive liquid biopsies was shown to be a promising tool for managing diseases, particularly cancer. The initial screening of disease-relevant transcripts is often performed by high-throughput next-generation sequencing (NGS), in here RNA sequencing (RNA-Seq). After complex processing of small RNA-Seq data, differential gene expression analysis is performed to evaluate miRNA biomarker signatures. To ensure experimental validity, biomarker candidates are commonly validated by an orthogonal technology such as reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR). This chapter outlines in detail the material and methods one can apply to reproducibly identify miRNA biomarker signatures from blood total RNA. After screening miRNA profiles by small RNA-Seq, resulting data is validated in compliance with the “Minimum Information for Publication of Quantitative Real-Time PCR Experiments” (MIQE) guidelines.

Key words microRNA, Small RNA sequencing, Biomarker, MIQE, RT-qPCR, NGS, RNA-Seq, Normalization, Validation, Standardization

1 Introduction

In recent years, next-generation sequencing (NGS) technology, in particular RNA sequencing (RNA-Seq), has proven its value as a powerful tool for transcriptomics, gene expression profiling and RNA based biomarker discovery [1]. Since this highly sensitive method has the power to detect genes with huge sample and target throughput, NGS is frequently utilized for analytical biomarker applications.

Biomarkers are well-characterized medical indicators, which can be measured in an accurate and reproducible manner and provide diagnostic, prognostic, or predictive utility. With the ability to objectively facilitate early diagnosis and prognosis of relevant

clinical outcomes across a variety of treatments and populations, and the potential to monitor health status and therapy success, transcriptional biomarkers are particularly suited for a multitude of clinical technologies and applications [2, 3]. The evaluation of the abovementioned approaches is based on differential expression levels of disease-relevant transcripts or combinations thereof.

Small non-coding RNAs, such as microRNAs (miRNAs), have post-transcriptional regulatory functions and therefore significant impact on physiological and pathological conditions [4]. Considering their stability and their easy and minimally invasive accessibility by liquid biopsies from a wide range of biofluids [5, 6], circulating miRNAs have been evolving as a favored source material for biomarker studies [7].

Transcriptomic biomarkers, previously identified by RNA-Seq methods, should be additionally validated in order to assume statistical consistency using a well-established and standardized method [1]. For this particular purpose, a highly sensitive and reliable reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR) approach is often applied [7]. RT-qPCR experiments include a cDNA synthesis via reverse transcription, cDNA amplification by PCR, and the detection and quantification of PCR products in real-time [8]. Using molecular probes, especially double-stranded DNA-binding dyes such as SYBR Green I, the formation of PCR products during each cycle can be monitored continuously based on fluorescence signals [9], which are directly proportional to the amount of newly generated template [10].

For the analysis of real-time PCR data, both absolute and relative quantification strategies are commonly used. While absolute quantification determines the copy number of input miRNAs by relating the PCR signal to a standard curve, relative quantification assesses the expression change of a target miRNA in relation to a non-regulated stable miRNA and further to a control group [11]. The latter method is usually used for biomarker applications, where relative changes in miRNA gene expression levels are more relevant than absolute transcript copy numbers.

To improve the evaluation of qPCR experiments and foster quality control, standardization and comparability of studies, a set of guidelines the “Minimum Information for Publication of Quantitative Real-Time PCR Experiments” (called MIQE guidelines), has been published [12]. Working according to the included checklist helps to ensure that results of RT-qPCR experiments are more reliable, consistent, comparable, and reproducible, in particular. Regarding transcriptional biomarkers, it is of great importance that comparable results are obtained across different clinics and laboratories in order to ensure the validity of evaluated biomarker signatures. Therefore, the MIQE guidelines have been considered in the protocol detailed below.

This chapter outlines the entire experimental procedure as well as data analysis process, ranging from sample handling to library preparation, small RNA-Seq, and RT-qPCR validation, required to reproducibly identify and validate miRNA biomarker signatures from blood total RNA.

2 Materials

For an MIQE-compliant documentation, manufacturers and catalogue numbers of all reagents and equipment are demanded for the items on the checklist, which should be reported. For publication purposes, it is also recommended to provide lot information for commercially available kits.

2.1 Sample Preparation

2.1.1 Whole Blood Sampling

1. PAXgene Blood RNA tubes (Cat.no. 762165 (BD), PreAnalytiX GmbH, Hombrechtikon, Switzerland).
2. Safety-Multifly-Needles, 20G (Cat.no. 85.1637.235, Sarstedt AG & Co, Nümbrecht, Germany).
3. Longneck blood culture adapter (Cat.no. 14.1207, Sarstedt AG & Co, Nümbrecht, Germany).
4. Tourniquet (Cat.no. 900-6441, Henry Schein Medical GmbH, Berlin, Germany).
5. Dry sterile gauze (Cat.no. V3-000, Medicounter GmbH, Norderstedt, Germany).
6. Alcohol swab for puncture site (Cat.no. 9803640, PAUL HARTMANN AG, Heidenheim, Germany).

2.1.2 RNA Isolation

1. Ethanol 100%, molecular biology grade (Cat.no. 10644795, Fisher Scientific GmbH, Schwerte, Germany).
2. Isopropanol 100% (Cat.no. 437423R, VWR International GmbH, Darmstadt, Germany).
3. PAXgene Blood RNA Kit (Cat.no. 762164, PreAnalytiX GmbH, Hombrechtikon, Switzerland).

2.1.3 RNA Quantification and Quality Control

1. Agilent 2100 Bioanalyzer with the RNA 6000 Nano Kit (Cat. no. G2939BA and 5067-1511, Agilent Technologies, Santa Clara, USA).
2. NanoDrop ND-1000 (Thermo Fisher Scientific Inc., Waltham, USA).

2.2 Library Preparation and Small RNA Sequencing

2.2.1 Library Preparation

1. NEBNext Multiplex Small RNA Library Prep Kit for Illumina (Cat.no. E7560S, New England Biolabs, Inc., Ipswich, USA).
2. MetaPhor Agarose (Cat.no. 50181, Lonza Group AG, Basel, Switzerland).
3. O'RangeRuler 20 bp DNA Ladder (Cat.no. SM1323, Thermo Fisher Scientific Inc., Waltham, USA).
4. GeneRuler Ultra Low Range DNA Ladder (Cat.no. SM1313, Thermo Fisher Scientific Inc., Waltham, USA).
5. Monarch PCR & DNA Cleanup Kit (Cat.no. T1030S, New England Biolabs Inc., Ipswich, USA).
6. Monarch DNA Gel Extraction Kit (Cat.no. T1020S, New England Biolabs, Inc., Ipswich, USA).
7. Agilent 2100 Bioanalyzer with the DNA 1000 and High Sensitivity DNA Kits (Cat.no. G2939BA, 5067-1504 and 5067-4626, Agilent Technologies, Santa Clara, USA).
8. 2100 Bioanalyzer Expert Software, version B.02.09.SI725 (Agilent Technologies, Santa Clara, USA).

2.2.2 Small RNA-Seq

1. HiSeq SR Rapid Cluster Kit v2 (Cat.no. GD-402-4002, Illumina Inc., San Diego, USA).
2. HiSeq Rapid SBS Kit v2, 50 cycles (Cat.no. FC-402-4022, Illumina Inc., San Diego, USA).
3. HiSeq 2500 Sequencing System (Cat.no. SY-401-2501, Illumina Inc., San Diego, USA).
4. HiSeq Control Software Package, version 2.2.68 (Illumina Inc., San Diego, USA).

2.3 Validation of Sequencing Results via RT-qPCR

2.3.1 Reverse Transcription (RT)

2.3.2 qPCR

1. miRCURY LNA RT Kit (Cat.no. 339340, QIAGEN GmbH, Hilden, Germany).
2. Mastercycler gradient (Cat.no. 5331, Eppendorf AG, Hamburg, Germany).
1. miRCURY LNA SYBR Green PCR Kit (Cat.no. 339347, QIAGEN GmbH, Hilden, Germany).
2. miRCURY LNA miRNA PCR Assays (Cat.no. 339306, QIAGEN GmbH, Hilden, Germany) (*see Table 1*).
3. FrameStar 384-well PCR plates, white wells, clear frame (Cat. no. 4ti-0381, 4titude Ltd., Berlin, Germany).
4. Adhesive sealing foils for PCR plates (Cat.no. 4ti-0560, 4titude Ltd., Berlin, Germany).

Table 1
PCR assays used for reference gene amplification

Primer name	Accession number	Primer sequence	ID number
hsa-miR-22-3p	MIMAT0000077	5' AAGCUGCCAGUUGAAGAACUGU	YP00204606
hsa-miR-30d-5p	MIMAT0000245	5' UGUAAACAUCCCCGACUGGAAG	YP00206047
SNORD48	NR_002745	5' TGATGATGACCCCAGGTAACTCT GAGTGTG	YP00203903
U6	NR_004394	5' GCGTCCCCGGTACGATTAGAA GAGACATAG	YP00203907
UniSp6	Not available	Proprietary data	YP00203954

2.3.3 Data Pre-processing, Normalization, and Relative Quantification

5. CFX384 Real-Time PCR Detection System (Cat.no. 184-5384 and 184-1000, Bio-Rad Laboratories GmbH, Munich, Germany).
6. CFX Manager Software, version 3.1 (Cat.no. 1845000, Bio-Rad Laboratories GmbH, Munich, Germany).

1. R, version 3.5.1 [13] (The R Foundation, Vienna, Austria).
2. GenEx 6.1—qPCR Data Analysis Software (MultiD Analyses AB, Gothenburg, Sweden).

3 Methods

The following steps are intended to be in high compliance with the MIQE guidelines [12], which have already been highlighted before. In addition to what is recommended in the MIQE guidelines regarding PCR experiments, also the steps before PCR, such as correct sample handling or RT, are equally essential and are also emphasized within the respective sections. In general, when working with RNA, several precautions have to be taken. Since RNases, whose biological function is to degrade any available RNA, are omnipresent and remarkably stable, an RNase-free laboratory environment must be ensured to avoid sample degradation. For example, instruments and working areas should be thoroughly cleaned, RNase-free water and consumables should be used, and gloves and lab coats should be worn at all times.

Since all pipetting steps are performed manually, master mix preparations are indispensable to minimize pipetting errors and technical variations.

3.1 Sample Preparation

1. Prior to sampling whole blood, a written informed consent of healthy individuals is obtained.
2. 2.5 ml of whole blood are drawn into PAXgene Blood RNA tubes via venipuncture according to clinical standard operating procedures (*see Notes 1 and 2*).
3. Instantly invert blood collection tubes ten times to allow for enhanced cell lysis and thorough mixing of blood and RNA-stabilizing additive.
4. After an upright overnight incubation (24 h) at room temperature, the tubes are stored at –20 °C.

3.2 RNA Isolation

1. Blood samples are thawed at room temperature for at least 2 h and inverted ten times.
2. For RNA purification, the PAXgene Blood RNA Kit is used, as it is designed for the integrated usage with PAXgene Blood RNA tubes. The isolation strategy is based on a silica-membrane technology, where total RNA > 18 nucleotides (nt), including miRNA, is bound to a spin column membrane, washed, and eluted with an appropriate buffer. The manufacturer's instructions are followed strictly (*see Note 3*).
3. RNA quality and quantity are assessed prior to sample storage at –80 °C.

3.3 RNA Quality and Quantity

1. Non-degraded RNA samples of high quality and integrity are important for any downstream application. Therefore, the RNA integrity number (RIN), ranging from 1 (heavily degraded RNA) to 10 (perfectly intact RNA), is determined using the Agilent 2100 Bioanalyzer [14]. The measurement is performed according to the manufacturer's instructions. With respect to subsequent RNA sequencing and RT-qPCR, a RIN higher than seven should be attained, indicating RNA of high integrity.
2. RNA yield, as well as successful RNA purification without contaminations pertaining to proteins or residues from the isolation procedure, is assayed using 1 µl of eluted RNA applied to a NanoDrop ND-1000 spectrophotometer. The 260/280 nm ratio should be between 1.8 and 2.2 for all RNA samples, and the 260/230 nm ratio should appear with a range from 2.0 to 2.2.

3.4 Library Preparation and Small RNA Sequencing

1. After quantifying RNA and testing for integrity and purity, samples are diluted to 200 ng RNA in a 6 µl reaction volume. Individual sequencing libraries are prepared using the NEB-Next Multiplex Small RNA Library Prep Kit for Illumina according to the manufacturer's instructions.

2. Adaptors are ligated onto both ends of all RNA molecules, which are subsequently reverse-transcribed to cDNA.
3. The ensuing PCR-based amplification step introduces multiplexing barcodes that tag each sample's cDNA with an individual sequence identifier, allowing the sequencing of multiple libraries in the same run. Experimental parameters such as the respective research question, sample type, RNA integrity and expected intra-group variability ultimately determine the appropriate degree of multiplexing, that is, how many samples will be pooled for a sequencing run.
4. After PCR amplification, size distribution and concentration of individual cDNA libraries are assessed by capillary electrophoresis using the Agilent 2100 Bioanalyzer and the DNA 1000 Assay.
5. The concentration of cDNA fragments carrying inserts in the miRNA size range of 130–150 base pairs (bp) is determined using the Bioanalyzer software's "Smear Analysis" feature.
6. Individual libraries are then combined so that each sample contributes the same amount of 130–150 bp cDNA to the pool.
7. Next, cDNA fragments are fractionated by subjecting the pool to an electrophoresis-based size selection step using a 4% agarose gel. Target cDNA is cut from the gel, purified, and subjected to a final quality control step (Agilent 2100 Bioanalyzer, High Sensitivity DNA Assay).
8. After verifying size distribution and purity, the library pool is quantified by qPCR and sequenced on the Illumina HiSeq2500 (50 cycles of single-end sequencing).

3.5 NGS Data Processing and Differential Gene Expression Analysis

1. Data pre-processing and quality control is performed based on raw sequencing data. Sequences are obtained in fastq file format, containing read sequences from small RNA inserts of interest as well as multiplexing barcodes and their corresponding quality scores (Phred score).
2. Since average miRNA lengths (~22 nt) will inevitably fall short of conventional sequencing run lengths (~50 bp), reads must be trimmed of 3'-adaptor sequences to facilitate correct alignment.
3. Furthermore, reads shorter than 16 nt, comprising degraded products and sequencing artifacts, should be excluded from the data set as well.
4. Pre-processing is achieved by running the Btrim software tool [15] with the following parameters: 3–15. Next, validity of sequencing is assessed by the quality control software FastQC [16] using default settings. Trimmed read data should be free

from low-quality sequences indicating sequencing errors (Phred score < 20) and display a distinct read length distribution centered around ~22 nt.

5. To translate individual sequences to counts of miRNAs necessary for statistical testing and all downstream analyses, trimmed reads must next be aligned against a reference database. Mature miRNA sequences are obtained from miRbase [17], and a mapping index is created by the bowtie-build tool within the bowtie software suite [18] using default parameters. Alignment against this miRNA index is carried out by employing the bowtie algorithm with the --SAM --best --nrc and -v 1 parameters, resulting in a single miRNA hit with a maximum of one mismatch for every sequencing read. miRNA readcount tables are then created from SAM file outputs of alignment by sorting, indexing and calling the sum of each miRNA using SAM-tools [19] with default parameters.
6. To address systematic variation such as differing sequencing depths or batch effects within the sequencing data set, read count data needs to be normalized. Employing the DESeq2 R package from Bioconductor [20], reads from each sample are corrected by a size factor obtained from the median expression ratio of geometric means. To account for biological variability and allow for exact instead of non-parametric tests, read data is then modeled on a negative binomial distribution with dispersion factors estimated via regression analysis over all miRNAs. Statistically significant miRNA expression changes are subsequently evaluated in pairwise comparison of treatment groups by Wald test and corrected for accumulation of alpha errors using the Benjamini–Hochberg method. Extensive DESeq2 sample workflows for various experimental setups can be found in the package’s vignette. In short, NGS data is loaded into R as data frames for both read count data (sample in columns, miRNAs in rows) and experimental design (treatments in columns, samples in rows). Next, normalization and distribution modeling are achieved by running the *DESeqDataSetFromMatrix* function with the following parameters: countData = {R object containing data frame with read count data}, colData = {R object containing data frame with experimental design}, design = ~{name of a column within experimental design} and running the *DESeq* function on the resulting data object. Finally, Wald statistics and p-value adjustment for two treatment groups can be called using the results function.

3.6 Selection of Reference Candidates

1. Normalized read counts for individual samples, previously used for differential gene expression analysis, serve as the basis for selecting qPCR reference candidates. As the hallmarks of

reference RNAs are robust expression levels and unaltered expression across samples and treatment groups, normalized counts should initially be filtered during data pre-processing in order to exclude miRNAs with very low and/or fluctuating expression. A mean expression of at least 100 reads per group seems to be a reasonable threshold for many experiments, depending on sequencing depth and number of multiplexed samples. Depleting miRNAs with expression levels below this threshold generates a new data set, which is subsequently screened for potential reference candidates.

2. Several algorithms aimed at identifying stably expressed genes in qPCR data have been developed for the selection of reference candidate genes, two of the most widely used being geNorm [21] and NormFinder [22]. When adjusting software settings, both can be used to assess stability of gene expression in RNA-Seq data instead of PCR data. Highly stable expression of a transcript is indicated by lower M values (geNorm) or stability values (NormFinder). Alternatively, the coefficient of variation (CV), calculated for each transcript across all samples, can be utilized to assess expression stability. As normalization with a single reference gene is usually insufficient [21–23], a set of multiple stably expressed transcripts should be selected, validated and averaged for normalization in a “reference index.” It is also important to mention that the selection of an appropriate “reference index” should be adapted and validated for each experiment separately, since miRNA expression can vary considerably between different tissues and disease conditions [24, 25].

3.7 Validation of Sequencing Results via RT-qPCR

Ideally, cDNA synthesis is performed in triplicate for each biological sample, and each of the triplicate cDNAs is quantified by real-time PCR in technical triplicates. However, this approach results in a tremendous workload and causes substantial costs with respect to consumables and time. While biological replicates are imperative prerequisites for every experimental setting, various approaches can be taken to reduce the number of technical replicates. For instance, artificial spike-ins with known concentrations introduced at different steps of the workflow (e.g. at RNA isolation and/or during cDNA synthesis) might help to monitor successful sample handling and allow for sample comparison and outlier identification.

The present experiment is based on relative quantification; thus, it is advisable to adjust the plate layout in a way that amplification reactions of control and treatment conditions for a certain gene can be screened in parallel. To still provide comparability across multiple interrelated measurements from different runs, inter-plate controls could be added to the plate layout (*see Note 4*).

1. For the reverse transcription of miRNAs into cDNA, the miRCURY LNA RT Kit and 10 ng of RNA template are used according to the manufacturer's instructions (*see Note 5*). In brief, a single reverse transcription reaction setup in a reaction volume of 10 μ l is pipetted into PCR clean tubes on ice and consists of the following: 2 μ l 5 \times miRCURY RT Reaction Buffer, 4.5 μ l RNase-free water, 1 μ l 10 \times miRCURY RT Enzyme Mix, 0.5 μ l UniSp6 RNA spike-in template, and 2 μ l template RNA previously diluted to 5 ng/ μ l. The lyophilized spike-in template is resuspended in 80 μ l nuclease-free water prior to usage.
2. A separate reaction, also referred to as noRT-control, is set up without 10 \times miRCURY RT Enzyme Mix to check for sample contamination with genomic DNA in subsequent qPCR reactions (*see Note 6*).
3. The cycling protocol is conducted with a thermal cycler and includes lid preheating, the reverse transcription reaction at 42 °C for 60 min and a heat inactivation step at 95 °C for 5 min. The cDNA is then stored at –20 °C until further processing or kept on ice for the ensuing qPCR.
4. Prior to setting up the amplification reaction by qPCR, lyophilized PCR primer assays are resuspended in 220 μ l nuclease-free water according to the manual.
5. cDNA templates are diluted 1:60 in nuclease-free water (*see Note 7*).
6. The reaction mix setup is pipetted into PCR clean plates in triplicates at room temperature. The 10 μ l reaction volume contains the following components: 5 μ l 2 \times miRCURY SYBR Green Master Mix, 1 μ l PCR primer mix, 3 μ l diluted cDNA template, 1 μ l RNase-free water (*see Note 8*). A separate reaction without any template, also referred to as no-template control (NTC), is set up for every primer, since a detectable amplification in these controls would point toward DNA contamination derived from kit components or improper pipetting. Due to an initially required heat activation of the kit-specific QuantiNova DNA polymerase, there is no need to pipette the reaction on ice. Next, the plates are sealed and spun down to eliminate air bubbles that might disturb fluorescence detection during amplification. Cycling conditions are chosen according to the manufacturer's advice (*see Table 2*). After each cycle as well as after every step of melting curve analysis, the fluorescence signal is acquired and collected at 515–530 nm (*see Note 9*).
7. Raw C_q (quantification cycle) values can be exported directly from the PCR recording software as data file. The pre-selected single threshold mode is chosen as the mode of determining C_q .

Table 2
Cycling conditions for qPCR

Step	Time	Temperature
1 Initial heat activation	2 min	95 °C
2 Denaturation	10 s	95 °C
3 Combined annealing/extension	60 s	56 °C
4 Repeat steps 2 and 3	40 cycles	
5 Melting curve analysis	Fast mode	60–95 °C/0.5 °C steps

values (see Note 8). The processing of obtained raw data is performed in R, version 3.5.1 [13].

8. In a first step, the last amplification cycles are treated as noise, since the probability of unspecific amplification rises with increasing cycle number, and C_q values higher than 37 are therefore discarded.
9. Next, outlier C_q values, meaning discrepancies from the characteristic distribution, are identified and should be treated as missing in further analyses. An outlier value is defined by a p-value of less than 0.01 based on the ISO 5725 guidelines [26]. Grubbs' test and the confirming Rosner's test are applied to detect outliers within the dataset using the R packages "outliers" [27] and "EnvStats" [28] and the functions *grubbs.test* and *rosnerTest* applying default settings. Missing values are replaced with the corresponding triplicate average. In the case that only one of the three replicates showed amplification, all replicates of the corresponding sample are eliminated from the final dataset. The thereby obtained C_q values of triplicate measurements are then averaged, creating a new dataset.
10. Checking the reference candidates previously selected based on NGS results for their true potential as reference miRNAs, the triplicate-averaged dataset is imported into the GeNorm and NormFinder algorithms, which both are implemented into GenEx software (by MultiD Analyses AB) and were already described in Subheading 3.6. The results, which depend on the respective dataset and experimental setup, are shown in Fig. 1a–c. The arithmetic mean of the best ranked reference miRNAs, termed reference index, is used in the following normalization step.
11. Once the selected reference genes are validated and an optimal combination is established, data normalization can be performed. This is the first step of relative quantification based on the $2^{(-\Delta\Delta C_q)}$ method according to Livak and Schmittgen [11]. The individual calculation steps are as follows:

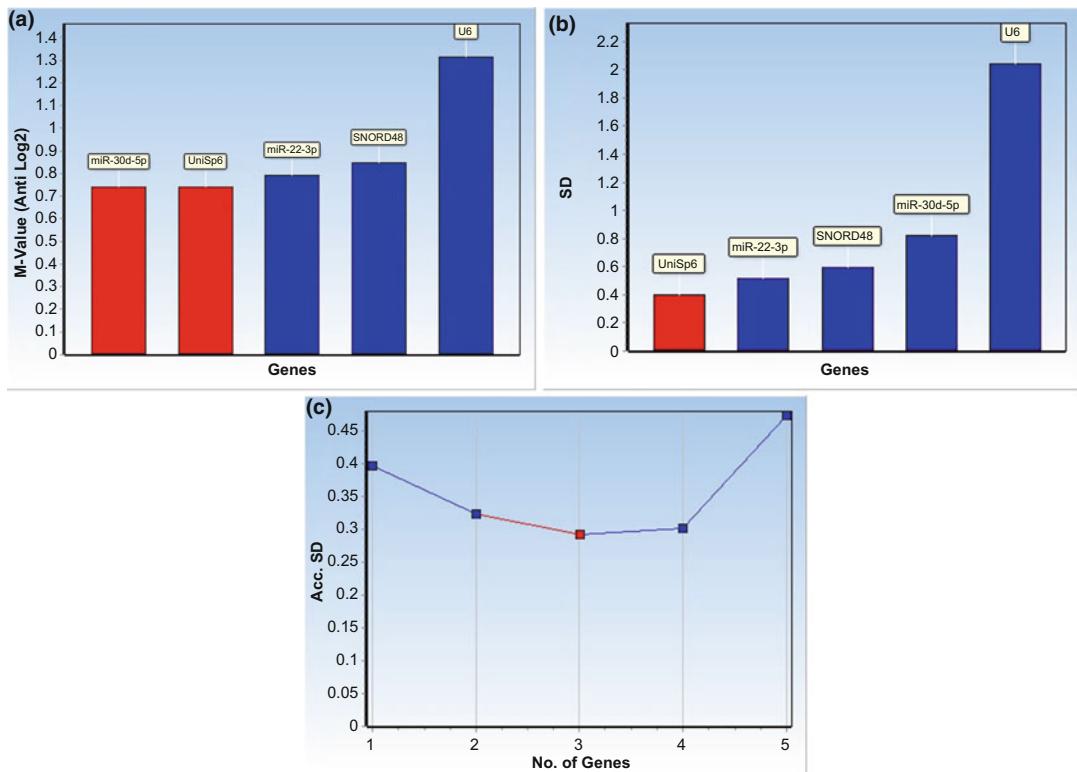


Fig. 1 GenEx output. **(a)** GeNorm, gene stability measure (M-Value), which is based on the arithmetic mean of all pairwise variations. **(b)** NormFinder, stability value (SD) based on an estimation of intragroup as well as intergroup variation. **(c)** NormFinder, recommended number of reference genes

$$\Delta C_q = C_q(\text{target gene}) - C_q(\text{reference index}) \quad (1)$$

$$\Delta\Delta C_q = \Delta C_q(\text{treatment}) - \Delta C_q(\text{control}) \quad (2)$$

$$\text{Fold Change} = 2^{(-\Delta\Delta C_q)} \quad (3)$$

To calculate fold changes of relative differential expression, it has to be considered that C_q values are already log-scaled. To obtain non-logarithmic fold changes, one has to raise basis 2 to the negative value of $\Delta\Delta C_q$ (see Eq. 3).

Statistical analyses are performed in R using the function *t.test* for two-tailed Student's *t*-test. *P* values smaller than 0.05 are considered as significant.

4 Notes

1. If the PAXgene Blood RNA tubes are the only tubes to be used during blood drawing, it is important to start with a disposal tube. Otherwise, the PAXgene Blood RNA tubes should be used at the end of the phlebotomy procedure.
2. In general, most clinical standard operating procedures for venipunctures include the following steps: assessment of the individual's physical condition, applying the tourniquet (on the upper arm), cleaning of the desired place of puncture with alcohol under sterile working conditions, second sterilization with alcohol only, penetrating the vein, release of the tourniquet, connecting collection tubes, removing needle, and applying pressure to the vein with flat hand until blood coagulation.
3. To increase RNA yield, it is recommended to apply the 40 μ l eluate on the same spin column for a second elution. In addition, incubating the spin column membrane on the benchtop for several minutes after the elution buffer has been applied can improve RNA yield.
4. Whenever it is planned to perform multiple PCR runs and analyze them in a merged experiment, it might be useful to add an inter-plate calibrator (IPC) to each run. This IPC should be constantly expressed or even constituted by an external spike-in. Applying the same sample-primer reaction to each of the PCR runs allows for comparison and calibration of different runs by calculating a calibration factor:

$$\text{Calibration Factor} = \text{IPC}_{\text{plate average}} - \text{IPC}_{\text{overall average}} \quad (4)$$

Lastly, calibration can be performed by subtracting the run-specific calibration factor from all C_q values pertaining to this run.

5. The principle of the miRCURY LNA miRNA PCR system is based on a two-step RT-qPCR comprising a universal RT reaction and an miRNA-specific amplification. In a universal first-strand cDNA synthesis, all mature miRNAs are polyadenylated and, thus serve as template sequences for reverse transcription using oligo-dT primers. Moreover, the introduction of technical variations is reduced to a minimum, and enormous time savings are ensured by a universal reaction setup. In the ensuing qPCR, miRNA-specific modified forward and reverse primers are used. The modification with locked nucleic acids (LNA) enables not only the adjustment of melting temperature but also increases the hybridization affinity, reduces background fluorescence, and increases sequence specificity of the primers [29]. For the detection of double stranded DNA, SYBR Green I is used, which emits green light at a wavelength

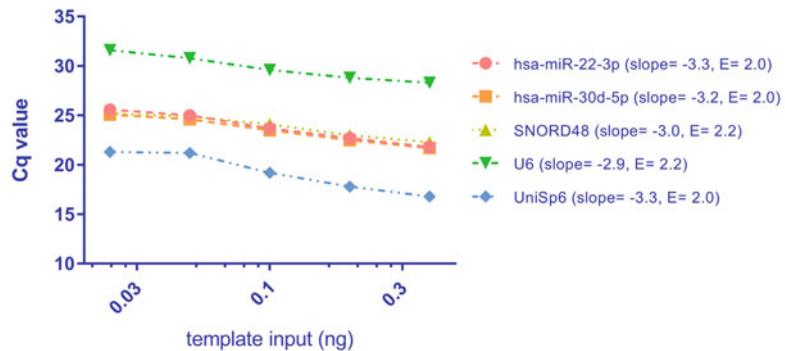


Fig. 2 Calibration curves of reference gene candidates. C_q values of different reference gene candidates are plotted against the logarithmized template input amount (ng)

of $\lambda_{\max} = 521$ nm upon intercalation into the minor groove during elongation.

6. In theory, there should be one accompanying noRT control for every sample. However, this would cause enormous costs in practice, since a lot more reagents and materials are needed. Thus, a pool of RNA samples from each group is often utilized in one RT reaction, which still allows for the detection of genomic DNA contaminations and saves both time and money.
7. Even though most RT-qPCR kits are designed for a certain input volume, the actual amount of RNA to be used can be different. Especially when using blood or blood compartments as starting material, the amplification of low abundant genes could be disturbed by residual PCR inhibitors [30], which might lead to reduced amplification efficiencies. To overcome this issue, the RNA input amount should be optimized. Therefore, a dilution series covering different cDNA input amounts of pooled samples is prepared, and the obtained C_q values are recorded (*see Fig. 2*). With decreasing dilution, the C_q values should get smaller. Based on the resulting curve, the slope can also be used to calculate primer efficiencies, whereby a slope of -3.3 corresponds to a perfect amplification factor (E) of 2, resulting in a primer efficiency of 100% based on the following equation:

$$\text{Amplification Factor } (E) = 10^{(-1/\text{slope})} \quad (5)$$

8. Since SYBR Green I does not distinguish between potentially occurring primer dimers, PCR by-products and specific amplification products, it is recommended to perform a melting curve subsequent to the real-time PCR analysis for primer quality control. With increasing temperature, the initial annealed and elongated double-stranded cDNA is increasingly

denatured, releasing SYBR Green I and thereby decreasing the fluorescence signal. In case of optimal primer specificity, only one peak should occur in the first derivative of the melting curve, whereas additional peaks would indicate the presence of unspecific amplification products or primer dimers.

9. When using real-time cyclers other than the CFX384Real-Time PCR Detection System, the addition of a passive reference dye, such as ROX, is advisable or is even a well-known necessity for instruments such as those from Applied Biosystems. By using a passive reference dye, it is possible to compensate for differences in fluorescence due to different well positions, slightly differing reaction volumes, and other device-dependent technical factors. Equally important and dependent on the respective PCR instrument are the software settings that have to be made before data export. According to the corresponding handbook, a manually set threshold and baseline might be necessary.

References

1. Buschmann D, Haberberger A, Kirchner B et al (2016) Toward reliable biomarker signatures in the age of liquid biopsies - how to standardize the small RNA-Seq workflow. *Nucleic Acids Res* 44:5995–6018. <https://doi.org/10.1093/nar/gkw545>
2. Pfaffl MW (2013) Transcriptional biomarkers. *Methods* 59:1–2. <https://doi.org/10.1016/jymeth.2012.12.011>
3. Strimbu K, Tavel JA (2010) What are biomarkers? *Curr Opin HIV AIDS* 5(6):463–466. <https://doi.org/10.1097/COH.0b013e32833ed177>
4. Wang J, Chen J, Sen S (2016) MicroRNA as biomarkers and diagnostics. *J Cell Physiol* 231 (1):25–30. <https://doi.org/10.1002/jcp.25056>
5. Ghai V, Wang K (2016) Recent progress toward the use of circulating microRNAs as clinical biomarkers. *Arch Toxicol* 90 (12):2959–2978. <https://doi.org/10.1007/s00204-016-1828-2>
6. Arneth B (2018) Update on the types and usage of liquid biopsies in the clinical setting: a systematic review. *BMC Cancer* 18(1):527. <https://doi.org/10.1186/s12885-018-4433-3>
7. Reithmair M, Buschmann D, Marte M et al (2017) Cellular and extracellular miRNAs are blood-compartment-specific diagnostic targets in sepsis. *J Cell Mol Med* 21:2403. <https://doi.org/10.1111/jcmm.13162>
8. Gibson U, Heid A, Williams P (1996) A novel method for real time quantitative RT-PCR. *Genome Res* 6:995–1001
9. Morrison T, Weiss J, Wittwer C (1998) Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *BioTechniques* 24:954–962
10. Arya M, Shergill IS, Williamson M et al (2005) Basic principles of real-time quantitative PCR. *Expert Rev Mol Diagn* 5:209–219. <https://doi.org/10.1586/14737159.5.2.209>
11. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C (T)) method. *Methods* 25(4):402–408. <https://doi.org/10.1006/meth.2001.1262>
12. Bustin SA, Benes V, Garson JA et al (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55:611–622. <https://doi.org/10.1373/clinchem.2008.112797>
13. R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>
14. Schroeder A, Mueller O, Stocker S et al (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* 7:3. <https://doi.org/10.1186/1471-2199-7-3>
15. Kong Y (2011) Btrim: a fast, lightweight adapter and quality trimming program for

- next-generation sequencing technologies. *Genomics* 98(2):152–153. <https://doi.org/10.1016/j.ygeno.2011.05.009>
16. Andrews S (2010) FastQC: a quality control tool for high throughput sequence data
 17. Kozomara A, Griffiths-Jones S (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* 42(Database issue):D68–D73. <https://doi.org/10.1093/nar/gkt1181>
 18. Langmead B, Trapnell C, Pop M et al (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25. <https://doi.org/10.1186/gb-2009-10-3-r25>
 19. Li H, Handsaker B, Wysoker A et al (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
 20. Love MI, Huber W et al (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. <https://doi.org/10.1186/s13059-014-0550-8>
 21. Vandesompele J, De Preter K, Pattyn F et al (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:RESEARCH0034
 22. Andersen CL, Jensen JL, Orntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 64 (15):5245–5250. <https://doi.org/10.1158/0008-5472.CAN-04-0496>
 23. Pfaffl MW, Tichopad A, Prgomet C et al (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper—excel-based tool using pair-wise correlations. *Biotechnol Lett* 26(6):509–515
 24. Lu J, Getz G, Miska EA et al (2005) MicroRNA expression profiles classify human cancers. *Nature* 435:834–838. <https://doi.org/10.1038/nature03702>
 25. Volinia S, Calin GA, Liu CG et al (2006) A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci U S A* 103:2257–2261. <https://doi.org/10.1073/pnas.0510565103>
 26. Burns MJ, Nixon GJ, Foy CA (2005) Standardisation of data from real-time quantitative PCR methods - evaluation of outliers and comparison of calibration curves. *BMC Biotechnol* 5:31. <https://doi.org/10.1186/1472-6750-5-31>
 27. Komsta L (2011) Outliers: tests for outliers, R package version 0.14. <https://CRAN.R-project.org/package=outliers>
 28. Millard SP (2013) EnvStats: an R package for environmental statistics. Springer, New York
 29. Kaur H, Arora A, Wengel J (2006) Thermodynamic, counterion, and hydration effects for the incorporation of locked nucleic acid nucleotides into DNA duplexes. *Biochemistry* 45:7347–7355. <https://doi.org/10.1021/bi060307w>
 30. Schrader C, Schielke A, Ellerbroek L et al (2012) PCR inhibitors - occurrence, properties and removal. *J Appl Microbiol* 113:1014–1026. <https://doi.org/10.1111/j.1365-2672.2012.05384.x>



Chapter 4

Enhanced Probe-Based RT-qPCR Quantification of MicroRNAs Using Poly(A) Tailing and 5' Adaptor Ligation

Valentin Vautrot and Isabelle Behm-Ansmant

Abstract

Probe-based quantitative PCR (qPCR) is a commonly used tool in the realm of real-time qPCR experiments since it is one of the most sensitive detection methods allowing an accurate and reproducible analysis. It uses real-time fluorescence from a fluorescently labeled probe that specifically targets the desired PCR product to measure DNA amplification at each cycle of the PCR. Coupled to a proper reverse transcription step, probe-based qPCR can be efficiently used for the analysis of the expression of difficult targets such as miRNAs. In this chapter, we describe the TaqMan® advanced miRNA assay in which, owing to a poly(A)-tailing step, the reverse transcription is advantageously performed at once for all the miRNAs in a given sample, and, coupled to the ligation of a 5' universal adapter, allows for a supplementary pre-qPCR amplification step increasing the sensitivity of the assay. Along this protocol, we also provide our general guidelines and advices to perform a reliable and successful quantitative analysis.

Key words TaqMan, MicroRNA, miR, miRNA, qPCR, RT-qPCR, Poly(A) tailing

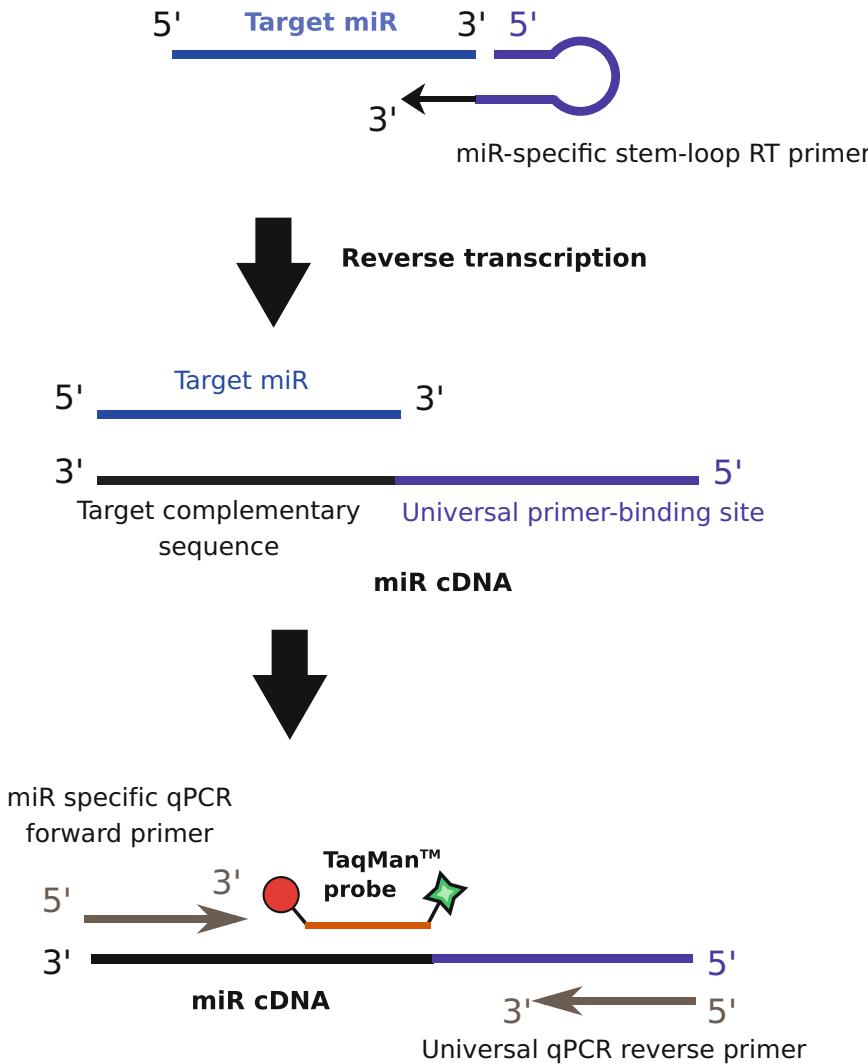
1 Introduction

MicroRNAs are endogenous small noncoding RNAs of 19 to 24 nucleotides (nts) in length that negatively regulate gene expression through base pairing with their target RNAs and interference with either their translation or stability. They are involved in the regulation of most of the fundamental cellular processes and their expression and function are tightly regulated in development and physiological homeostasis. Dysregulation of miRNAs is critical to pathogenesis of multiple human diseases and they constitute promising targets for the development of new diagnosis and therapeutic strategies [1–3]. Nevertheless, efficient detection and quantification of miRNAs remain challenging due to their small size, their low abundance (cellular miRNA concentration can be as low as 1000 molecules per cell) [4, 5], their sequence similarity with their own precursors (pri- and pre-miRNAs), and also the high degree of similarity between members of a same miRNA family

(members can differ by only one base). Currently, quantitative real-time PCR (qPCR), microarrays and next-generation sequencing (NGS) are the most commonly used techniques for quantifying miRNAs, with qPCR still being considered the gold standard for accurate, sensitive, fast and simple measurement of miRNA expression. During the past years several design strategies have been proposed for PCR-based miRNA detection. Since the main difficulty resides in the short length of miRNAs, most of these strategies aimed at increasing the size of the cDNA when performing the reverse transcription step in order to improve the efficiency and the specificity of the PCR amplification. This was firstly achieved by using a stem-loop primer to perform the reverse transcription [6]. The stem-loop primer consists in a constant region that forms a stem-loop structure and a variable six nucleotide-long extension. This extension is the reverse complement of the last 6 nucleotides at the 3' end of the miRNA and provides specificity, whereas the stem-loop part of the primer extends the miRNA to more than ~60 nts and includes a universal primer-binding site [6] (Fig. 1).

Quantitative PCR analysis can then be performed by combining the universal reverse primer with a miRNA-specific forward primer. This technique has several advantages: (a) it allows specific detection of mature miRNAs and not their precursors, since the stacked bases in the stem of the stem-loop create spatial constraints that greatly reduce the efficiency of the primer binding on miRNA precursors, (b) it is accurate and sensitive, and (c) total RNAs and even cell lysates can be used as starting material and there is no need to enrich for miRNAs. In order to improve even more the sensitivity and the specificity of the assay, probe-based qPCR was performed instead of the classical dye-based qPCR (SYBR® green). Probe-based qPCR uses real-time fluorescence from a fluorescently labeled probe that specifically targets the desired PCR product to measure its amplification at each cycle of the PCR even in the presence of nonspecific PCR products. Probe designs vary a lot but the most common type, named hydrolysis (e.g., TaqMan®) probes, incorporate a 5' reporter fluorophore (Applied Biosystems™ FAM™ or VIC™ dye) and a 3' quencher (Applied Biosystems™ NFQ™ dye coupled with Minor Groove Binder (MGB) in order to enhance the T_m of the probe and reduce its length) [7, 8] on a short DNA oligonucleotide complementary to the target sequence (Fig. 2).

Fluorescence resonance energy transfer (FRET) prohibits emission of the fluorophore while the probe is intact. During each PCR cycle, the 5' flap endonuclease activity of the Taq DNA polymerase hydrolyzes the probe as the primer is extended and the target sequence is amplified. This cleavage event moves away the reporter fluorophore from the quencher and leads to the emission of a fluorescent signal proportional to the number of copies generated



qPCR amplification

Fig. 1 The first solutions of miR quantification with TaqMan® probes include a reverse transcription primer specific for the miR target, for only 6 nts but with great specificity. Indeed, its 5' extremity forms an overhanging stem-loop structure that might enhance stability of the duplex with its target [6]. This stem-loop sequence is exogenous from the human and murine genomes. It also contains an adapter sequence protected inside the stem-loop structure, that is used after denaturation during PCR amplification step as a hybridization site for a universal primer, coupled to a forward primer specific of the miRNA target

(Fig. 2). The coupling of the stem-loop based RT and the TaqMan®-based qPCR (TaqMan® miRNA assay) is very efficient and allows specific detection of mature miRNAs even when present at low abundance. However, its main drawback is that it requires a specific stem-loop primer and thus a specific RT reaction for each

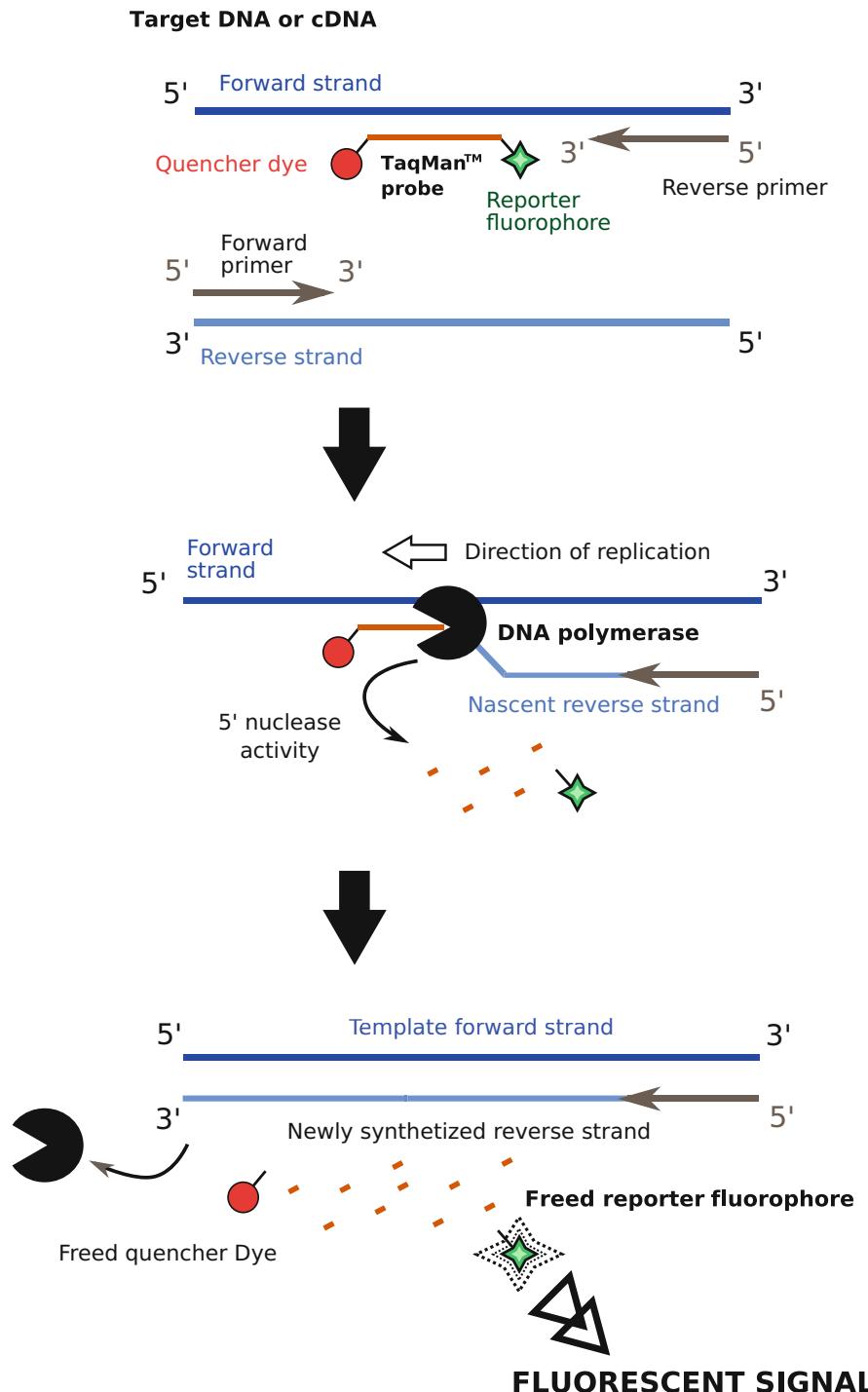


Fig. 2 The TaqMan® probe is specifically designed for hybridization to the DNA or cDNA target that needs to be quantified. At its 5' extremity, a fluorescent reporter dye (FAM™) was added (reporter fluorophore). While the probe integrity is maintained, its signal is blocked by a nonfluorescent quencher (NFQ) at its 3' end, where there is also a Minor Groove Binder modification (not drawn) that allows the increase of the melting

miRNA of interest which does not allow optimal normalization of the qPCR data according to the general MIQE guidelines for quantitative PCR and has a high cost in terms of reagents and biological samples [9]. Multiplexing protocols have been described but are only recommended for the detection of a few miRNA species at the same time since they reduce the efficiency of the RT step and thus slightly decrease the sensitivity of the assay.

In the last years, protocols have evolved to include a universal RT step in order to save biological samples, improve the normalization of the qPCR data and reduce the cost of the experiments. A universal stem-loop primer (USLP) method, in which the 3' extension of the stem-loop primer has been replaced by an 8 nt-long random sequence in order to combine with all mature miRNAs, has been proposed [10]. It works nicely with highly abundant miRNAs, but the sensitivity for low abundant miRNAs was found to be decreased, more especially that the authors performed traditional dye-based qPCR (SYBR® Green) and thus cannot fully guaranty the specificity of their assays. One could imagine to combine this USLP with probe-based qPCR in order to improve the specificity, but the sensitivity will still remain lower as compared to specific RT reactions. Another strategy that is commonly used is the poly (A) tailing of mature miRNAs in order to reverse transcribe all of them at the same time. At our knowledge two companies currently commercialize kits based on this strategy: Qiagen/Exiqon (miRCURY LNA™ Universal RT microRNA PCR, (Fig. 3)) and Life Technologies (TaqMan® Advanced MicroRNA Assay, (Fig. 4)).

In both cases the poly(A) tailing is performed using a template independent poly(A) polymerase and the reverse transcription step relies on an oligo(dT) adaptor primer that includes a tag at its 5' extremity. In the case of the miRCURY LNA™ Universal RT microRNA PCR protocol, this tag is used during the qPCR reaction to specifically amplify the miRNA of interest. Indeed, in this protocol the reverse primer of the qPCR is complementary to this tag, the poly (T) region and the 3' extremity of the miRNA, whereas the qPCR forward primer is complementary to the 5' extremity of the miRNA (Fig. 3). Importantly both the reverse and forward primers contain Locked Nucleic Acid (LNA) bases in order to be shorter and to fit on the miRNA without overlapping. This protocol is very efficient for abundant miRNAs nevertheless due to the universal RT its sensitivity for poorly expressed miRNAs

Fig. 2 (continued) temperature and the shortening of the length of the probe [7, 8]. As the 3' end of the probe is locked by these modifications, its replication is in theory not possible. The probe hybridization site is comprised between the forward and the reverse PCR primers. Thus, after the passage of the DNA polymerase during PCR amplification, the probe is hydrolyzed thanks to the 5' flap nuclease activity of the Taq DNA polymerase, which results in the emission of a quantifiable fluorescent signal from the reporter dye freed from the quencher activity

① Polyadenylation step



② Reverse transcription step



③ QPCR amplification

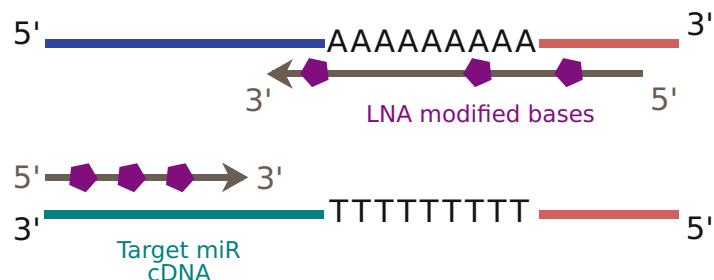


Fig. 3 The miRCURY LNA™ Universal RT microRNA PCR assay is composed of three main steps. (1) the addition of a poly(A) tail to nonpolyadenylated RNAs thanks to the poly(A) polymerase (PAP). (2) the poly(A) tail is then used for the hybridization of a universal RT primer, comprising an oligo-dT sequence backed up with another universal tag sequence, that allows the addition of this sequence in the final reverse transcription product. (3) The resulting product is then used directly as a template for a qPCR amplification reaction using SYBR® Green for quantification. For this reaction, the forward and reverse primers used carry several LNA (Locked Nucleic Acids, in purple) that achieve a comparable duplex stability with a shorter primer length, thus, coupled with the universal tag sequence added downstream, sufficiently compensating for the small size of miRNA templates

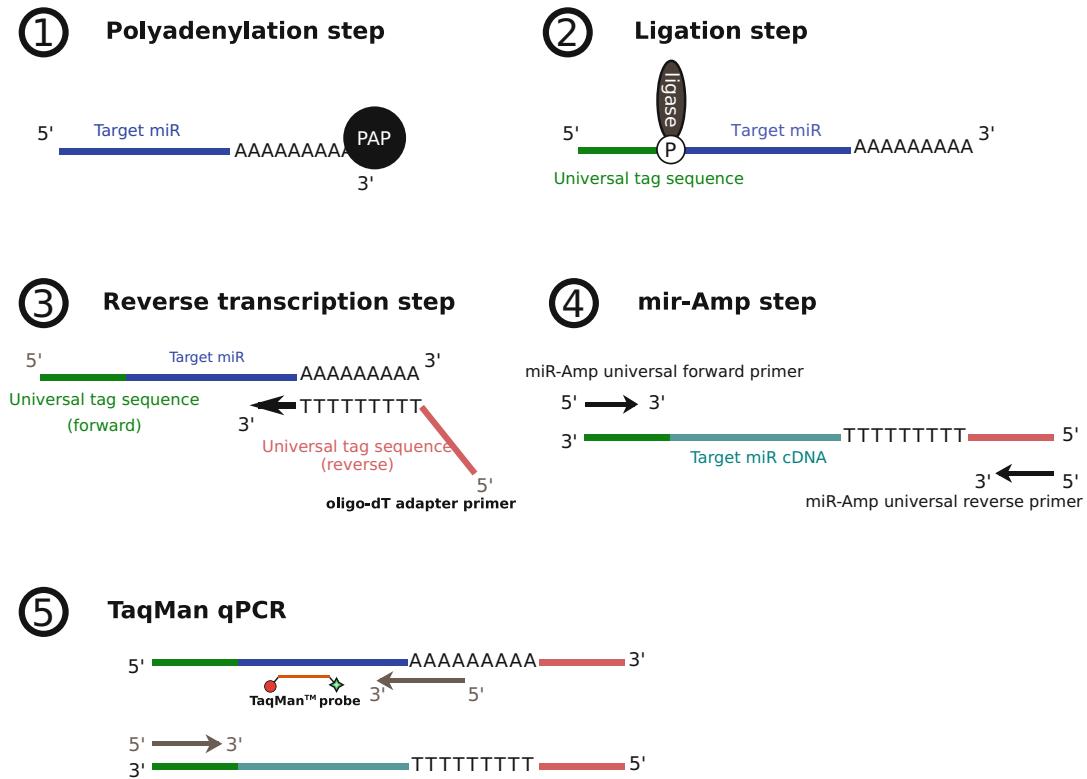


Fig. 4 The TaqMan® Advanced miRNA Assay is made up of five main steps: (1) the addition of a poly(A) tail to nonpolyadenylated RNAs thanks to the poly(A) polymerase (PAP). (2) The ligation of a universal adapter (ligation adapter) at the 5' extremity of 5'-phosphorylated RNAs thanks to a RNA ligase. (3) The poly(A) tail is then used for the reverse transcription step using an oligo-dT primer 5'-extended with a universal adapter sequence, that allows the addition of this adapter sequence in the final reverse transcription product. (4) The 5'- and 3'-universal adapter sequences are then used as hybridization sites for forward and reverse universal primers, respectively, in an additional amplification step (miR-Amp). (5) The resulting product can be used as a template in a classical TaqMan® qPCR reaction

even if being good is still lower as compared to RT specific protocols. In the case of the TaqMan® Advanced MicroRNA Assay (Fig. 4) the strategy is different: the poly(A) tailing is followed by a ligation step to graft a universal adapter at the 5' end of the miRNA sequences taking advantage of the 5'-phosphate extremity of mature miRNAs. The reverse transcription is then performed using a tagged oligo-dT primer that allows addition of a universal PCR amplification sequence at the 3' end of the cDNA sequence. Both adapters sequences (5' and 3') are then used in a supplemental amplification step (termed miR-Amp). This step is supposed to minimize the introduction of bias in relative quantitation as it requires the same couple of primers for each target and hence the efficiency is the same for each miRNA. Probe-based qPCR using TaqMan® probes is then performed to detect and quantify specific

miRNAs as previously described. Thus, the TaqMan® Advanced MicroRNA Assay has the double advantage of getting rid of the bias previously introduced by separated specific reverse transcription for each miRNA, and enhancing the sensitivity by adding a supplemental amplification step. It constitutes a practical, reliable, and sensitive technique for microRNA expression analysis and is presented further in the following sections.

2 Materials

2.1 Poly(A) Tailing Reaction

1. Standard thermal cycler and corresponding plates or tube strips.
2. Microcentrifuge adapted for thermal cycler strip tubes and thermal cycler, 96- or 384-well.
3. The following reagents from the Advanced TaqMan® cDNA synthesis kit: 10× Poly A Buffer; ATP 10 mM solution; Poly (A) Enzyme solution (5 U/μl). Prepare also the 50% PEG 8000 solution. microRNAs extracted from biological samples: 10 ng of total RNAs isolated from cells or tissues (in a maximum volume of 2 μl), or 2 μl of total RNAs isolated from whole blood or plasma (*see Note 1*). Make sure your RNA extraction method or kit is well adapted to the type of biological material you want to experiment on, and allows the purification of small RNAs (*see Note 2*). Also make sure at all times that RNA is properly handled to avoid RNase contamination and RNA degradation (*see Notes 3 and 4*). Avoid performing those reactions on RNAs isolated from tissues containing large amount of retrotranscriptase or PCR activity inhibitors.
4. Eventual exogenous control synthetic oligodexynucleotides, that must be phosphorylated at the 5' extremity, and prepared according to the manufacturer's instructions (*see Note 5*).

2.2 Ligation Reaction

The following reagents from the Advanced TaqMan® miRNA cDNA synthesis kit: 5× Ligase Buffer; RNA Ligase solution (10 U/μl); 50% PEG 8000 solution; 25× Ligation Adaptor solution.

2.3 Reverse Transcription Reaction

The following reagents from the Advanced TaqMan® miRNA cDNA synthesis kit: 10× RT Enzyme Mix; 5× RT Buffer; 20× Universal RT Primer solution; 100 mM dNTP Mix (25 mM each).

2.4 miR-Amp Reaction

The following reagents from the Advanced TaqMan® miRNA cDNA synthesis kit®: 20X miR-Amp Primer Mix; 2X miR-Amp Master Mix.

2.5 TaqMan® qPCR Reaction

1. The following reagents from the TaqMan® advanced miRNA assays, specifically designed for your target(s) miR(s) of interest, endogenous and exogenous controls (*see Notes 5–7*): 2× Fast Advanced Master Mix; 20X TaqMan® Advanced miRNA Assay; RNase-free water.
2. Real-time qPCR instruments compatible with the TaqMan® advanced miRNA assay, and their compatible optical plates and manufacturer software (e.g., 7500 Real-Time PCR System, 7500 Fast Real-Time PCR System, StepOnePlus™ Real-Time PCR Systems, ViiA™ 7 Real-Time PCR System, QuantStudio® 6 Flex, 7 Flex, and 12K Flex Real-Time PCR Systems).

3 Methods

Carry out all procedures at room temperature unless otherwise specified. Also make sure until the cDNA synthesis reaction is over that RNA is properly handled to avoid RNase contamination and RNA degradation (*see Notes 3 and 4*). When handling any solution containing enzymes, use vortex at slow speeds or prefer flickering to homogenize your solutions.

3.1 Poly(A) Tailing Reaction

1. Thaw your RNA samples and cDNA synthesis reagents on ice.
2. Put the tube of 50% PEG 8000 at room temperature now for a better handling in time during the adaptor ligation reaction (*see Note 8*).
3. Add your eventual spike-in synthetic oligonucleotide, that will serve as an exogenous control, into your sample, at a final concentration of 1–10 pM (*see Note 5*).
4. In a 1.5 ml microcentrifuge tube, prepare a reaction mix with the following volumes per sample: 0.5 µl of 10× Poly(A) buffer, 0.5 µl of ATP, 0.3 µl of poly A enzyme, and 1.7 µl of RNase-free water for a total volume of 3 µl per reaction. When preparing this reaction mix for several samples, account for a 10% of pipetting error and increase the volumes in this proportion.
5. Homogenize and short-spin the tubes in a microcentrifuge.
6. In a PCR tubes strip (or plate for a large number of samples), mix 2 µl of your input isolated RNA (*see Subheading 2.1*) with 3 µl of your prepared reaction mix.
7. Short-spin in a centrifuge adapted for format of PCR tubes (strips or plate).
8. Using a thermal cycler, incubate for 1 cycle of 45 min at 37 °C followed by 10 min at 65 °C and hold the reaction at 4 °C when finished.

3.2 Ligation Reaction

1. In a 1.5 ml microcentrifuge tube, prepare a reaction mix with the following volumes per sample: 3 µl of 5× DNA ligase buffer, 4.5 µl of 50% PEG 8000 solution, 0.6 µl of 25× Ligation Adaptor solution, 1.5 µl of RNA ligase solution, and 0.4 µl of RNase-free water, for a total volume of 10 µl per reaction. When preparing this reaction mix for several samples, account for a 10% of pipetting error and increase the volumes in this proportion. Be cautious when pipetting the PEG 8000 solution (*see Note 8*).
2. Homogenize and short-spin the tubes in a microcentrifuge.
3. Add 10 µl of your reaction mix into the PCR tubes obtained after the poly(A) tailing reaction, for a final volume of 15 µl.
4. Short-spin in a centrifuge adapted for the format of PCR tubes used.
5. Using a thermal cycler, incubate for 1 cycle of 1 h at 16 °C and hold the reaction at 4 °C when finished.

3.3 Reverse Transcription Reaction

1. Do not denature the RNA samples before the reverse transcription step.
2. In a 1.5 ml microcentrifuge tube, prepare a reaction mix with the following volumes per sample: 6 µl of 5× RT buffer, 1.2 µl of dNTP mix (25 mM each), 1.5 µl of 20× universal RT primer solution, 3 µl of 10× RT enzyme mix, and 3.3 µl of RNase-free water, for a total volume of 15 µl per reaction. When preparing this reaction mix for several samples, account for a 10% of pipetting error and increase the volumes in this proportion.
3. Homogenize and short-spin the tubes in a microcentrifuge.
4. Add 15 µl of your reaction mix into the PCR tubes obtained after the previous ligation reaction, to attain a total volume of 30 µl.
5. Short-spin in a centrifuge adapted for the format of PCR tubes used.
6. Using a thermal cycler, incubate for 1 cycle of 15 min at 42 °C, followed by a stopping step of 5 min at 85 °C, and hold the reaction at 4 °C when finished.
7. Following this step, reaction products can be stored at –20 °C.

3.4 miR-Amp Reaction

1. In a 1.5 ml microcentrifuge tube, prepare a reaction mix with the following volumes per sample: 25 µl of miR-Amp Master Mix, 2.5 µl of miR-Amp Primer mix, and 17.5 µl of RNase-free water, for a total volume of 45 µl per reaction. When preparing this reaction mix for several samples, account for a 10% of pipetting error and increase the volumes in this proportion.
2. Homogenize and short-spin the tubes in a microcentrifuge.

Table 1
PCR temperature cycles for the miR-Amp reaction

Step	Temperature (°C)	Duration	Number of cycles
Enzyme activation	95	5 min	1
Denaturation	95	3 s	14
Annealing/extension	60	30 s	
Stop reaction	99	10 min	1
Hold	4	Hold	1

3. In a new PCR tubes strip or PCR plate, dispose for each sample of 45 µl of this reaction mix and complete to a 50 µl reaction with 5 µl of the products obtained after the previous reverse transcription step.
4. Short-spin in a centrifuge adapted for the PCR tubes format used.
5. Using a thermal cycler, incubate according to the following temperature cycles (Table 1).
6. Following this step, reaction products can be stored at –20 °C for up to 2 months.

3.5 qPCR Using TaqMan Advanced miRNA Assay

Please work using general recommendations for PCR and qPCR to avoid contamination and obtain the best results (*see Notes 9 and 10*).

Before proceeding to the qPCR, take some time to choose your endogenous controls, indispensable for efficient relative comparison of expression in any qPCR experiment (*see Note 11*).

1. Dilute the cDNA template obtained during the miR-Amp step into 0.1× TE buffer or RNase free water to a 1:10 ratio.
2. In a 1.5 ml microcentrifuge tube, prepare a reaction mix with the following volumes per sample: 10 µl of 2× Fast Advanced Master Mix, 1 µl of 20X TaqMan® Advanced miRNA assay, and 1 µl of RNase-free water, for a total volume of 15 µl per reaction. When preparing this reaction mix for several samples, account for a 10% of pipetting error and increase the volumes in this proportion.
3. In a new qPCR optical plate, dispose for each sample of 15 µl of this reaction mix and complete to a 20 µl reaction with 5 µl of diluted cDNA (*see Notes 11 and 12*) to place your control assays).
4. Vortex and short-spin in a centrifuge adapted for the PCR tubes format used.

Table 2
**qPCR temperature cycles for the TaqMan® Advanced miRNA assay,
according to the real-time PCR instrument used**

		StepOnePlus™, ViiA™ 7, and QuantStudio® systems		7500 and 7500 fast systems	
Step	Temperature (°C)	Duration (s)	Number of cycles	Duration (s)	Number of cycles
Enzyme activation	95	20	1	20	1
Denaturation	95	1	40	3	40
Annealing/ extension	60	20		30	

- Set up the qPCR plate according to your real-time PCR instrument's manufacturer instructions and following to the cycles exposed in Table 2.
- Verify the aspect of the amplification curves (*see Note 13*) and proceed to quantification (*see Note 14*).

4 Notes

- As in whole blood and plasma samples the amount of starting material can greatly vary among individuals (and given difficulties of quantification related to the very low amounts of purified material), it is recommended to use a fixed input volume of 2 µl of RNA volume extracted in the same conditions, instead of a measured starting concentration.
- Commercial miRNA extraction kits are available for several types of biological samples (tissue, blood, cultured cells) and allow the purification of RNAs with a size cutoff under 200 nt long. If using other kits or methods, make sure that it allows recovery of small fragments of RNA in a PCR-compatible buffer. Please note that commercial kits often allow better reproducibility, which is a crucial parameter in qPCR experiments. If you have limited choice over the RNA isolation methods, keep in mind that precipitation methods following TRIzol® or phenol/chloroform extraction can have a different efficiency on RNA precipitation depending on their size. For the isolation of microRNA, we would recommend the use of ethanol precipitation over isopropanol, often coupled to a neutral carrier agent like glycogen or glycoblue, as it allows a

better precipitation of smaller fragments of RNA, especially when samples are chilled (i.e., -80°C) during the precipitation period. RNAs should then be resuspended in RNase-free water or Tris-EDTA (TE) Buffer, pH 8.0. The addition of a fixed amount of an exogenous synthetic RNA oligonucleotide to your biological sample during the RNA isolation procedure can help you to monitor the efficiency of your extraction protocol (more details on those exogenous controls are provided in **Note 5**). The quality of your RNA preparations can be assessed by using microfluidic chips dedicated to the analysis of small RNA samples such as the Agilent small RNA chips with the 2100 Bioanalyzer system. We highly recommend to use small RNA chips allowing to resolve small nucleic acids in the size range of 6 to 150 nucleotides rather than chips dedicated to total RNA analysis. Indeed, even though the last ones allow objective measurement of RNA quality with RIN (RNA Integrity Number), this RIN value mainly reflects the integrity of ribosomal RNAs, which are the dominant RNA species in total RNAs rather than the integrity of less abundant miRNAs and we prefer to visualize and specifically quantify our small RNA species on gel.

3. Main recommendations for handling RNA samples include the following ones:
 - Use sterile and RNase-free labware (including disposables) and reagents.
 - When not in use, put the RNA samples on ice to slow down potential RNase activity.
 - Wear disposable gloves to avoid contamination with RNases from the skin, and clean lab coat, to handle samples and reagents.
 - Use workspaces and equipment dedicated to RNA handling (including pipettes), and clean regularly your working surfaces with RNase removal products such as RNaseZap from Ambion (or 0.5% SDS followed by 3% H_2O_2).
4. If necessary, you can add RNase inhibitors into your starting RNA samples according to the manufacturer's instruction to avoid RNase activity, though miRNAs are usually more stable than longer RNAs.
5. To monitor the influence of sample preparation on the expression of miRNAs and even on your endogenous controls, and to verify the absence of polymerase inhibitors, you can add synthetic exogenous oligonucleotides sequences or purposefully designed constructs with known concentration/copy number as controls [11]. Those spike-in controls have to be phosphorylated at their 5' extremities to be compatible with the TaqMan[®]

Advanced miRNA assays. Applied Biosystems proposes several predesigned assays for such sequences not present in humans (e.g., based on the sequences of ath-miR159a, cel-lin-4-5p, cel-miR-2-3p, cel-miR-238-3p, cel-miR-39-3p, cel-miR-54-3p, and cel-miR-55-3p (according to IDs in miRbase v22)). Those spike-in controls can be either combined with the biological sample during the RNA isolation procedure or added just before starting the poly(A)-tailing reaction.

6. For the current available assays list, please visit www.lifetechnologies.com/advancedtaqmanmirna
7. The normalization using endogenous controls is at the center of every quantification technique used in qPCR, and the slightest variation can have a negative impact on expression estimation. To limit this risk, the use of several endogenous controls is the sine qua non condition for a reliable quantification. If available, they must be chosen using literature data or expression databases in the concerned cells or tissues, or at default, in preestablished lists of miRNAs referenced as generally having a stable expression in the cellular model or tissue of interest. On the other hand, it was shown repeatedly that expression of genes typically referenced as housekeeping genes can vary a great deal depending on sample conditions (time of sampling, growth conditions, state of tissues, etc.) [12–14]. Thus it is excluded to use only one endogenous control, as still seen occasionally in some studies, and even recommended, whatever controls were chosen, to test their variation in the different conditions possible in the experiment. To facilitate those tests, several software solutions exist to help determine the best endogenous controls among several tested based on their stability of expression. The initial geNorm solution, once available preprogrammed in an Excel sheet [14], is now fully incorporated in the paid qbase+ software (*see* <https://genorm.cmgg.be/>). Nevertheless, a handful of free equivalent alternatives can be found, that give similar results: Normfinder (<https://www.moma.dk/normfinder-software>) [15], Bestkeeper (<https://www.gene-quantification.de/bestkeeper.html>) [16], and, if you are familiar with the python programming language, Eleven, a free and open source implementation of geNorm in python 2.7 (<https://pypi.org/project/eleven/>).
8. PEG 8000 is very viscous and if not careful during pipetting it may result in large errors. To limit those errors, ensure that the PEG solution had enough time to attain room temperature, proceed slowly and wait 5–10 s after each manipulation of the pipette plunger. You may also use the reverse pipetting technique or positive displacement pipettes if possible.

9. Prepare the qPCR reactions in a dedicated area, with material free of any contact with high copy number nucleic acids of any kind (e.g., that of PCR end products, preparations of genomic DNA or plasmidic DNA, and bacteria manipulation); use filter tips and wear a clean lab coat to avoid contamination.
10. Avoid as much as possible the exposure of the TaqMan® probes (comprised in the TaqMan® Advanced miRNA assay tube) to the light before the experiment, as it can impact the intensity of the fluorescent signal beforehand. To this end and for better results, prepare the qPCR reaction plate extemporaneously.
11. If using different reaction volumes, scale all components of the reaction mix proportionally. Reaction volumes <10 µl are not recommended. Make sure you do not exceed volume capacity of your qPCR plate wells.
12. Because of the differences that can arise in quantification otherwise, it is better to have all the endogenous and exogenous control assays deposited on the same qPCR plate, and from the same sample preparation.
13. Most of the thresholds chosen in the amplification curve to establish the C_q value are, in general, successfully determined automatically by the instrument software algorithm. Nevertheless, it is always a good idea to manually check those thresholds on the graphical representations of the curves. It is important that this threshold is established in the beginning of the exponential growth phase of the fluorescence signal, that is, the straight part before the inflection point in the often used log representation of the amplification curve.
14. The quantification can be made using several techniques. One popular and easy to use with TaqMan® is the $2^{(-\Delta\Delta C_q)}$ or $\Delta\Delta C_q$ method [17, 18], because TaqMan® probes are optimized for maximum efficiency, provided that you stay in the starting RNA concentration range and the condition given by the manufacturer. Whatever method you choose, be sure to generally follow the MIQE guidelines.

Acknowledgments

This work was supported by a public grant overseen by the French National Research Agency (ANR) as part of the second “Investissements d’Avenir” program FIGHT-HF (reference: ANR-15-RHU-0004).

References

1. Basak I, Patil KS, Alves G et al (2016) MicroRNAs as neuroregulators, biomarkers and therapeutic agents in neurodegenerative diseases. *Cell Mol Life Sci* 73:811–827. <https://doi.org/10.1007/s00018-015-2093-x>
2. Bertoli G, Cava C, Castiglioni I (2016) MicroRNAs as biomarkers for diagnosis, prognosis and theranostics in prostate cancer. *Int J Mol Sci* 17:421. <https://doi.org/10.3390/ijms17030421>
3. Sastre B, Cañas JA, Rodrigo-Muñoz JM et al (2017) Novel modulators of asthma and allergy: exosomes and microRNAs. *Front Immunol* 8:826. <https://doi.org/10.3389/fimmu.2017.00826>
4. Bissels U, Wild S, Tomiuk S et al (2009) Absolute quantification of microRNAs by using a universal reference. *RNA* 15:2375–2384. <https://doi.org/10.1261/rna.1754109>
5. Bosson AD, Zamudio JR, Sharp PA (2014) Endogenous miRNA and target concentrations determine susceptibility to potential ceRNA competition. *Mol Cell* 56:347–359. <https://doi.org/10.1016/j.molcel.2014.09.018>
6. Chen C, Ridzon DA, Broomer AJ et al (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* 33: e179. <https://doi.org/10.1093/nar/gni178>
7. Afonina I, Zivarts M, Kutyavin I et al (1997) Efficient priming of PCR with short oligonucleotides conjugated to a minor groove binder. *Nucleic Acids Res* 25:2657–2660
8. Kutyavin IV, Lukhtanov EA, Gamper HB et al (1997) Oligonucleotides with conjugated dihydropyrroloindole tripeptides: base composition and backbone effects on hybridization. *Nucleic Acids Res* 25:3718–3723
9. Bustin SA, Benes V, Garson JA et al (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55:611–622. <https://doi.org/10.1373/clinchem.2008.112797>
10. Yang L, Wang S, Tang L et al (2014) Universal stem-loop primer method for screening and quantification of microRNA. *PLoS One* 9: e115293. <https://doi.org/10.1371/journal.pone.0115293>
11. Nolan T, Hands RE, Ogunkolade W et al (2006) SPUD: a quantitative PCR assay for the detection of inhibitors in nucleic acid preparations. *Anal Biochem* 351:308–310. <https://doi.org/10.1016/j.ab.2006.01.051>
12. Bustin SA, Benes V, Garson J et al (2013) The need for transparency and good practices in the qPCR literature. *Nat Methods* 10:1063–1067. <https://doi.org/10.1038/nmeth.2697>
13. Bustin SA (2014) The reproducibility of biomedical research: sleepers awake! *Biomol Detect Quantif* 2:35–42. <https://doi.org/10.1016/j.bdq.2015.01.002>
14. Vandesompele J, De Preter K, Pattyn F et al (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:research0034
15. Andersen CL, Jensen JL, Ørntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 64:5245–5250. <https://doi.org/10.1158/0008-5472.CAN-04-0496>
16. Pfaffl MW, Tichopad A, Prgomet C et al (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper—Excel-based tool using pair-wise correlations. *Biotechnol Lett* 26:509–515
17. 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>
18. Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative C (T) method. *Nat Protoc* 3:1101–1108



Chapter 5

A Novel System to Discriminate HLA-C mir148a Binding Site by Allele-Specific Quantitative PC R

Priscilla Biswas, Eddi Di Marco, and Mauro S. Malnati

Abstract

The levels of expression of the HLA-class I molecules are critical for modulating T/NK lymphocytes effector functions. Among HLA molecules, HLA-C, the most recent developed form of class I antigens, is subjected to multiple post transcriptional level of regulation that affect its cell surface expression.

We describe a new method of allele-specific real-time PCR that monitor the integrity/disruption of the binding site of the microRNA Hsa-miR-148a, a key factor associated to the levels of HLA-C expression in the Caucasian populations.

Key words Allele-specific real-time PCR, Single-nucleotide polymorphism, miRNA 148a binding site, Epigenetic regulation, HLA-C expression

1 Introduction

The human leukocyte antigen (HLA) class I molecules present pathogen-derived antigens to CD8 T lymphocytes and regulate the functions of natural killer (NK) cells. While CD8 T lymphocytes sense pathogen-derived epitopes [1] present in the peptide binding groove of HLA-B class I molecules and, to a lesser extent, of HLA-A and C molecules, NK cells are mainly regulated by the level of cell surface expression [2, 3] of HLA- class I alleles. In humans, the tuning of NK activity is based on the interaction between the killer immunoglobulin-like receptors (KIRs), the HLA-C molecule [4, 5] and some HLA-A and B alleles ($\approx 40\%$ of the alleles). KIRs belong to the immunoglobulin (Ig) superfamily and are encoded by at least 15 distinct polymorphic gene loci located in a region of 150–200 kb of the leukocyte receptor complex (LRC) on human chromosome 19q13.42 [6, 7] and display either inhibitory (L) or activating (S) functions [8, 9]. They are inherited as different haplotypes, resulting in a centromeric and a telomeric KIR profile that varies in number and composition of genes [9, 10]. Interestingly, evolution drove at least 5 KIRs to recognize HLA-C alleles,

the most recently developed form among the HLA molecules [5]. Compared to its HLA-A and B counterparts, HLA-C is expressed at lower cell surface levels, is less polymorphic, and has evolved to have more extensive interactions with KIRs [11–14]. HLA-C expression varies widely in an allele-specific manner [15, 16] influencing disease outcome, especially as observed in the case of HIV-1 infection [15, 17–19]. Thus, high HLA-C protein expression has been associated with the frequency of HLA-C peptide-restricted CTL responses and increased frequency of viral escape mutations [15, 20], suggesting that higher HLA-C expression exerts a selection pressure on the virus [15], which is in line with the recently discovered virus-mediated downregulation of HLA-C expression [21].

In contrast, high levels of HLA-C expression correlate with increased risk of Crohn's disease [16], psoriasis [22], and, in cases of unrelated haematopoietic transplantation, with poor outcome and graft-versus-host disease [23]. The divergent effects of HLA-C expression on infectious and autoimmune diseases, combined with evidence for the recent origin of mutations that influence expression [24], suggest a dynamic evolutionary balance between positive and negative gene regulation, which can shift with the epidemiological cycling of specific pathogens.

HLA-C expression is modulated by multiple factors acting at several levels, from transcription to epigenetic post transcriptional regulation such as miRNA binding or protein assembly due to a strong peptide selectivity mediated by the antigen-binding cleft. Thus, integration of these factors consequently leads to a net effect that determines abundance at the cell surface [25, 26]. Among the factors influencing HLA-C expression levels a polymorphism (SNP) that maps in the 3' UTR of HLA-C alleles (rs67384697 G/del) affects the binding of the microRNA Hsa-miR-148a [18]. This SNP was shown to partially influence cell surface expression of HLA-C with weakly expressed alleles possessing an intact miR148a binding site (263G) and highly expressed alleles having a deletion in the site (263del), thus escaping the regulation by that microRNA [18]. Here we describe a novel, fast, and reproducible method based on an allele-specific real-time PCR technique that allows to determine the status of the binding site for microRNA Hsa-miR-148a on all HLA-C alleles.

2 Materials

2.1 DNA Extraction

1. MagDEA DNA 200 GC extraction KIT.
2. DNase-free microtubes and microtips with filters.
3. DNase-free H₂O for solutions and negative control samples.
4. TE buffer (10 mM Tris adjusted to pH 8.0 with HCl, 1 mM EDTA) for DNA storage and dilutions (*see Note 1*).

Plate Setup

1. 96-well plates for fast-run qPCR (Thermo-Fisher System)
2. DNA from study samples.
3. Ethnic Diversity DNA Panel (ECACC).
4. Mastermix for PCR reaction containing Taq polymerase and all other necessary components including the reference dye ROX (Kapa Probe Fast qPCR Kit) (*see Note 2*).
5. Primers (two forward and one reverse).
6. Probes (two conjugated with FAM, one with VIC fluorochromes).

The primers/probes are as follows:

miR148-Frw_195-212 (18-mer): CTCCCCGATCATCTTCC. It will be briefly called F195 (*see Note 3*).

miR148-Frw_202-224 (22-mer): ATCATCTTCCTGTTCCA-GAGAA. It will be briefly called F202 (*see Note 4*).

miR148-Rev_384-406 (24-mer): CAGGTCTTATTGCTCTCTAACAC. It will be briefly called R384 (*see Note 5*).

263-del_probe (20-mer): 5' VIC-TGTCTCAACTTACGTGT AC-quencher 3' (*see Note 6*).

263-ins_probe (17-mer): 5' FAM-AACATTCATGGTGCGCTG-quencher 3' (*see Note 7*).

C*07/17-Ins_probe (18-mer): 5' FAM-AGCTCAGTGC ACCATGAA-quencher 3'. It will be referred to as C*07 ins probe (*see Note 8*) (Fig. 1).

Primers and probes were set at 100 µM stock solutions and stored at -20 °C. Appropriate working solutions were prepared when needed.

The instrument used for the quantitative PCR runs is the Applied Biosystems™ QuantStudio™ 12K Flex with its dedicated software for the analyses. The fast-mode was chosen throughout the study.

3 Methods

3.1 DNA Extraction

The DNA of healthy donors used in this study derived from either whole blood or peripheral blood mononuclear cells (PBMC). DNA from whole blood samples was extracted using the Magtration System 12GC, automated platform and MagDEA DNA.

200 GC extraction KIT according to the manufacturer's instruction (Precision System Science PSS Co. Ltd). PBMC were prepared by Ficoll gradient density centrifugation and dry pellets of about 1 to 2×10^6 were stored at -20 °C. DNA was then extracted

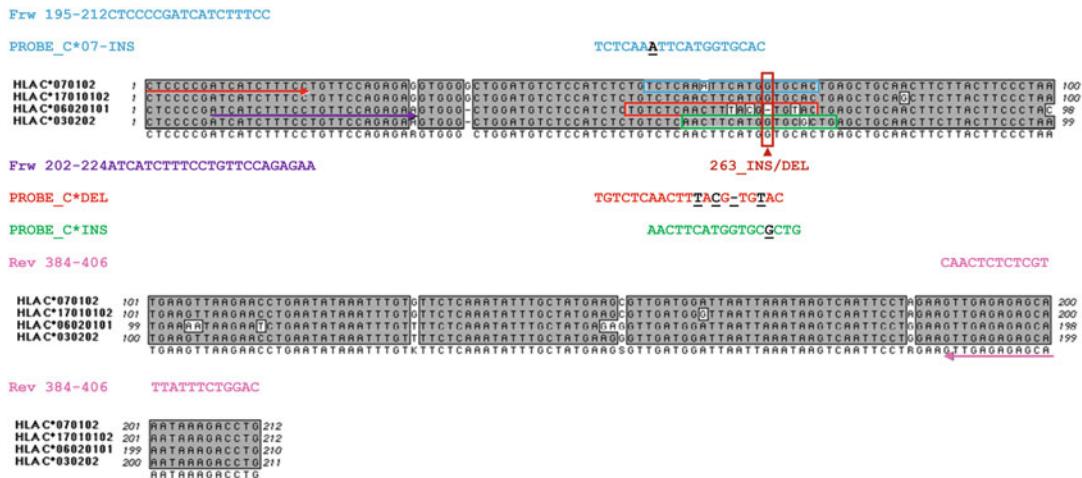


Fig. 1 Nucleotidic alignments of HLA-C*07, C*17, C*06 and C*03 containing the primers' and probes' positioning

by phenol-chloroform procedure, as elsewhere described [27]. Finally one set of samples was commercially available and provided as purified DNA from healthy donors of different ethnicities (European Collection of Cell Cultures -ECAAC plate EDP-1). DNA concentrations were measured by UV-spectroscopy determining the 260/280 DNA absorbance ratio (optimal ratio = 1.9) and adjusted to the required concentrations with TE buffer.

3.2 Preparation of the Reaction Mix

The volume of the reaction for the fast-run is 20 µl/well. Triplicates or at least duplicates were run for each sample. Single probe reactions were performed with the C7 probe and F195 primer whereas dual probe reactions were performed with the ins and del probes and F202 primer; the R384 was used in each reaction. Calculate the number of wells needed for each PCR reaction, then prepare the required mixes in eppendorf tubes. From the 100 µM stock solutions prepare working solutions of primers at 20 µM (1:5 dilution) and probes at 5 µM (1:20 dilution) with TE buffer. Add 1 and 0.9 µl of working solution for probe and primers, respectively, for each well. The final concentration is 900 nM for the primers and 250 nM for the probes. The mastermix is at a 2× concentration, thus add 10 µl for each well. Finally adjust the volume with the required amount of dH₂O. The DNA at the required concentration will be seeded in 5 µl/well, but only when the mixes are already loaded into the 96-well plate. An example of the preparation of a mix for 50 wells for a dual probe reaction is displayed below.

50 wells \times 20 μ l = 1 ml	
mastermix 2 \times	500 μ l
F202 primer 20 μ M	45 μ l
R384 primer 20 μ M	45 μ l
Ins probe 5 μ M	50 μ l
Del probe 5 μ M	50 μ l
dH ₂ O	60 μ l
DNA	250 μ l
	1000 μ l = 1 ml

3.3 Preparation of Standard Curve and Plate Loading

To verify whether all reagents work correctly it is important to perform few experiments utilizing serial dilutions of DNA extracted from HLA class I-typed reference samples. At least three/four dilutions of reference DNA need to be tested for each system covering a range of DNA concentration between 1/2 ng and 20/50 ng (Fig. 2).

To correctly manipulate the target DNA a flow laminar hood is preferred otherwise a bench should be properly cleaned with denaturating agents such as NaClO and SDS or commercially available preassembled solutions. It is mandatory that this area should be in a different room where clinical samples are processed or loaded in the PCR plate.

A dedicated set of calibrated micropipettes and gloves must be used during the entire procedure. Example of preparation of a 4 point standard curve from 20 to 2.5 ng/w is shown below.

1. Set 4 eppendorf tubes (of 1 ml volume) each containing 100 μ l ml TE buffer.
2. Label tubes with the final concentration of the DNA (20, 10, 5, 2.5 ng/w).
3. Warm-up HLA-typed reference DNA, previously set at 8 ng/ μ l, for 5 min at 70 °C.
4. Cool on ice for 20 s.
5. Pipet 100 μ l of the stock DNA into the first eppendorf tube emptying the tip once (without resuspending to avoid bubbles) and avoid to submerge the tip into the buffer; then immediately close the DNA stock and remove it from the bench.
6. Mix by inversion (10–15 times) or by gentle vortexing.
7. Spin briefly to remove droplets from the cap of the tube.
8. Repeat the complete procedure 4 times always starting from the last diluted solution; in the last tube the volume will be 200 μ l.

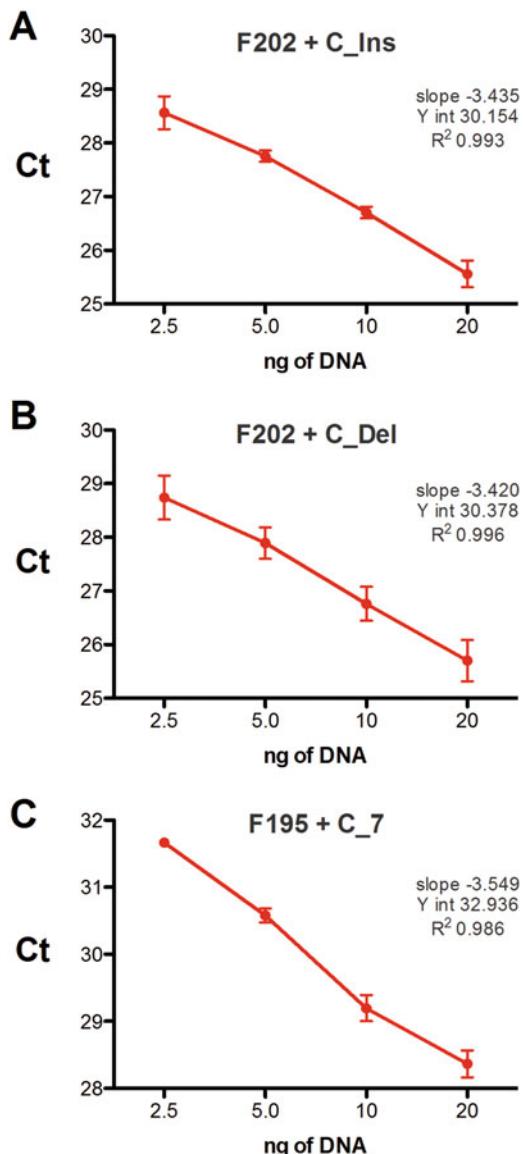


Fig. 2 Amplification profile of the rs67384697G/del (microRNA Hsa-miR-148a binding site) system. The panels depict the three different allele-specific PCR systems set up in this study for HLA-C_{ins} alleles (**a**), HLA-C_{del} alleles (**b**) and HLA-C*07 (**c**). The DNA used was from a C*03-C0*3 donor (**a**), a C*12-C*15 donor (**b**) and a C*07-C*07 donor (**c**). The standard curves range from 2.5 ng/w to 20 ng/w of DNA (x axis) and their threshold cycles (Ct) are reported on the y-axis. The panels show also the slope, the Y-intercepting cycle and the R^2 coefficient of each curve

9. The volume to be plated of each dilution is 5 µl/w.
10. Change gloves after each dilution point.

3.4 Plate Loading

Prepare a 96-well template page where the location of each sample is clearly written in order to easily load the plate. Seed 15 µl/well of each mix prepared, changing microtip every 4–6 wells. Seed 5 µl/well of the DNA samples, changing one microtip/each well. When seeding the different DNA concentrations of a standard curve plate the less concentrated first followed by the more concentrated ones. Finally, add 5 µl/well of dH₂O for the negative control wells for each mix. After sealing the plate with an adhesive cover, briefly centrifuge it to assure that the small reaction volume is placed deep into the V-bottom wells of the 96-well plate.

3.5 PCR Runs in the QuantStudio™ 12K Flex Instrument

The fast-mode PCR runs last only 30 min. The conditions are the following: 20 s for DNA denaturation and enzyme activation at 95 °C, 40 cycles of annealing extension lasting 20 s at 60 °C intermingled by 40 cycles of denaturation at 95 °C for 20 s.

4 Notes

1. The purpose of TE buffer is to solubilize DNA while protecting it from degradation. The slight alkalinity of TE buffer prevents chances of DNA acid hydrolysis. EDTA further inactivates nucleases by binding to metal cations required by these enzymes. A stock can be prepared (for 500 ml stock: 1 M Tris–HCl pH 8: 5 ml; 0.5 M EDTA–HCl pH 8: 1 ml; dH₂O: 494 ml), autoclaved to sterilize and stored at RT.
2. The choice of the mastermix is a critical point because the reaction efficiency of the mastermixes varies widely due to the different processivity of Taq polymerases contained and the salt concentrations. Here we report an experiment performed with two mixes whose employ two different mastermixes (primers, probes, and DNA were the same) with a titration of DNA (Donor 1 HLA-C*8, -C*14). Thus consistency of mastermix choice throughout a set of experiments should be mandatory (Fig. 3).
3. This primer works for HLA-C*07 and C*17 alleles. This primer was chosen analysing 912 HLA-C specific sequences. The complete description of the strategy for primers/probes design is reported elsewhere [28].
4. This primer is specific for all the HLA-C alleles except for C*07 and C*17 alleles. It derives from the analysis of 682 HLA-C sequences belonging to all the HLA-C alleles besides C*07/C*17.

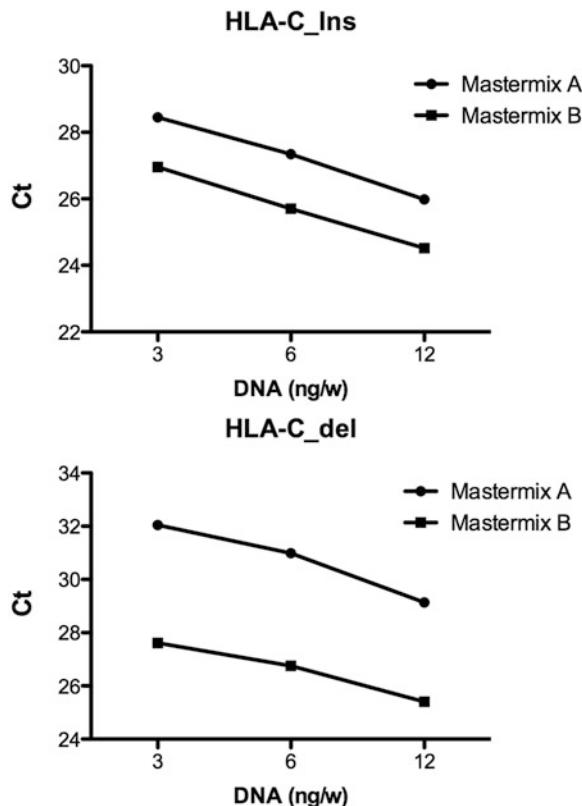


Fig. 3 Comparison of mastermix-related efficiencies. HLA-C_{ins} (upper panel) and HLA-C_{del} (lower panel) PCR standard curves data were generated with two mixes in which the only difference was the mastermix (A or B). The DNA was from a C*08-C*14 donor

5. This reverse primer is used to amplify all HLA-C alleles (C*07, C*17, all other ins, and del alleles). The sequence is conserved in all the 908 HLA-C specific sequences analyzed.
6. This probe is specific for all the HLA-C del alleles: it is conserved in all the 378 HLA-C sequences belonging to C*02, *05, *06, *08, *12, *15, *16 alleles analyzed.
7. This probe is specific for HLA-C*01, *03, *04, *14, *18 alleles. It derives from the analyses of 396 HLA-C sequences belonging to the C*01, *03, *04, *14, *18 alleles. This probe has one mismatch in the 3'end of the sequence (position 14) and can be used with C*17 alleles as described [28].
8. This probe is specific for C*07 alleles. It is conserved in the entire 222 HLA C*07 alleles sequences analyzed. This probe can be used for detecting 263 G/del SNP of C*17 alleles as described [28].

References

1. Margulies DH, McCluskey J (2003) Antigen processing and presentation. In: Pauls WE (ed) *Fundamental immunology*. Lippincott Williams & Wilkins, Philadelphia, pp 571–612
2. Moretta A, Bottino C, Vitale M et al (1996) Receptors for HLA class-I molecules in human natural killer cells. *Annu Rev Immunol* 14:619–648
3. Long EO (1999) Regulation of immune responses through inhibitory receptors. *Annu Rev Immunol* 17:875–904
4. Parham P (2005) MHC class I molecules and KIRs in human history, health and survival. *Nat Rev Immunol* 5:201–214
5. Parham P, Norman PJ, Abi-Rached L et al (2012) Human-specific evolution of killer cell immunoglobulin-like receptor recognition of major histocompatibility complex class I molecules. *Philos Trans R Soc Lond B Biol Sci* 367:800–811
6. Wende H, Colonna M, Ziegler A et al (1999) Organization of the leukocyte receptor cluster (LRC) on human chromosome 19q13.4. *Mamm Genome* 10:154–160
7. Biassoni R, Ugolotti E, De Maria A (2010) Comparative analysis of NK cell receptor expression and function across primate species: perspective on antiviral defenses. *Self Nonself* 1:103–113
8. Biassoni R, Malnati MS (2018) Human natural killer receptors, co-receptors, and their ligands. *Curr Protoc Immunol* 121(1):e47
9. Middleton D, Favel G (2009) The extensive polymorphism of KIR genes. *Immunol* 129:8–19
10. Pyo C, Guethlein LA, Vu Q et al (2010) Different patterns of evolution in the centromeric and telomeric regions of group A and B haplotypes of the human killer cell Ig-like receptor locus. *PLoS One* 5:e15115
11. Bashirova AA, Martin MP, McVicar DW, Carrington M (2006) The killer immunoglobulin-like receptor gene cluster: tuning the genome for defense. *Annu Rev Genomics Hum Genet* 7:277–300
12. Snary D, Barnstable CJ, Bodmer WF, Crumpton MJ (1977) Molecular structure of human histocompatibility antigens: the HLA-C series. *Eur J Immunol* 7:580–585
13. Zemmour J, Parham P (1992) Distinctive polymorphism at the HLA-C locus: implications for the expression of HLA-C. *J Exp Med* 176:937–950
14. Apps R, Meng Z, Del Prete GQ et al (2015) Relative expression levels of the HLA class-I proteins in normal and HIV-infected cells. *J Immunol* 194:3594–3600
15. App R, Qi Y, Carlson JM et al (2013) Influence of HLA-C expression level on HIV control. *Science* 340:87–91
16. Kulkarni S, Qi Y, O'huiggin C et al (2013) Genetic interplay between HLA-C and MIR148A in HIV control and Crohn disease. *Proc Natl Acad Sci U S A* 110:20705–20710
17. Fellay J, Shianna KV, Ge D et al (2007) A whole-genome association study of major determinants for host control of HIV-1. *Science* 317:944–947
18. Kulkarni S, Savan R, Qi Y et al (2011) Differential microRNA regulation of HLA-C expression and its association with HIV control. *Nature* 472:495–498
19. Thomas R, Apps R, Qi Y et al (2009) HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA-C. *Nat Genet* 41:1290–1294
20. Blais ME, Zhang Y, Rostron T et al (2012) High frequency of HIV mutations associated with HLA-C suggests enhanced HLA-C-restricted CTL selective pressure associated with an AIDS-protective polymorphism. *J Immunol* 188:4663–4670
21. Apps R, Del Prete GQ, Chatterjee P et al (2016) HIV-1 Vpu mediates HLA-C downregulation. *Cell Host Microbe* 19:686–695
22. Majorczyk E, Matusiak L, Nowak I et al (2014) A single nucleotide polymorphism -35 kb T > C (rs9264942) is strongly associated with psoriasis vulgaris depending on HLA-Cw/06. *Hum Immunol* 75:504–507
23. Petersdorf EW, Gooley TA, Malkki M et al (2014) HLA-C expression levels define permissible mismatches in hematopoietic cell transplantation. *Blood* 124:3996–4003
24. O'h Uigin C, Kulkarni S, Xu Y et al (2011) The molecular origin and consequences of escape from miRNA regulation by HLA-C alleles. *Am J Hum Genet* 89:424–431
25. Sibilio L, Martayan A, Setini A et al (2008) A single bottleneck in HLA-C assembly. *J Biol Chem* 283:1267–1274
26. Kaur G, Gras S, Mobbs JI et al (2017) Structural and regulatory diversity shape HLA-C protein expression levels. *Nat Commun* 8:15924

27. Malnati MS, Scarlatti G, Gatto F et al (2008) A universal real-time PCR assay for the quantification of group-M HIV-1 proviral load. *Nat Protoc* 3:1240–1248
28. Biswas P, Ugolotti B, Di Marco E et al A fast and reliable method for detecting SNP rs67384697 (miRNA 148a binding site) by a single run of allele-specific real-time PCR. Submitted for publication



Chapter 6

Detection of Yellow Fever Virus by Quantitative Real-Time PCR (qPCR)

Gisela Freitas Trindade, Sheila Maria Barbosa de Lima, Constança Britto, and Alice Gomes Fernandes-Monteiro

Abstract

The recent resurgence of yellow fever virus (YFV) activity in the tropical regions of Africa and South America has sparked renewed interest in this infamous arboviral disease. Also, the development and production of viral vaccines involve several steps that need the monitoring of viral load throughout the process (antigen production, purification, and inactivation). Currently, these steps are followed by plaque lysis titration assay, whose results take about 7–10 days to come out and thus resulting in a laborious and time-consuming approach. With the advent of quantitative real-time PCR (qPCR), we have a faster method to be applied during vaccine production and also to be effectively used for the diagnosis of YFV infection. The technique herein standardized proved to be effective for determining YF viral load both *in vivo* and *in vitro*, thus becoming a very important tool for laboratory analysis to verify the vaccination status of individuals, beyond acting as a quality control for vaccine production and diagnosis.

Key words Yellow fever virus, qPCR, Viral vaccines, Viral load, Molecular diagnosis, TaqMan fluorogenic probe

1 Introduction

Despite the availability of a safe and effective vaccine, yellow fever (YF), a zoonotic flaviviral disease caused by YF virus (YFV) remains a cause of major epidemics in Africa and South America [1–4]. YFV has reemerged in certain areas, and recent estimates indicate that 43–52% of the population in YFV-endemic regions still require vaccination [5]. Although the patterns by which epidemics emerge from the enzootic cycle are well recognized, the timely implementation of emergency immunization campaigns is often hindered by delays in performing a specific laboratory diagnosis. It is not unusual that outbreaks numbering hundreds or thousands of clinically defined cases occur without a specific YFV testing. In addition, YF may presents with signs and symptoms that are similar to dengue virus infections, a related flavivirus endemic to all regions

with YFV transmission [6]. Numerous conditions can mimic YF in their clinical presentations and occurrence in outbreaks [2, 4, 7, 8]. Other viral agents of hepatitis and hemorrhagic fever are especially important in the differential diagnosis in areas of South America and Africa where YF is endemic [7].

The development and production of YF viral vaccines, in general, involve several steps that need viral load monitoring throughout the process. These steps range from the production of antigens, purification, inactivation and lyophilization, up to preclinical testing and clinical trials, and once the product is licensed, a process of continuous pharmacological surveillance is necessary [7]. Usually, the detection of YFV is carried out by virus cultivation on susceptible cells followed by immunofluorescence, ELISA, immunohistochemical examination or PCR [8–11]. For quantifying the load of infectious particles in cell culture or serum samples, plaque assay is still the commonly used method. However, plaque titration is cumbersome and time-consuming technique that takes at least 5 days to get the result [12]. With the development of quantitative real-time PCR, a rapid and accurate alternative to quantify viral load is available [13].

Here, we describe a real-time qPCR assay based on the NS5 polyprotein sequence of the 17DD yellow fever virus vaccinal strain used in Latin America and African countries [12].

2 Materials

There are different types of RNA extraction protocols in the literature and the choice of protocol should be based on the best RNA yield for the type of biological material to be processed. It also depends on the laboratory infrastructure. RNA can be extracted from biological fluids such as serum, saliva, urine and liquor, or from fresh or paraffin-shaped tissue. RNA sample should be stored at –80 °C and not be thawed more than three times to avoid degradation. For reduced volume samples, wash buffer or RNase free ultrapure water should be added to complete the final volume. For RNA extraction of samples with large volumes, it is recommended a previous step for sample concentration. As negative control for sample processing, human serum of healthy individual is included.

2.1 RNA Extraction and cDNA Synthesis

1. The protocol is carried out using QIAamp® Viral RNA Mini kit, QIAGEN according to the manufacturer's instructions (*see Notes 1 and 2*).
2. Complementary DNA (cDNA) is synthesized using the High-Capacity cDNA Reverse Transcription mix (Applied Biosystems).

3. Micropipettes.
4. RNase-free Filter tips.
5. Sterile tubes (1.5 mL) for sample processing and RNA collection.

2.2 Reference Controls

TaqMan® RNase P Control Reagents kit can be used as an endogenous internal control in a multiplex qPCR format, for the analysis of human clinical samples (*see Note 3*).

It should be emphasized that either, the EXO-IPC DNA or the IAC can be used in a multiplex format (a same reaction for the detection/quantification of YFV and the reference control) or individually (singleplex), in any kind of biological sample (serum, urine, blood, cultivated virus). In these cases, the samples should be “contaminated” with the exogenous control DNA prior to RNA extraction (*see Subheading 3.3, step 5*).

2.3 TaqMan Quantitative Real-Time PCR (qPCR)

In our study, the qPCR assay is directed to the NS5 region of the yellow fever virus (YFV), as previously described [9, 12] (Fig. 1). Reactions are performed with the TaqMan® Universal PCR Master Mix (Applied Biosystems). The following primers and probe are used to detect and quantify YFV- NS5 region (*see Note 4*):

Forward: 5' GCA CGG ATG TAA CAG ACT GAA GA 3'.

Reverse: 5' CCA GGC CGA ACC TGT CAT 3'.

Fluorogenic probe: 5' FAM-CGA CTG TGT GGT CCG GCC CAT C-TAMRA 3'.

When testing human clinical samples, the reaction is performed in multiplex using the TaqMan® RNase P Control Reagents kit (VIC/TAMRA probe labeled) (*see Subheading 2.2*).

Individual qPCR reactions (singleplex) directed to the Exogenous Internal Positive Control EXO-IPC (VIC/TAMRA probe labeled) were run in parallel, in the same plate for the testing of in vitro serum-free YFV samples. Individual reactions are also performed for protocols that use the heterologous cloned sequence as Internal Amplification Control—IAC. The following primers and probe are used to quantify the aquaporin gene of *Arabidopsis thaliana* [14]:

IAC Forward: 5'ACC GTC ATG GAA CAG CAC GTA 3'.

IAC Reverse: 5'CTC CCG CAA CAA ACC CTA TAA AT 3'.

IAC probe: 5'Cy5-TTG GAG CAT CTG TTC TTG AAG GTG TTT TAG C-BHQ3'.

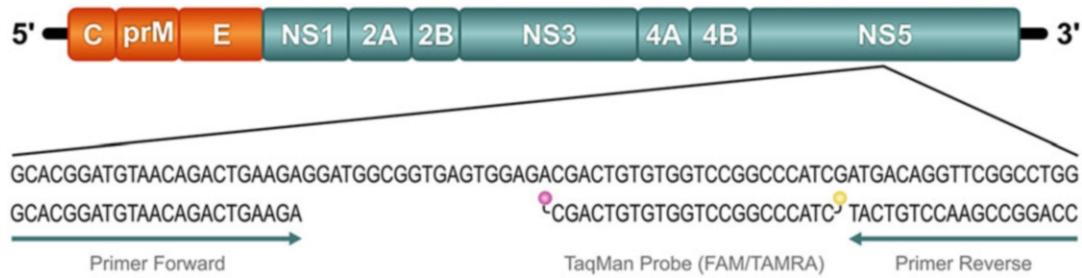


Fig. 1 Schematic representation of the yellow fever genome region amplified by qPCR. The orange blocks represent sequences that encode structural proteins (C, prM, E), and green blocks encode the nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5). The NS5 amplified sequence is demonstrated with the positioning flanking primers and the internal fluorogenic probe labeled with reporter 5'FAM (pink dot) and quencher 3'TAMRA (yellow dot)

3 Methods

3.1 RNA Extraction

RNA extraction can be performed either manually or automated. The principle of the method is also quite variable; it can be carried out totally in-house or using commercial kits. In our protocol, the QIAGEN kit is used following the manufacturer's recommendations.

1. The first step is the addition of 560 µL of AVL Buffer freshly prepared containing the carrier RNA (lysis buffer) into a 1.5 mL microcentrifuge tube.
2. A volume of 140 µL sample is added to the AVL Buffer–carrier RNA in the microcentrifuge tube. It is very important to mix the solution for 15 s by pulse-vortexing (*see Note 5*).
3. The mixture is incubated for 10 min at room temperature (15–25 °C) (*see Note 6*).
4. After incubation, a quick centrifugation is needed to prevent drops remaining in the lid, to avoid contamination between samples or loss of volume.
5. Add 560 µL ethanol (96–100%) in the sample mixture (*see Note 7*), and mix for 15 s by pulse-vortexing. After mixing, centrifuge the tube to prevent drops remaining from inside the lid.
6. Apply 630 µL of the **step 5** solution in the column previously placed into a 2 mL collection tube, without wetting the circlet. Close the cap and centrifuge for 1 min at $6000 \times g$ (8000 rpm) (*see Note 8*). Place the column into a new 2 mL collection tube, discard flow-through and the used collection tube.
7. Add 500 µL AW1 Buffer into the column (*see Note 9*). Centrifuge for 1 min at $6000 \times g$ (8000 rpm). Replace the column into a new 2 mL collection tube and discard the previous tube containing the filtrate.

8. Add 500 μ L AW2 Buffer and centrifuge for 3 min at maximum speed ($20,000 \times g$, 14,000 rpm).
9. Discard the collection tube containing the filtrate to eliminate any possible carryover by residual AW2 Buffer, and place the column into a new 2 mL collection tube. Centrifuge at full speed for 1 min (see Note 10).
10. Discard the collection tube and set the column into a clean 1.5 mL tube. Carefully add 60 μ L AVE Buffer into the column and incubate at room temperature for 1 min.
11. Centrifuge for 1 min at $6000 \times g$ (8000 rpm).

Viral RNA is stable for up to 1 year when stored at -30°C to -15°C or at -80°C to -65°C . Do not manually process more than 24 samples at a time to avoid cross-contamination. It is important to carefully open the lids and change gloves whenever necessary to avoid contamination. Include positive control with approximately 100 copies/reaction and negative samples, as controls.

3.2 cDNA Synthesis

Complementary DNA (cDNA) is synthesized using the High-Capacity cDNA Reverse Transcription mix, Applied Biosystems, containing 10 \times Reverse-Transcription (RT) Buffer, 10 \times RT Random Primers, 25 \times dNTP mix (100 mM), reverse transcriptase (50 U/ μ L), RNase inhibitor, and nuclease-free water. Store all components at -15°C to -25°C . Sterile tubes of 1.5 mL and 0.2 mL are necessary.

Reverse transcription reaction is performed with random primers in 20 μ L extracted RNA added to 20 μ L of High-Capacity cDNA Reverse Transcription mix (Applied Biosystems), as follows:

1. Pipet 4.0 μ L of 10 \times Buffer into a 0.2 mL microcentrifuge tube.
2. Add 2.0 μ L of 25 \times dNTP mix to this tube.
3. Add 4.0 μ L of 10 \times RT Random Primers, 2.0 μ L reverse transcriptase (50 U/ μ L) (see Note 11), 1.0 μ L RNase inhibitor, and 7.0 μ L nuclease-free water (see Note 12).
4. After the addition of 20 μ L RNA sample, reaction is performed in the thermal cycler using the following conditions: 50 $^{\circ}\text{C}$ for 2 min; 37 $^{\circ}\text{C}$ for 120 min; 4 $^{\circ}\text{C}$ on hold.
5. The obtained cDNA (40 μ L) is stored at -80°C until use.

3.3 TaqMan Quantitative Real-Time PCR (qPCR)

The qPCR assays are directed to a sequence of the NS5 region of the yellow fever virus (YFV), as previously described [9, 12] (see Subheading 2.3 for primers and probe sequences). Reactions can be performed in any real-time PCR thermal cycler available in the laboratory.

Primers (nM)			
Reverse	Forward		
	100	300	600
100	100/100	300/100	600/100
300	100/300	300/300	600/300
600	100/600	300/600	600/600
Probe (nM)			
50	75	100	125
			150
			175

Fig. 2 Example of standard curve for the absolute quantification of yellow fever virus using plasmid DNA containing the NS5 amplified sequence. Each point of the curve is obtained by serial dilutions of plasmid DNA revealing a reportable range of 10^7 to 10^2 copies/reaction. The values of Slope; Intercept (Y-inter); R^2 and Efficiency (Eff%) should be evaluated to validate the test as described in the text

1. In order to standardize the best concentration of oligonucleotide primers, three distinct concentrations of each oligonucleotide were combined (Fig. 2).

For these optimizing assays, the dilution corresponding to 10^4 virus copies per reaction was established and the concentration of TaqMan probe was fixed at 150 nM. Once the establishment of the optimal concentration of oligonucleotides set, it was selected the probe concentration that resulted on fluorescent signal emission at the lowest Ct (threshold cycle, where the reaction crosses the threshold of detection). It will be at this point that the signal strength values will be collected by the equipment and used for the absolute quantification of samples.

2. For this test, six different concentrations of TaqMan probe were used, ranging from 50 to 175 nM (Fig. 2).
3. For each probe concentration, the same sample constituted by 10^4 virus copies/reaction was tested in 8 replicates. This testing revealed the optimal primers and probe concentrations to be used in the laboratory routine.
4. Following optimization, concentrations of 300/300 nM for primers set and 150 nM for the TaqMan probe presented the lowest variation and the best performance on the quantification results.
5. qPCR assays were set up with 0.3 μ M of each primer, 0.15 μ M TaqMan fluorogenic probe, 1× TaqMan® Universal PCR Master Mix (Applied Biosystems), and 5 μ L cDNA template in a final 25 μ L reaction volume.

6. Thermal conditions were 50 °C, 2 min; 95 °C, 10 min and 40 cycles of 95 °C, 15 s and 60 °C, 1 min, in the ABI Prism 7500 (Applied Biosystems, Foster City, CA).
7. For generating the standard curve, an 83 bp amplified cDNA fragment obtained from the NS5 YFV genome region was cloned into TOPO TA Cloning Vector (Invitrogen), according to the manufacturer's instructions (*see Note 13*).
8. Serial dilutions from 10⁷ to 10² plasmid copies per reaction in 100 ng/µL Yeast RNA solution (Invitrogen) were used to construct the calibration curves for the qPCR assays (*see Note 14*).

Routinely, each assay was designed in order to include negative (human serum of healthy individual) and no-template controls (NTC), together with a positive control (human serum of healthy individual spiked with cultivated YFV to reach 10² copies/reaction) (*see Note 15*). Samples to be analyzed should be tested in at least three replicates per run. Standard samples (10² to 10⁷ plasmid copies/reaction) are included in duplicate in every run.

When testing human clinical samples, the reaction was performed in **multiplex** using the TaqMan® RNase P Control Reagents kit (*see Subheading 2.2*). Reaction conditions were the same as described above, with the addition of 0.5× RNase P detection reagent (TaqMan probe VIC/TAMRA).

Individual qPCR assays (**singleplex**) directed to the Exogenous Internal Positive Control EXO-IPC were run in parallel, in the same plate for the testing of in vitro serum-free YFV samples (*see Subheading 2.2*). In this case, YFV samples were spiked with EXO-IPC DNA before processing, to reach a final concentration of (1×) and reaction was conducted according to the manufacturer's specifications. It is expected that homogeneous threshold cycle (Ct) values for detection/quantification of the EXO-IPC DNA are obtained for all analyzed samples.

Alternatively, when using the heterologous cloned sequence as an internal amplification control (IAC), serum-free YFV samples were previously spiked with 5 µL of linearized IAC (40 pg/µL) (*see Subheading 2.2*). This quantity of IAC was chosen because it renders a Ct value around 20, which is positioned in the middle of the linear range of IAC amplification, as reported [14]. Individual qPCR assays (**singleplex**) directed to the IAC reference control were run in parallel, in the same plate for the testing of in vitro serum-free YFV samples.

3.4 Analysis of Quality Parameters

It is strongly recommended the construction of a panel of reconstituted samples containing different concentrations of YFV spiked in serum of healthy individual, to evaluate the performance of the test, twice a year. The panel should contain samples that represent

the broad spectrum of viral load, and should be distributed into aliquots and stored at -80°C . This recommendation is needed to guarantee the quality of the standard curve parameters (linearity of the test, reportable range), and to monitor the reproducibility and efficacy of the qPCR test. The reproducibility and repeatability should be evaluated in three independent runs, in different days, with different operators. The panel of reconstituted samples is also used to determine both, the limit of detection (LOD) and the limit of quantification (LOQ) of the method (*see Note 16*).

To provide an accurate quantification of viral load, specific parameters respective to the dynamic range of the assay were used as acceptance criteria. As for instance, the regression coefficient of linearity (R^2) should be at minimum 0.990; the slope should be comprehended in the range of -3.100 and -3.600 ; the intercept should not surpass 40.00; negative controls and no-template controls (NTC) have to be negative (not detected). The threshold must be fixed in order to standardize the quantification parameters between reactions in different runs; in our protocol the threshold was established in 0.02 (Figs. 3 and 4). Another parameter regards the Ct (threshold cycle) value correspondent to the highest plasmid copy containing the NS5 viral region in the standard curve (10^7 copies per reaction), that should be positioned between Ct 17 and 18 to be considered an acceptable quantification reaction. The baseline was inferred automatically by the equipment.

All reference controls used, either in multiplex or singleplex, should give acceptable signals for human RNase P (mean

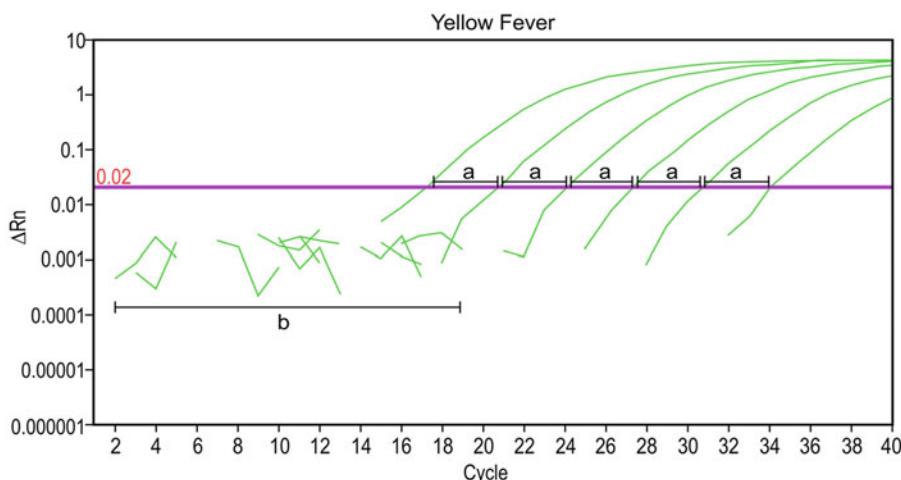


Fig. 3 Yellow fever virus amplification plot. Each curve represents the amplification of different dilution points of the standard curve considering the threshold fixed in 0.02 (purple line). The regular interval between tenfold diluted points demonstrates precise dilution and an accurate standard curve calibration (indicated as “a”). Baseline noise represents unspecific fluorescence signal emission of not detectable negative samples (indicated as “b”)

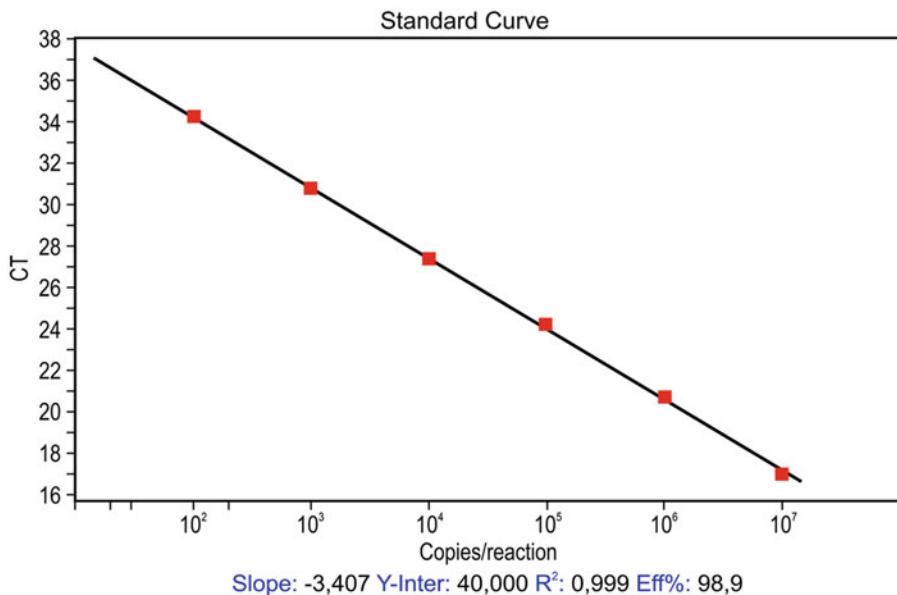


Fig. 4 Optimization of the concentrations of primers and TaqMan probe directed for the NS5 yellow fever virus region. The optimal concentrations are in bold

Ct 30.00 ± 1.65), EXO-IPC DNA (mean Ct 29.38 ± 2.30) and IAC amplification (mean Ct 20.00 ± 1.80), thus indicating that no false negative results were generated by PCR inhibitors or due to the loss or quality of the extracted RNA and subsequent cDNA synthesis, nor were there errors in sample dispensing into the PCR reaction (*see Note 17*).

4 Notes

1. The kit combines the selective binding properties of a silica-based membrane with the speed of microspin and is highly suited for simultaneous processing of many samples. The sample is first lysed under highly denaturing conditions to inactivate RNases to ensure isolation of intact viral RNA. Buffering conditions are then adjusted to promote optimum binding of the RNA to the membrane, and the sample is loaded onto the column. The RNA binds to the membrane, and contaminants are washed away in two steps using two different washing buffers. High-quality RNA is eluted in a special RNase-free buffer, ready for direct use or safe storage. The purified RNA is free of protein, nucleases and other contaminants and inhibitors.
2. Preparation of carrier RNA adding 310 µL AVE Buffer to the tube containing 310 µg lyophilized carrier RNA to obtain a

solution of 1 µg/µL. After, dissolve the carrier RNA thoroughly, divide into 150 µL aliquots, and store at -30 °C to -15 °C. Do not freeze-thaw the aliquots of carrier RNA more than three times. Then, preparation of lysis buffer (**AVL Buffer–carrier RNA**) to be used in the extraction batch is performed calculating the volume of AVL Buffer–carrier RNA solution needed per batch of samples by selecting the number of samples to be simultaneously processed according to the following equation:

$$n \times 0.56 \text{ mL} = y \text{ mL}$$

$$y \text{ mL} \times 10 \text{ } \mu\text{L/mL} = z \text{ } \mu\text{L}$$

where: n = number of samples to be simultaneously processed

y = calculated volume of AVL Buffer

z = volume of carrier RNA–AVE Buffer to be added to the AVL Buffer.

AVL Buffer–carrier RNA should be prepared fresh, and is stable at 2–8 °C for up to 48 h. This solution develops a precipitate when stored at 2–8 °C that must be dissolved by warming at 80 °C before use. Do not warm AVL Buffer–carrier RNA solution more than six times. Do not incubate at 80 °C for more than 5 min. Frequent warming and extended incubation will cause degradation of carrier RNA, leading to reduced recovery of viral RNA. This is particularly the case with low-titer viral samples. Mix by turning over the tube, about 10 times. Do not vortex the tube to avoid foaming. Finally, **AW1 and AW2 Buffer** needs to be prepared: **AW1 Buffer**: Before using for the first time, add 130 mL of ethanol (96–100%). AW1 Buffer is stable for 1 year when stored closed at room temperature (15–25 °C), until the kit expiration date. While **AW2 Buffer** is ready adding 160 mL of ethanol (96–100%) to the AW2 Buffer concentrate.

3. This control aims to ensure the quality of the test and excludes the possibility of false negative due to the presence of eventual inhibitors or the quality and integrity of RNA samples and subsequent cDNAs. The commercial kit contains a set of specific primers and TaqMan probe (VIC/TAMRA) for the detection of human RNase P gene. Alternatively, TaqMan® Exogenous Internal Positive Control Reagents EXO-IPC DNA, is employed for assays using in vitro serum-free samples. The EXO-IPC kit contains a synthetic DNA, without homology with any DNA sequence available on public databanks, together with a set of specific primers and TaqMan probe (VIC/TAMRA). Another option for a nonconstitutive control could be the use of recombinant plasmid containing an inserted sequence of a heterologous extrinsic organism, as for instance

the aquaporin gene of *Arabidopsis thaliana* used as an internal amplification control (IAC), with its respective set of primers and TaqMan probe (see Subheading 2.3). For PCR purposes, the recombinant plasmid is linearized using the restriction enzyme *Pst*I [14].

4. The design of specific oligonucleotides and internal fluorogenic probe is fundamental for the accurate quantification of genetic material. TaqMan probes presenting the same oligonucleotide sequence but labeled with different fluorophores at their 5' extremities can generate distinct fluorescent signals in the analysis of a same sample, resulting in differences on the quantification values. The qPCR assay can be performed as a one-step or two-step reaction, and ideally, primers design should be focused in a conserved region of the viral genome.
5. To assure an efficient lysis process, it is fundamental that the sample is narrowly mixed with the AVL Buffer-carrier RNA to generate a homogeneous solution. Frozen samples that were only been thawed no more than three times can also be used.
6. Incubation should not exceed the recommended incubation time, or it will affect the RNA quality.
7. Ethanol substitution could compromise the results reducing the quantities and quality of purified RNA. Do not use denatured alcohol that contains other substances such as methanol or methylethylketone. For sample volumes larger than 140 µL, the volume of ethanol should be proportionally increased (e.g., a 280 µL sample will require 1120 µL ethanol). To guarantee an efficient binding, it is necessary that the ethanol is completely mixed with the sample to provide a homogeneous solution.
8. All columns must be closed during the spin process to avoid cross-contamination during centrifugation. After centrifugation, if the solution has not entirely passed through the membrane, repeat the centrifugation step at higher speed until all the solution has passed through. Centrifugation is performed at $6000 \times g$ (8000 rpm). Centrifugation at full speed will not affect the efficiency or purity of viral RNA.
9. There is no need to increase the volume of AW1 Buffer even if the sample volume is higher than 140 µL.
10. If rotor vibrates during centrifugation deceleration resulting on flow-through contact between AW2 Buffer and the column, it could cause problems at the elution step, so, in this situation, an additional **step 9** should be performed.
11. To ensure maximum efficiency, the enzyme should be kept on ice.

12. In order to obtain better result and to facilitate pipetting smaller volumes, the reaction mix should be performed in a solely 1.5 μ L tube, for all samples, and then be distributed into 0.2 mL tubes, followed by the individual addition of RNA sample.
13. The standard curve used was derived from plasmid DNA; however there are other approaches such as Synthetic RNA Curve, Armored RNA and Synthetic Gene that can be designed to simultaneously detect more than one target.
14. The standard curve should be prepared in a segregate area to avoid *amplicon* contamination. Prepare aliquots from each point of the standard curve with sufficient volume to be used only once.
15. Each dilution point of the standard curve is applied side by side in duplicate in the plate assay. Positive controls should be positioned in the opposite side of the no-template control (NTC) to avoid contamination.
16. **LOD** (analytical sensitivity)—The lowest concentration of an analyte in a specimen that can be consistently detected (e.g., $\geq 95\%$ of the specimens tested) with precision under routine laboratory conditions and in a defined specimen type. The detection limit is usually below the range of linearity of an assay and could not be assumed as higher than the quantification limit [15]. **LLOQ**—The lower limit of linearity is frequently referred to as the lower limit of quantification (LLOQ) and the upper limit of linearity as the upper limit of quantification (ULOQ). The upper limit of linearity may be restricted by the highest available concentration in a sample or by the saturation of the signal generated by the instrument [15]. **Precision**—Closeness of agreement between independent test/measurement results obtained under stipulated *conditions* [16]. It is important to consider that a measurement may be very precise (replicates have the same result) but not very accurate (the real value is much different). The ideal assay is both precise and accurate. Precision can be established by testing multiple replicates across the dynamic range of the assay in the same run and in different runs and applying suitable statistical tests [15]. **Accuracy**—Accuracy refers to the closeness of the agreement between the result of a single measurement and the value of the analyte [16].
17. Control samples should be prepared and aliquoted to be used only once.

Acknowledgments

This study was funded by grants from FAPERJ (CNE E-26/202.931/2015) and PAEF/IOC-Fiocruz/CNPq (25030.000379/2015-16). C. Britto is a research fellow of CNPq (305589/2015-6). The authors thank the Program for Technological Development in Tools for Health (PDTIS-Fiocruz), Bio-Manguinhos for real-time PCR and DNA sequencing platform facilities, and Heloisa Maria Nogueira Diniz from the Production and Image Processing Service of the Oswaldo Cruz Institute, for preparing the figures.

References

1. Brès PLJ (1986) A century of progress in combating yellow fever. Bull World Health Organ 64:775–786
2. Monath TP (1987) Yellow fever: a medically neglected disease. Report on a seminar. Rev Infect Dis 9:165–175
3. Monath TP (1989) Yellow fever. In: Monath TP (ed) Arboviruses: epidemiology and ecology. CRC Press, Boca Raton, FL, pp 139–231
4. Monath TP, Ballinger ME, Miller BR et al (1989) Detection of yellow fever viral RNA by nucleic acid hybridization and viral antigen by immunocytochemistry in fixed human liver. Am J Trop Med Hyg 40:663–668
5. Domingo C, Patel P, Yillah J et al (2012) Advanced yellow fever virus genome detection in point-of-care facilities and reference laboratories. J Clin Microbiol 50:4054–4060. <https://doi.org/10.1128/JCM.01799-12>
6. World Health Organization (2016) Yellow fever. World Health Organization, Geneva. <http://www.who.int/mediacentre/factsheets/fs100/en/>. Accessed 4 Apr 2017
7. Monath TP (1990) Yellow fever. In: Monath TP (ed) The arboviruses: epidemiology and ecology, vol 5. CRC Press, Boca Raton, FL, pp 139–231
8. WHO (1986) Prevention and control of yellow fever in Africa. World Health Organization, Geneva, pp 1–94
9. Fernandes-Monteiro AG, Trindade GF, Yamamura AM et al (2015) New approaches for the standardization and validation of a real-time qPCR assay using TaqMan probes for quantification of yellow fever virus on clinical samples with high quality parameters. Hum Vaccin Immunother 11:1865–1871. <https://doi.org/10.4161/21645515.2014.990854>
10. Tsai TF (1992) Arboviruses. In: Rose NR, Conway de Marco E, Fahey JL, Friedman H, Renn GM (eds) Manual of clinical laboratory immunology, 4th edn. Am Soc for Microbiol, Washington, DC, pp 606–618
11. Monath TP, Nystrom RR (1984) Detection of yellow fever virus in serum by enzyme immunoassay. Am J Trop Med Hyg 33:151–157
12. Mantel N, Aguirre M, Gulia S et al (2008) Standardized quantitative PCR assays for quantitation of yellow fever and chimeric yellow fever-dengue vaccines. J Virol Methods 151:40–46. <https://doi.org/10.1016/j.jviromet.2008.03.026>
13. Mackay IM, Arden KE, Nitsche A (2002) Real-time PCR in virology. Nucleic Acids Res 30:1292–1305
14. Duffy T, Bisio M, Altcheh J et al (2009) Accurate real-time PCR strategy for monitoring bloodstream parasitic loads in Chagas disease patients. PLoS Negl Trop Dis 3:e419. <https://doi.org/10.1371/journal.pntd.0000419>
15. Burd EM (2010) Validation of laboratory-developed molecular assays for infectious diseases. Clin Microbiol Rev 23:550–576. <https://doi.org/10.1128/CMR.00074-09>
16. International Organization for Standardization (2006) Statistics—vocabulary and symbols, part 1. Probability and general statistical terms. ISO 3534-1. International Organization for Standardization, Geneva, Switzerland



Chapter 7

Evaluation of the Abundance of Fungi in Wastewater Treatment Plants Using Quantitative PCR (qPCR)

Paula Maza-Márquez, Elisabet Aranda, Jesús González-López, and Belén Rodelas

Abstract

Assessment of the abundance of fungi in environmental samples by quantitative PCR (qPCR) of community DNA is often a difficult task due to biases introduced during PCR amplification, resulting from the differences associated with length polymorphism and the varying number of copies of the rRNA operon among fungal species, the lack of specificity of the primers targeting the different regions of the rRNA operon, or their insufficient coverage of the fungal lineages. To overcome those limitations, it is crucial to test and select the specific primers sets which provide the more accurate approximation to the quantification of the targeted fungal populations in a given set of samples. Fungi are a significant fraction of the microbiota in wastewater treatment plants (WWTPs), but the activated sludge microbial communities comprise many other eukaryotic microorganisms whose molecular markers are often coamplified by primers initially designed as fungal-specific. Here, the use of the FungiQuant primer set is recommended for the quantification of fungal molecular markers (18S rRNA genes) by qPCR in activated sludge samples and the full protocol is described.

Key words Wastewater treatment, Fungi, qPCR, Molecular markers, PCR primers

1 Introduction

The microbiota of activated sludge in wastewater treatment plants (WWTPs) comprises populations of both prokaryotic and eukaryotic microorganisms whose abundance, diversity and functions have been thoroughly characterized throughout the last 30 years, particularly after the introduction of cultivation-independent molecular approaches in microbial ecology [1]. The vast majority of the available studies is focused on Bacteria, which are regarded as the most abundant microorganisms in activated sludge and the main actors of organic matter, N and P removal [2], while the analysis of the eukaryotic microorganisms was only seldom approached, and consequently an overview of their importance for the biological wastewater treatment processes is yet missing.

Nonetheless, several studies based in either cultivation-dependent or molecular methods have stated that fungi are a significant fraction of the microbial community in the activate sludge of biological WWTPs based in different technologies, such as conventional activated sludges (CAS), membrane bioreactors (MBR), or A²O systems, where they can significantly contribute to organic matter degradation, particularly of the more recalcitrant molecules, and removal of N through denitrification [3–7].

In recent years, the application of real-time quantitative PCR (qPCR) has become the molecular method of choice for the absolute quantification of prokaryotic and eukaryotic microorganisms in a wide range of natural and engineered environments [8]. However, the accurate estimation of the abundance of fungal populations by qPCR is currently limited by several factors. The major difficulties are associated with the selection of the most suitable molecular marker and the design of primers for PCR amplification displaying enough specificity and coverage for all the lineages of fungi [9, 10]. The sequences encoding the ribosomal rRNAs are the most frequently used molecular markers for the phylogenetic identification of both prokaryotic and eukaryotic microorganisms [11]. The genes for the rRNAs of the small (18S) and large (25–28S) subunits and the internal transcribed spacer (ITS) region, comprising the two highly variable ITS1 and ITS2 regions and the more conserved 5.8S rRNA gene (Fig. 1), have been thoroughly tested as fungal molecular markers for fungi; however, mycologists did not yet designate a standard universal DNA marker for fungi meeting all the ideal requirements [11]. The ITS1 and ITS2 regions do not code for ribosome components and are highly variable among fungi of even closely related species, and for these reasons were proposed as the primary fungal DNA barcode markers [12]; however, this selection is controversial by several facts. Intraspecific variability is a limitation for the identification of closely related species, particularly within the *Ascomycota*, which hampers the characterization of the fungal communities from environmental DNA [13, 14]. In addition, the ITS1 and ITS2 regions display length polymorphism, which may introduce PCR amplification biases toward the shorter sequences [15].

Many primers targeting the fungal rRNA markers (18S, 25–28S, and ITS regions) have been designed, evaluated, and compared since the 1990s, including universal primers aimed for the coverage of all fungal lineages and Phylum-specific primers for the selective amplification of *Ascomycota* or *Basidiomycota* sequences [9, 16–23]. However, it must be taken into account that not all of these primers have been tested for qPCR on environmental DNA samples, and very few were specifically designed for such purpose. Several studies revealed that some of these primer

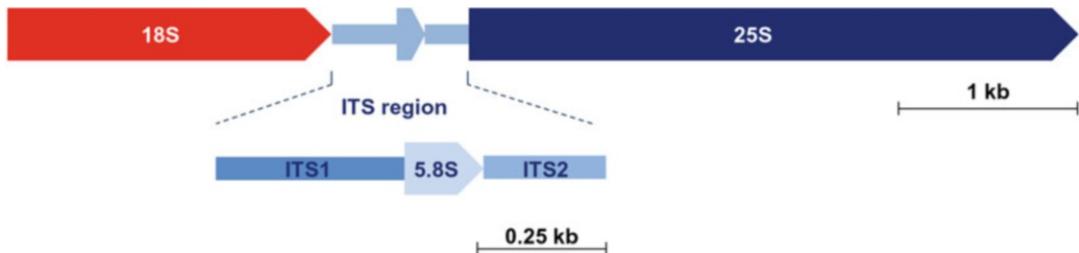


Fig. 1 Map showing the structure of the locus RDN37-1 of chromosome XII of *Saccharomyces cerevisiae* strain S288C (sequence available at the *Saccharomyces* Genome Database website, <http://www.yeastgenome.org>), illustrating the locations of the 5.8S, 18S, and 25S rRNA genes and the internal transcribed spacer (ITS) region

pairs, particularly those targeting the 18S rRNA, amplify sequences from nonfungal eukaryotic organisms such as plants, metazoa, protozoa, and other protists [3, 10, 24, 25]. It must be also taken into consideration that quantitative biases can be introduced due to the varying number of copies of the ribosomal operon among fungal species, which ranges between 10 and 200 [26].

When the abundance of the fungal populations is evaluated in environmental DNA samples by qPCR, coamplification of DNA from nontarget organisms is a major source of quantitative bias, which can lead to huge overestimations [3]. Consequently, the selection of primers becomes a key step when the qPCR approach is applied on any set of environmental DNA samples, and it is recommended to comparatively evaluate the performance of different primer sets [3]. This is an issue of particular importance when fungi are targeted in activated sludge samples, since the microbial community is rich in eukaryotic microorganisms such as ciliated protozoa, which carry a very large number of copies of the rRNA operons and whose rRNA molecular markers are often coamplified by primers initially designed as fungal-specific [3].

Studies addressing the absolute abundance of fungi in WWTPs by means of qPCR are scarce. The molecular markers and primer pairs which have been previously tested are summarized in Table 1. In a recent study, Maza-Márquez et al. [3] comparatively evaluated the abundance of fungal populations using three different sets of fungal-specific universal primers pairs: NS1-Fung, FungiQuant (targeting the 18S rRNA gene), and ITS3-ITS4 (targeting the ITS2 region). The results of the study revealed differences up to 2 orders of magnitude of the numbers of copies of fungal markers per liter activated sludge or g total suspended solids depending on the primer set used, showing that primers NS1-Fung and ITS3-ITS4, previously known to coamplify DNA of nonfungal organisms [10, 24, 25, 34], widely overestimate the abundance of fungi. The FungiQuant primer set was specifically designed for the evaluation of the abundance of fungi by qPCR by Liu et al. [19], and its coverage and specificity have been validated in several other studies

Table 1

Universal fungal primers sets used for the quantification of the abundance of fungi by means of qPCR in WWTP samples

Molecular marker	Primer pairs		References of primer design	References of applications for qPCR of fungi
	Name	Sequence (5'-3')		
18S rRNA gene	NS1 Fung	GTAGTCATATGCTTGTCTC ATTCCCCGTTACCCGTTG	May et al. [27]	[3, 28]
18S rRNA gene	FungiQuant-F FungiQuant-R	GGRAAACTCACCAAGGTCCAG GSWCTATCCCCAKCACGA	Liu et al. [19]	[3, 4, 29, 30]
18S rRNA gene	LR3 LR0R	CCGTGTTCAAGACGGG ACCCGCTGAACITTAAGC	Penton et al. [31]	[4]
18S rRNA gene	FF390 FR1	CGATAACGAACGAGACCT AICCATTCAATCGGTAIT	Vainio and Hantula [32]	[7]
ITS2 spacer	ITS-3 ITS-4	GCATCGATGAAGAACGCAGC TCCTCCGCTTATTGATATGC	White et al. [23]	[3]
ITS2 spacer	gITS7 ITS-4	GTGARTCATCGARTCTTG TCCTCCGCTTATTGATATGC	Ihrmark et al. [18] White et al. [23]	[33]

[35, 36]. In this Chapter, the full protocol for quantification of fungal molecular markers (18S rRNA genes) in activated sludge samples by means of qPCR based in the FungiQuant primer set is described.

2 Materials

2.1 Activated Sludge Sample Collection

1. 200 ml of activated sludge (*see Note 1*).
2. Disposable sterile plastic bottles (250 ml).
3. Ice bucket.
4. Microcentrifuge tubes, 2.0 ml.
5. Tabletop microcentrifuge.

2.2 Extraction and Purification of DNA (See Note 2)

1. FastPrep24® benchtop homogenizer (MP-BIO, USA) for the lysis of biological samples and extraction of nucleic acids.
2. FastDNA®-2 ml SPIN Kit for Soil (MP-BIO, USA).
3. Agarose for molecular biology (low EEO).
4. TBE (Tris–borate–EDTA) electrophoresis buffer, molecular biology grade. TBE buffer can be either purchased ready to use, or prepared as a 10× stock solution containing 1 M Tris

base, 1 M boric acid, and 0.02 M Na₂EDTA in molecular biology grade water [37]. Dilute 100 ml to 1 L to prepare the working concentration for electrophoresis.

5. NanoDrop® ND-1000 spectrophotometer (Thermo Scientific, USA) or equivalent instrument for the quantification of nucleic acids.
1. Genomic DNA of *Candida albicans* strain ATCC 10231 (*see Note 3*).
2. Hot Start Taq DNA polymerase (5 U/μl) with 10× reaction buffer (with KCl).
3. 25 mM MgCl₂.
4. 10 nM deoxynucleotide (dNTP) solution mix.
5. 20 mg/ml bovine serum albumin (BSA), molecular biology grade.
6. Molecular biology grade water (nuclease- and protease-free water).
7. 10 mM FungiQuant primers (forward and reverse), in molecular biology grade water (Table 1) (*see Note 4*).
8. Standard PCR thermal cycler (in the experiments described here, an Eppendorf® Mastercycler thermal cycler (Eppendorf, Germany, was used).
9. TOPO® TA cloning® system with One Shot® TOP10 chemically competent *Escherichia coli* cells (Invitrogen, Life Technologies Corporation, USA), or equivalent commercial cloning system.
10. LB (Luria–Bertani) agar plates, amended with 50 μg/ml ampicillin (*see Note 5*).
11. 40 mg/ml of 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-gal) in dimethyl formamide.
12. LB broth vials (10 ml) amended with 50 μg/ml ampicillin (*see Note 5*).
13. Plasmid DNA purification kit (in the experiments described here, the WideUSE® Plasmid Purification Kit (Canvax, Spain) was used).
14. HindIII restriction endonuclease (20,000 U/ml).

2.4 Detection Chemistry for qPCR Experiments

1. 10,000× SYBR Green fluorescent dye.
2. Dimethyl sulfoxide (DMSO).
3. SYBR Green at a final concentration of 20× in DMSO (*see Note 6*).

2.5 qPCR

1. Hot Start Taq DNA polymerase (5 U/ μ l) with 10 \times reaction buffer (with KCl).
2. 25 mM MgCl₂.
3. 10 mM deoxynucleotide (dNTP) solution mix.
4. 20 mg/ml bovine serum albumin (BSA), molecular biology grade.
5. Molecular biology grade water (nuclease- and protease-free water).
6. 10 mM FungiQuant primers (forward and reverse), in molecular biology grade water (Table 1) (*see Note 4*).
7. MicroAmp® Optical 96-Well Reaction Plate (0.2 ml) and adhesive covers (Thermo Scientific, USA).
8. Tabletop microplate centrifuge.
9. QuantStudio® 3 Real-Time PCR System (Thermo Scientific, USA).

3 Methods

Work under aseptic conditions, using powder-free gloves and a flow hood cabinet to minimize sample contamination, and maintain all the reactives in ice (unless indicated otherwise).

3.1 Activated Sludge Sample Collection

1. Add 2 ml of each homogenized activated sludge sample to a 2-ml microcentrifuge tube.
2. Centrifuge the tubes at 14,000 \times g for 1 min and discard the supernatants.
3. Add another additional 2 ml of each of the activated sludge samples to the corresponding tubes and repeat step 2, in order to collect the solids from a total volume of 4 ml of each sample in a single tube.
4. Proceed immediately with the genomic DNA extraction step.

3.2 Extraction of Genomic DNA

1. Add the lysing matrix beads suspension of the FastDNA®-2 ml SPIN Kit for Soil to the pellets of activated sludge samples obtained in Subheading 3.1.
2. Follow the manufacturer's protocol for DNA extraction. After cell wall lysis and removal of the insoluble cellular material, purify the nucleic acids using a silica-based procedure, using the spin filters provided with the kit. After finishing the extraction, write down the exact final volume of eluted DNA yielded from each 4 ml activated sludge sample, since it may be needed for further calculations.

3. Check the environmental DNA integrity by electrophoresing 3 µl subsamples of the purified DNA in an 1% agarose gel, prepared by melting agarose in 1× TBE electrophoresis buffer.
4. Measure DNA quality and concentration using 1 µl subsamples of the eluted DNA, using the NanoDrop® ND-100 spectrophotometer. Concentration is given by the OD measure at 260 nm. An OD 260/280 ratio of approximately 1.8 indicates a good quality of the purified DNA [38].

3.3 Construction of Plasmids Carrying Fungal Molecular Markers (To Be Used as Quantification Standards)

1. Use the FungiQuant primers (Table 1) to amplify the target 18S rRNA sequence from the *Candida albicans* ATCC 10231 genomic DNA. Use 50 ng–100 ng of template DNA by appropriately diluting the genomic DNA stock in molecular biology grade water, and add it to a PCR reaction mixture containing 0.125 µl of Hot-start Taq DNA polymerase (5 U/µl), 0.5 µl of each primer (10 µM), 2.5 µl of 10× Taq Buffer (with KCl), 1.5 µl MgCl₂ (25 mM), 0.5 µl of dNTPs (10 µM), 0.0625 µl of BSA (20 mg/ml), and molecular biology grade water up to a final volume of 25 µl (see Table 3). Include a negative control reaction lacking template DNA.
2. Place the tubes into a standard PCR thermal cycler and run a thermal cycling program with the following conditions: initial denaturation at 94 °C for 3 min, followed by 35 amplification cycles (denaturation at 94 °C for 30 s, annealing at 62 °C for 30 s, extension at 72 °C for 45 s), and a final elongation step at 72 °C for 10 min.
3. After the PCR amplification, check that the PCR products contain a single band of the expected amplicon size (351 bp), by running a subsample of 3 µl in a 2% agarose gel (see Note 7).
4. Clone the resulting pool of amplicons into the pCR2.1-TOPO® vector provided by the TOPO® TA cloning® system, then transform the competent *E. coli* cells with the cloning reaction, following the manufacturer's guidelines.
5. Select the *E. coli* clones carrying plasmids with inserts by a white/blue screening, by plating the transformation in an LB-ampicillin agar plate, previously spread with 40 µl of 40 mg/ml of X-gal using a sterile Drigalski spatula. At least two different volumes of the transformation should be used, to ensure well-isolated colonies on at least one plate.
6. Incubate at 37 °C for 24 h, select at least ten isolated white or light-blue colonies, and transfer each of them into a fresh LB-ampicillin agar plate and an LB-ampicillin broth vial. Discard all dark-blue colonies (carrying no inserts).
7. Incubate the plates and vials at 37 °C for 24 h.

8. Isolate plasmid DNA from each LB-ampicillin broth culture, using the WideUSE® Plasmid Purification Kit or equivalent, following the manufacturer's protocol.
9. Check the integrity of the isolated plasmid DNAs, running 3-μl subsamples in a 1% agarose gel.
10. Verify the plasmids by via Sanger sequencing analysis or diagnostic restriction digest. Select one of the *E. coli* clones carrying the plasmid with the expected insert sequence as standard. The *E. coli* cells can be long-term preserved in 20% glycerol stocks at -80 °C.
11. Linearize the selected standard plasmid DNA (60 μl) by restriction digest with the HindIII endonuclease, using the buffer provided with the enzyme and following the manufacturer's guidelines. HindIII was selected because the cloned fragment does not include its target sequence, but it can be substituted by other endonucleases whose target sequences are also unique to the multicloning site of the vector.
12. Check that the plasmid DNA is fully linearized, by running a 3-μl subsample in a 1% agarose gel. A single band of the appropriate length (vector + insert) must be observed.
13. Purify the linearized plasmid DNA using phenol-chloroform (1:1, v/v). Transfer the plasmid DNA to a microcentrifuge tube, and add an equal volume of phenol-chloroform. Mix the content of the tube by vortexing 20 s and centrifuge the mixture 5 min at 14,000 × *g*. Carefully transfer the aqueous phase containing the purified DNA to a clean tube. Discard the interface and organic phase.
14. Measure the quality and concentration of the purified linearized plasmid standard DNA by spectrophotometry using NanoDrop® ND-100, as described in step in Subheading 3.2.
15. Calculate the number of copies per μl of purified linearized standard plasmid DNA, as described previously [39]:

$$\text{No.of copies per } \mu\text{l} = (n \times 6.023 \times 10^{23}) / (L \times 1.0 \times 10^9 \times 650)$$

where *n* is the mass of linearized standard plasmid DNA in nanograms (measured in Subheading 3.3, step 12) and *L* is the length of the linearized standard plasmid DNA in base pairs (length of vector plus insert). A free online tool for these calculations is available (<https://cels.uri.edu/gsc/cndna.html>).

3.4 Preparation of Serial Dilutions of the Standard Plasmid

1. Once the number of molecules per μl of linearized standard plasmid DNA is known, prepare a dilution at a concentration of 10⁸ molecules per μl, which will carry the same amount of copies of the target 18S rRNA amplicon. At least 8 points of

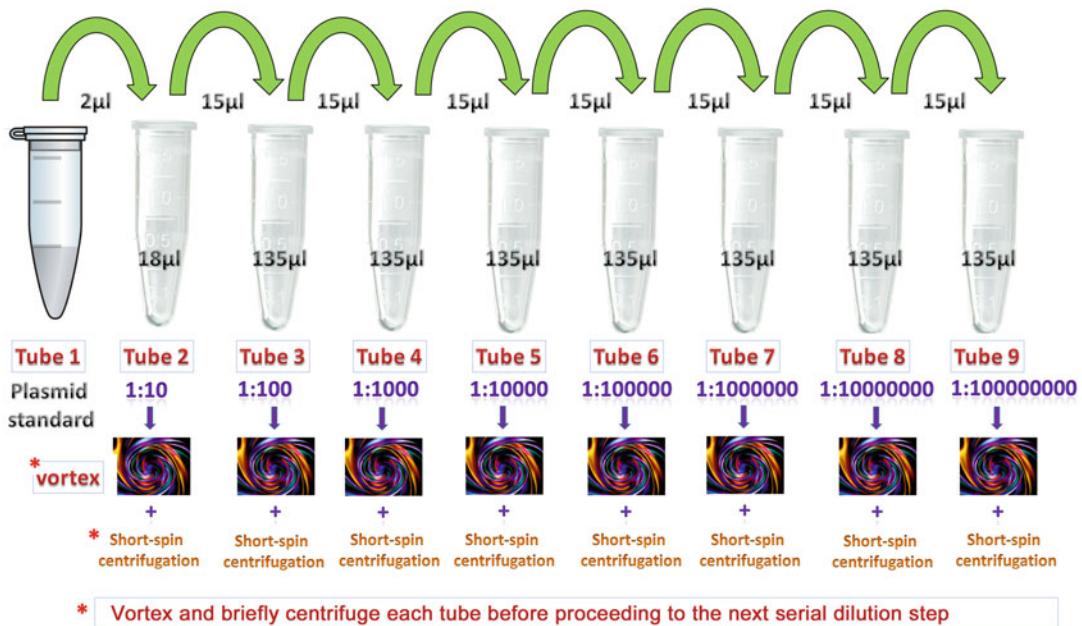


Fig. 2 Serial dilution scheme for the construction of qPCR standard curves

serial dilutions of the linearized standard plasmid DNA are needed in order to construct the standard calibration curve. The serial dilution scheme is depicted in Fig. 2. All the dilutions must be triplicated (*see Note 8*).

2. To prepare the 10^{-1} dilution: add 2 µl of linearized standard plasmid DNA (10^8 copies per µl, tube 1) into a new tube (tube 2) with 18 µl nuclease-free water, gently vortex to mix, and centrifuge briefly in order to collect the volume at the bottom of the tube.
3. To prepare the 10^{-2} dilution: add 15 µl from tube 2 to tube 3 with 135 µl nuclease-free water, gently vortex, and centrifuge briefly.
4. To prepare the 10^{-3} dilution: add 15 µl from tube 3 to tube 4 with 135 µl nuclease-free water, gently vortex, and centrifuge briefly.
5. To prepare the 10^{-4} dilution: add 15 µl from tube 4 to tube 5 with 135 µl nuclease-free water, gently vortex, and centrifuge briefly.
6. To prepare the 10^{-5} dilution: add 15 µl from tube 5 to tube 6 with 135 µl nuclease-free water, gently vortex, and centrifuge briefly.
7. To prepare the 10^{-6} dilution: add 15 µl from tube 6 to tube 7 with 135 µl nuclease-free water, gently vortex, and centrifuge briefly.

8. To prepare the 10^{-7} dilution: add 15 µl from tube 7 to tube 8 with 135 µl nuclease-free water, gently vortex, and centrifuge briefly.
9. To prepare the 10^{-8} dilution: add 15 µl from tube 8 to tube 9 with 135 µl nuclease-free water, gently vortex, and centrifuge briefly.

3.5 Preparation of Dilutions of the DNA Extracted from the Activated Sludge Samples

1. At least two different dilutions of the DNA samples (1/10 and 1/20) must be prepared (see Note 9). All the dilutions must be triplicated (see Note 8).
2. To prepare the 1/10 dilution: add 1 µl of DNA sample to 9 µl nuclease-free water, gently vortex, and centrifuge briefly.
3. To prepare the 1/20 dilution: add 1 µl of DNA sample to 19 µl nuclease-free water, gently vortex, and centrifuge briefly.

3.6 Preparation of the qPCR Master Mix

1. Prepare a qPCR master mix for a final reaction volume of 25 µl per tube, using the reagents and volumes shown in Table 2 (see Note 10).
2. Mix well the master mix by vortexing or using the force of pipetting, and briefly centrifuge it to collect all the volume at the bottom of the tube.

3.7 qPCR Experiment

1. Add 23 µl of the qPCR master mix to all wells needed of the MicroAmp Optical 96-Well Reaction Plate, and subsequently add 2 µl of each DNA sample dilution to quantify or standard linearized plasmid dilution to the corresponding wells. Include at least three negative control wells lacking DNA template.
2. Seal the plate using the provided adhesive film covers and centrifuge it briefly in a tabletop microplate centrifuge.
3. Place the plate into the QuantStudio® 3 Real-Time PCR System and start the thermal cycling program described in Table 3.

Table 2
qPCR cycling conditions used for the quantification of the abundance of Fungi in activated sludge samples using the FungiQuant primer set

Initial denaturalization	95 °C, 3 min
Denaturalization	94 °C, 30 s
Primer annealing	62 °C, 30 s
Elongation	72 °C, 45 s
Melting curve	60 °C–95 °C + 2 °C/min Fluorescence measured each 15 s
Hold	25 °C

Table 3
qPCR reaction components

Reactives	Volume per single reaction tube (μl)	Volume for a master mix for ca. 85 samples (μl) ^a
10× Taq buffer (with KCl)	2.5	250
MgCl ₂ (25 mM)	1.5	150
dNTPs (10 mM)	0.5	50
Bovine serum albumin (20 mg/ml)	0.0625	6.25
SYBR Green I diluted in DMSO (20×)	0.125	12.5
Forward FungiQuant primer (10 μM)	0.5	50
Reverse FungiQuant primer (10 μM)	0.5	50
Hot Start Taq DNA polymerase (5 U/ μl)	0.125	12.5
DNA template (unknown samples)	2	200
Molecular biology grade water (DNase-/RNase/proteases-free water)	17.19	1718.75

^aPlease see Note 10

4. Once the experiment is finished, analysis of the PCR products by electrophoresis in a 2% agarose gel (see Fig. 3) and melting curve analysis are recommended, in order to assure specificity of the amplification and quantification reliability (a single band in the gel and a single melting peak in the curve).
5. Plot the quantification cycle (C_q) values against the logarithm of the numbers of copies of template of the dilutions of the standard DNA, in order to construct the standard curve (see Note 11). At least one of the dilutions used of the DNA samples to be quantified must fall within the range of the standard curve.
6. For further reading on the guidelines to generate consistent and reproducible data from qPCR experiments and the minimum information that must be reported in publications, see Bustin et al. [40].
7. The number of copies of fungal markers in the samples can be expressed as copies per L of activated sludge by the following calculation:

$$\text{No.of copies per } L \text{ activated sludge} = X \times D \times V \times 125$$

where X is the No. of copies of 18S rRNA measured in the 2 μl DNA samples by qPCR, D is the dilution factor of the sample (10 or 20), and V is the volume in μl of DNA extracted from each 4 ml activated sludge sample.

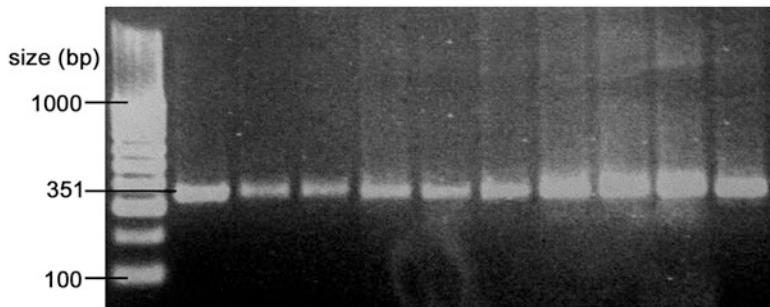


Fig. 3 Amplicons of the fungal 18S rRNA gene generated with the FungiQuant primer pair (351 bp) by qPCR, electrophoresed in a 2% agarose gel

4 Notes

1. Homogenized activated sludge samples must be collected at the WWTP facility using sterile bottles. Samples must be carried at 4 °C to the laboratory, after which they must be immediately processed.
2. The extraction of genomic DNA from environmental samples is a critical step in molecular ecology approaches, in order to achieve a true representation of the diversity or abundance of the targeted community [41]. The FastDNA®-2 ml SPIN Kit for Soil and the FastPrep®24 apparatus are strongly recommended for nucleic acids extraction and purification of activated sludge samples [42]. The FastPrep®24 instrument homogenizes samples using mechanical force (bead-beating) combined with chemical reagents, delivering high yields of genomic DNA. The FastDNA®-2 ml SPIN Kit for Soil provides 2.0 ml tubes for sample homogenization containing Lysing Matrix E®, a mixture of ceramic and silica particles designed to efficiently disrupt the cells of all types of soil microorganisms, including those regarded as difficult to lyse such as gram-positive bacteria, endospores, yeast, algae, nematodes, and fungi (https://www.mpbio.com/includes/protocol/FastDNA_Spin_Kit_for_soil.pdf). The kit is designed for soil but fits the purpose for activated sludge samples, since all the aforementioned organisms are common part of the activated sludge microbiota [2]. Several types of lysing matrices with different recommendations of use depending on the nature of the samples and the target organisms are available from the manufacturer (https://www.mpbio.com/index.php?cPath=2_77_425). In the final step, the genomic DNA is cleaned by a silica-based procedure (GeneClean SPIN), providing an efficient and high yield of purified genomic DNA.

3. Genomic DNA of *Candida albicans* strain ATCC 10231 was used for the construction of the plasmid standard carrying the targeted 18S rDNA amplicon. Alternatively, genomic DNA of other collection strains can be used.
4. The stock solution of the primers must be prepared at a concentration of 100 mM in small volumes and stored at –20 °C. Prepare the working solution at a concentration of 10 mM by appropriately diluting the stock, using molecular biology grade water. Keep the stock solution in ice throughout the process and avoid subjecting it to multiple freeze–thaw cycles.
5. SYBR Green I diluted in dimethyl sulfoxide DMSO (20×) is a dye which emits fluorescence only when it binds to double-stranded DNA, and is used for the quantification of amplicons generated during qPCR, since fluorescence measured by the qPCR instrument increases after each PCR cycle as the number of copies increases. It must be stored at –20 °C protected from light, that is, by using aluminum foil.
6. LB broth contains 10 g peptone, 5 g yeast extract, and 10 g sodium chloride per L of distilled water [43]. LB agar is made by adding 15 g agar per 1 L of medium before autoclaving. Ampicillin is prepared as a concentrated stock in distilled water and added using a 0.22 µm pore-size membrane filter in the appropriate amount for a final concentration of 50 µg/ml, after autoclaving and cooling the media at 45 °C. Thoroughly mix the agar medium before pouring the plates and leave them at room temperature until agar solidifies. Ensure that the medium surface is completely dry before using the plates.
7. Agarose gel electrophoresis is a simple and reliable method to check the specificity of PCR amplification, allowing for the detection of nonspecific amplicons or primer dimers. If a single DNA band is not observed, the PCR reaction must be optimized. Troubleshooting if primers dimers occur: increase the amount of template. Troubleshooting if nonspecific bands occur: increase the temperature of annealing. If a single band cannot be obtained after trying optimization, the desired PCR amplicon can be purified using a DNA gel extraction kit.
8. Accurate pipetting is essential in this step. Prepare at least three repetitions for each standard sample or activated sludge DNA sample, in order to generate statistically representative datasets.
9. To ensure that the fluorescence generated during the amplification of the activated sludge DNA samples falls within the range of the standard quantification curve, it is recommended to prepare and process at least two different dilutions.
10. About 15% more than the needed volume of master mix must be prepared, since some loss by pipetting occurs.

11. The quantification cycle (C_q) or threshold cycle (C_t) value is the number of PCR cycles required by each sample to emit fluorescence over a given threshold level. For accurate and reproducible results, the threshold must be set at the level where the amplification rate is higher during the exponential phase [44].

Acknowledgments

This work was supported by Junta de Andalucía, Plan Andaluz de Investigación (Environmental Microbiology Group, RNM-270).

References

1. Ferrera I, Sánchez O (2016) Insights into microbial diversity in wastewater treatment systems: how far have we come? *Biotechnol Adv* 3:790–802. <https://doi.org/10.1016/j.biotechadv.2016.04.003>
2. Seviour RJ, Nielsen PH (2010) Microbial communities in activated sludge plants. In: Seviour RJ, Nielsen PH (eds) *Microbial ecology of activated sludge*, 2nd edn. IWA Publishing, London, pp 95–125
3. Maza-Márquez P, Vílchez-Vargas R, González-Martínez A et al (2018) Assessing the abundance of fungal populations in a full-scale membrane bioreactor (MBR) treating urban wastewater by using quantitative PCR (qPCR). *J Environ Manag* 223:1–8. <https://doi.org/10.1016/j.jenvman.2018.05.093>
4. Gallardo-Altamirano MJ, Maza-Márquez P, Peña-Herrera JM et al (2018) Removal of anti-inflammatory/analgesic pharmaceuticals from urban wastewater in a pilot-scale A^2O system: linking performance and microbial population dynamics to operating variables. *Sci Total Environ* 643:1481–1492. <https://doi.org/10.1016/j.scitotenv.2018.06.284>
5. Liébana R, Arregui L, Belda I et al (2015) Membrane bioreactor wastewater treatment plants reveal diverse yeast and protist communities of potential significance in biofouling. *Biofouling* 31:71–82. <https://doi.org/10.1080/08927014.2014.998206>
6. Yang Q, Angly FE, Wang Z et al (2011) Wastewater treatment systems harbor specific and diverse yeast communities. *Biochem Eng J* 58–59:168–176
7. Yang Q, Wang J, Wang H et al (2012) Evolution of the microbial community in a full-scale printing and dyeing wastewater treatment system. *Bioresour Technol* 117:155–163. <https://doi.org/10.1016/j.biortech.2012.04.059>
8. Smith CJ, Osborn AM (2009) Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiol Ecol* 67:6–20. <https://doi.org/10.1111/j.1574-6941.2008.00629.x>
9. Prévost-Bouré NC, Christen R, Dequiedt S et al (2011) Validation and application of a PCR primer set to quantify fungal communities in the soil environment by real-time quantitative PCR. *PLoS One* 6(9):e24166. <https://doi.org/10.1371/journal.pone.0024166>
10. Wurzbacher C, Rösel S, Rychla A et al (2014) Importance of saprotrophic freshwater fungi for pollen degradation. *PLoS One* 9(4):e94643. <https://doi.org/10.1371/journal.pone.0094643>
11. Lindahl BD, Nilsson RH, Tedersoo L et al (2013) Fungal community analysis by high-throughput sequencing of amplified markers—a user's guide. *New Phytol* 199:288–299. <https://doi.org/10.1111/nph.12243>
12. Schöch CL, Seifert KA, Huhndorf S et al (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc Natl Acad Sci U S A* 109:6241–6246. <https://doi.org/10.1073/pnas.1117018109>
13. Badotti F, de Oliveira FS, Garcia CF et al (2017) Effectiveness of ITS and sub-regions as DNA barcode markers for the identification of Basidiomycota (Fungi). *BMC Microbiol* 17:42. <https://doi.org/10.1186/s12866-017-0958-x>
14. Kiss L (2012) Limits of nuclear ribosomal DNA internal transcribed spacer (ITS) sequences as species barcodes for Fungi. *Proc*

- Natl Acad Sci U S A 109:E1811. <https://doi.org/10.1073/pnas.1207143109>
15. Bellemain E, Carlsen T, Brochmann C et al (2010) ITS as an environmental DNA barcode for fungi: an *in silico* approach reveals potential PCR biases. BMC Microbiol 10:189. <https://doi.org/10.1186/1471-2180-10-189>
 16. Anderson IC, Cairney JWG (2004) Diversity and ecology of soil fungal communities: increased understanding through the application of molecular techniques. Environ Microbiol 6:769–779. <https://doi.org/10.1111/j.1462-2920.2004.00675.x>
 17. Bokulich NA, Mills DA (2013) Improved selection of internal transcribed spacer-specific primers enables quantitative, ultra-high-throughput profiling of fungal communities. Appl Environ Microbiol 79:2519–2526. <https://doi.org/10.1128/AEM.03870-12>
 18. Ihrmark K, Bödeker ITM, Cruz-Martinez K et al (2012) New primers to amplify the fungal ITS2 region-evaluation by 454-sequencing of artificial and natural communities. FEMS Microbiol Ecol 82:666–677. <https://doi.org/10.1111/j.1574-6941.2012.01437.x>
 19. Liu CM, Kachur S, Dwan MG et al (2012) FungiQuant: a broad-coverage fungal quantitative real-time PCR assay. BMC Microbiol 12:255. <https://doi.org/10.1186/1471-2180-12-255>
 20. Martin KJ, Rygiewicz PT (2005) Fungal-specific PCR primers developed for analysis of the ITS region of environmental DNA extracts. BMC Microbiol 5:28. <https://doi.org/10.1186/1471-2180-5-28>
 21. Mitchell JI, Zuccaro A (2006) Sequences, the environment and fungi. Mycologist 20:62–74
 22. Taylor DL, Walters WA, Lennon NJ et al (2016) Accurate estimation of fungal diversity and abundance through improved lineage-specific primers optimized for Illumina amplicon sequencing. Appl Environ Microbiol 82 (24):7217–7722. <https://doi.org/10.1128/AEM.02576>
 23. White TJ, Bruns T, Lee S et al (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) PCR protocols: a guide to methods and applications. Academic Press, New York, pp 315–322
 24. Evans TN, Seviour RJ (2012) Estimating biodiversity of fungi in activated sludge communities using culture-independent methods. Microb Ecol 63:773–786. <https://doi.org/10.1007/s00248-011-9984-7>
 25. Hoshino YT, Morimoto S (2010) Soil clone library analyses to evaluate specificity and selectivity of PCR primers targeting fungal 18S rDNA for denaturing-gradient gel electrophoresis (DGGE). Microbes Environ 25:281–287
 26. Wallander H, Ekblad A, Godbold DL et al (2007) Evaluation of methods to estimate production, biomass and turnover of ectomycorrhizal mycelium in forests soils - a review. Soil Biol Biochem 57:1034–1047
 27. May LA, Smiley B, Schmidt MG (2001) Comparative denaturing gradient gel electrophoresis analysis of fungal communities associated with whole plant corn silage. Can J Microbiol 47:829–841
 28. Niu L, Li Y, Xu L et al (2017) Ignored fungal community in activated sludge wastewater treatment plants: diversity and altitudinal characteristics. Environ Sci Pollut Res 24:4185–4193. <https://doi.org/10.1007/s11356-016-8137-4>
 29. González-Martínez A, Siivonen M, Muñoz-Palazón B et al (2018) Microbial ecology of full-scale wastewater treatment systems in the polar Arctic circle: Archaea. Bacteria and Fungi. Sci Rep 8:2208. <https://doi.org/10.1038/s41598-018-20633-5>
 30. González-Martínez A, Muñoz-Palazón B, Maza-Márquez P et al (2018) Performance and microbial community structure of a polar Arctic circle aerobic granular sludge system operating at low temperature. Bioresour Technol 256:22–29. <https://doi.org/10.1016/j.biortech.2018.01.147>
 31. Penton CR, Louis DS, Cole JR et al (2013) Fungal diversity in permafrost and tallgrass prairie soils under experimental warming conditions. Appl Environ Microbiol 79:7063–7072. <https://doi.org/10.1128/AEM.01702-13>
 32. Vainio EJ, Hantula J (2000) Direct analysis of wood-inhabiting fungi using denaturing gradient gel electrophoresis of amplified ribosomal DNA. Mycological Res 104:927–936
 33. Wei Z, Liu Y, Feng K et al (2018) The divergence between fungal and bacterial communities in seasonal and spatial variations of wastewater treatment plants. Sci Total Environ 628–629:969–978. <https://doi.org/10.1016/j.scitotenv.2018.02.003>
 34. Maza-Márquez P, Vilchez-Vargas R, Kerckhof FM et al (2016) Community structure, population dynamics and diversity of fungi in a full-scale membrane bioreactor (MBR) for urban wastewater treatment. Water Res 105:507–519. <https://doi.org/10.1016/j.watres.2016.09.021>

35. Biyeyeme Bi Mve MJ, Cloutier Y, Lacombe N et al (2017) Comparison of methods to evaluate the fungal biomass in heating, ventilation, and air-conditioning (HVAC) dust. Environ Monit Assess 189:8. <https://doi.org/10.1007/s10661-016-5682-8>
36. Wymore AS, Compson ZG, Liu CM et al (2013) Contrasting rRNA gene abundance patterns for aquatic fungi and bacteria in response to leaf-litter chemistry. Freshw Sci 32:663–672
37. TBE electrophoresis buffer (10×) (2010) Cold Spring Harb Protoc. <http://cshprotocols.cshlp.org/content/2010/6/pdb.rec12231.full>
38. Thermo Scientific (2009) T042-Technical Bulletin NanoDrop® Spectrophotometers. <http://www.nhm.ac.uk/content/dam/nhmwww/our-science/dpts-facilities-staff/Coreresearchlabs/nanodrop.pdf>
39. Jackson GA, Livingston RS, Riley LK et al (2013) Development of a PCR assay for the detection of *Spiromyces muris*. J Am Assoc Lab Animal Sci 52:165–170
40. Bustin SA, Benes V, Garson JA et al (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55:611–622. <https://doi.org/10.1373/clinchem.2008.112797>
41. Maza-Márquez P, González-Martínez A, Rodelas B et al (2017) Full-scale photobioreactor for biotreatment of olive washing water: structure and diversity of the microalgaebacteria consortium. Bioresour Technol 238:389–398. <https://doi.org/10.1016/j.biotech.2017.04.048>
42. Guo F, Zhang T (2013) Biases during DNA extraction of activated sludge samples revealed by high throughput sequencing. Appl Microbiol Biotechnol 97:4607–4616
43. Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor, New York
44. Forlenza M, Kaiser T, Savelkoul HFJ et al (2012) The use of real-time quantitative PCR for the analysis of cytokine mRNA levels. Methods Mol Biol 820:7–23. https://doi.org/10.1007/978-1-61779-439-1_2.



Chapter 8

Early Detection of Fungal Plant Pathogens by Real-Time Quantitative PCR: The Case of *Diplodia sapinea* on Pine

Nicola Luchi, Alberto Santini, Francesca Salvianti, and Pamela Pinzani

Abstract

This chapter reports the use of real-time quantitative PCR to detect *Diplodia sapinea*, a fungal plant pathogen that causes shoot tip dieback and tree mortality on pine trees. This molecular approach represents a reliable and sensitive tool to detect fungal pathogens in DNA extracted from plant tissues and its use can be also recommended to study fungal behavior in host tissues by quantifying fungal growth in the latent phase, when symptoms in the host are not present yet.

Key words Fungal pathogens, Latent infections, qPCR, Diplodia shoot blight, Conifers

1 Introduction

During the last years it has been drawn attention to the impact of the climate change on natural ecosystems. These drastic changes in climate, mainly characterized by increasing temperatures and a change in rainfall regimes, also have an effect on forest ecosystems, bringing changes in different trophic levels [1–3]. In addition to plants, different organisms, such as phytopathogenic fungi, are directly involved in this phenomenon. The increase of temperatures can determine variations in growth and aggressiveness of pathogens and also in their spreadability in new geographical area, in which they cannot thrive so far [4, 5].

In addition, as a consequence of globalization, we are assisting to a growing trade of goods and movement of people. This process is leading to a continuous birth of new pest–host associations and, therefore, to the raise of new epidemic diseases. In order to prevent the spread of new diseases or to manage it, once the disease agents are established, stakeholders are increasingly demanding for reliable, sensitive, specific and fast tools for early pathogen detection. For this aim in the last years plant pathology has borrowed and adapted typical clinical molecular markers-based tools.

Some of the pathogenic microorganisms that cause plant diseases live in the plant tissues in a latent phase, without showing any symptom in the host. In this phase a strict synergy between the host and microorganism is established, allowing the survival of the latter using the nutrients made available by the plant. Moreover, when physiological changes take place in the plant, caused for example by water stress, the microorganism shows its pathogenic action, with subsequent presence of symptoms and damage on the plant [6].

Diplodia sapinea (Fr.) Fuckel (syn. *Diplodia pinea* (Desm.) Kickx, *Sphaeropsis sapinea* (Fr.:Fr.) Dyko and Sutton) is a fungal pathogen that causes shoot blight and canker disease of different conifer species [7–9] (Fig. 1). The pathogen can live as an



Fig. 1 (a) *Diplodia sapinea* infection on pine (*Pinus nigra*) trees; (b) Symptomatic shoot infected with *D. sapinea*; (c) Healthy pine shoots with latent infections of *D. sapinea*; (d) *D. sapinea* mycelium growing on PDA media (potato dextrose agar)

endophytic fungus in healthy plant tissues, causing disease when trees are subjected to biotic or abiotic stress [9, 10] (Fig. 1).

Conventional methods to detect *D. sapinea* in latent phase consist in isolation of the fungus by placing a small piece of healthy plant tissue on culture media and, after incubation, identifying the pathogen by using morphological and molecular methods. However, these approaches are time-consuming, and their accuracy depends on the expertise of the person making the diagnosis.

Over the last few years, new diagnostic methods have been developed to detect fungal organisms colonizing plant and wood tissue. These new methods are mainly based on the detection of fungal DNA in a complex matrix such as plant tissues. Most of these techniques are already well established in clinical research, but they are now becoming accessible from an economic point of view to plant biology as well. During the last years the use of molecular techniques, based on real-time PCR assay allowed the development of a new protocols able to detect a plant pathogen [11–13], even before the presence of symptoms on the host [14–17] and the to detect of fungal pathogens that are difficult to culture [18]. More recently the sensitivity of qPCR has highlighted an interdisciplinary approach between clinical cancer research and plant pathology in the study of latent invaders [19].

The aim of this work is to develop a real-time PCR assay to detect *D. sapinea* latent pathogen in apparently healthy plant tissues. The chapter deals with preparation of samples, DNA extraction, and optimization of a real-time PCR TaqMan assay to detect *D. sapinea* DNA. The chapter concludes with a brief description of the duplex assay optimized for the pathogen (*D. sapinea*) and the plant (*Pinus nigra*) in order to study the infectious disease process within the host.

2 Materials

2.1 Sample Collection

1. Plastic bags.
2. Scissors.
3. Ice.
4. Ethanol (75–100%).
5. NaOCl 2%.
6. Sterile water.
7. Disposable gloves.

2.2 DNA Extraction from Plant and Mycelium

1. Plant DNA extraction kit (e.g., E.Z.N.A. Plant DNA kit).
2. TissueLyser System.
3. Stainless grinding balls.

4. 1.5 ml microcentrifuge tubes.
5. 2 ml microfuge tubes.
6. Isopropanol.
7. Ethanol.
8. Pipets and pipet tips.
9. Heating blocks.
10. Microcentrifuge.
11. Vortexer.
12. Disposable gloves.
13. Spectrophotometer (e.g., NanoDrop).

2.3 Real-Time PCR Assay

1. qPCR instrument (e.g., StepOne Plus, Life Technologies).
2. Plasticware specific for qPCR instrument.
3. Primers and TaqMan probe (as described below).
4. PCR Master Mix.
5. PCR grade water.
6. Pipettes and pipet tips.
7. Disposable gloves.

3 Methods

3.1 Sample Collection

1. Asymptomatic pine (*Pinus nigra*) shoots are collected and examined to detect the presence of *Diplodia sapinea* as latent pathogen in healthy needles and shoot tissues (see Note 1) (Fig. 1). Symptomatic plant tissues, with visible symptoms of *D. sapinea* (see Note 2), are also included as positive control (Fig. 1).
2. Each pine shoot portion is surface-sterilized by using ethanol, NaOCl, and sterile water according to Stanosz et al. [6].
3. From each shoot small fragments of needles and shoot tissue are collected (see Note 3).

3.2 DNA Extraction

1. DNA from plant material is extracted by using the E.Z.N.A. Plant DNA kit protocol (Omega).
2. Fresh plant material (ca. 80 mg) is placed in 2-ml microfuge tubes containing 400 µl lysis buffer P1 and stainless grinding balls.
3. Plant material was then disrupted by using TissueLyser System (Qiagen) (see Note 4).
4. After incubation at 65 °C for 10 min, 140 µl of buffer P2 is added to the lysis mixture, vortex and centrifuge at 10,000 × g for 10 min.

5. Clear lysate was transferred to a new 1.5 ml microcentrifuge tube and 0.7 volumes of isopropanol was added and vortex to precipitate DNA (*see Note 5*).
6. Centrifuge at $14,000 \times g$ for 2 min to pellet DNA and discard supernatant (*see Note 6*).
7. Add 300 μ l sterile water (heated at 65 °C), vortex to resuspend the pellet and incubate 65 °C for 5 min.
8. Add 5 μ l RNase A and vortex to mix.
9. Add 150 μ l buffer P3 and 300 μ l ethanol (100%) and vortex immediately.
10. Transfer the mixture from **step 9** to the HiBind DNA Mini Column, centrifuged at $10,000 \times g$ for 1 min and discard the flow-through to remove all precipitates and cell debris.
11. Add 650 μ l DNA Wash buffer and centrifuge at $10,000 \times g$ for 1 min and discard the flow-through (*see Note 7*).
12. Transfer the HiBind DNA Mini Column to a new 1.5 ml microcentrifuge tube and add 50 μ l Elution buffer (heated at 65 °C) to elute the DNA (*see Note 8*).
13. Centrifuge at $10,000 \times g$ for 1 min to elute the DNA (*see Note 9*).
14. DNA is then ready for UV DNA quantification and real time PCR processing.

3.3 Quantification of *D. sapinea* by qPCR

1. Absolute quantification of *D. sapinea* DNA is carried out amplifying the target region of small subunit of ribosomal RNA gene by using the following primers and TaqMan probe:
 Forward primer: 5'-GTAAAAACTGACGTTGAGGGACG-3';
 Reverse primer: 5'- CATAATTGTCTGCCCGGACTACT-3';
 TaqMan Probe: 5'-FAM-AGGCTCGGGTAGCGAATA
 GGATTAGATAACCC-TAMRA-3'.
2. *Diplodia sapinea* DNA is obtained from an axenic culture (*see Note 10*) and the concentrations measured at 260 nm using a NanoDrop spectrophotometer.
3. The standard curve is generated by using a five fold serial dilution (ranged from 25 ng/ μ l to 0.04 ng/ μ l).
4. For each run include the suitable control samples and NTC (No Template Control).
5. Prepare the real-time PCR reaction in a final volume of 25 μ l by using:
 - (a) 300 nM forward primer;
 - (b) 300 nM reverse primer;
 - (c) 200 nM TaqMan probe;
 - (d) 12.5 μ l TaqMan Universal Master Mix;

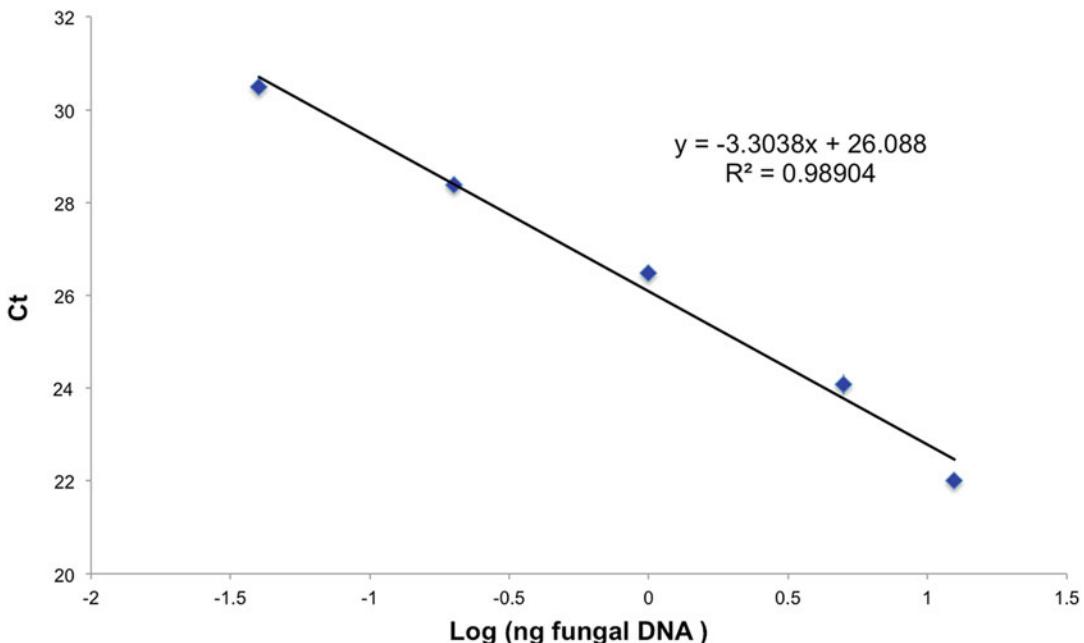


Fig. 2 Standard curve for *D. sapinea* generate by plotting the value of threshold cycle (C_t) in ordinates vs. concentration of *D. sapinea* DNA serial dilution. The value of each dilution is the average \pm S.D. (standard deviation) of three replicates. The unknown samples are calculated by the interpolation of respective C_t value on this standard curve

6. Add 5 μ l template DNA.
7. Run each sample in triplicate according to the following PCR protocol: 50 °C (2 min); 95 °C (10 min); 45 cycles of 95 °C (0.5 min); 60 °C (1 min).
8. The concentration of unknown samples will be calculated by interpolation of the data on the standard curve (Fig. 2).
9. For each kind of sample (mycelium, symptomatic tissue, healthy tissue, and needles) the amount of *D. sapinea* DNA is expressed as pg *D. sapinea* DNA/100 mg pine tissue (Fig. 3).

3.4 Duplex qPCR

1. A duplex assay is developed to detect the variation of pathogen (*D. sapinea*) DNA and host (pine) DNA during the infection process (see Note 11).
2. Absolute quantification of *D. sapinea* DNA is carried out by using primers and probe described in Subheading 3.3.
3. Quantification of *Pinus nigra* DNA is carried out amplifying the region of 28S ribosomal DNA using the following primers and TaqMan probe:
 - Forward primer: 5'-AACCAACACGGGCTGCTG-3';
 - Reverse primer: 5'-GGCAAACCGACAGGCAGTAA-3';
 - TaqMan Probe: 5'-VIC-ATCGTCACAAGCGCGGGCG-TAMRA-3'.

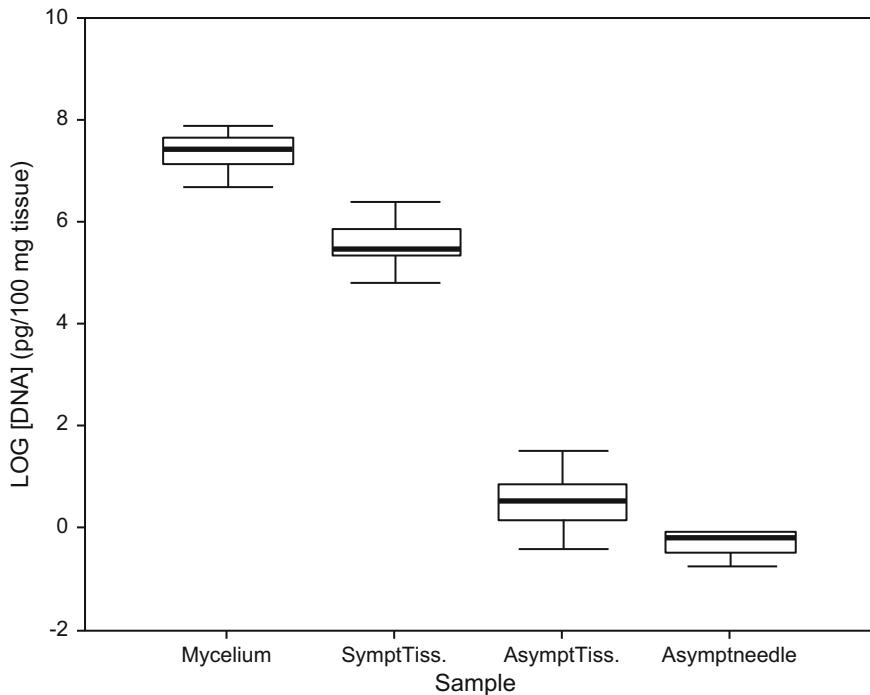
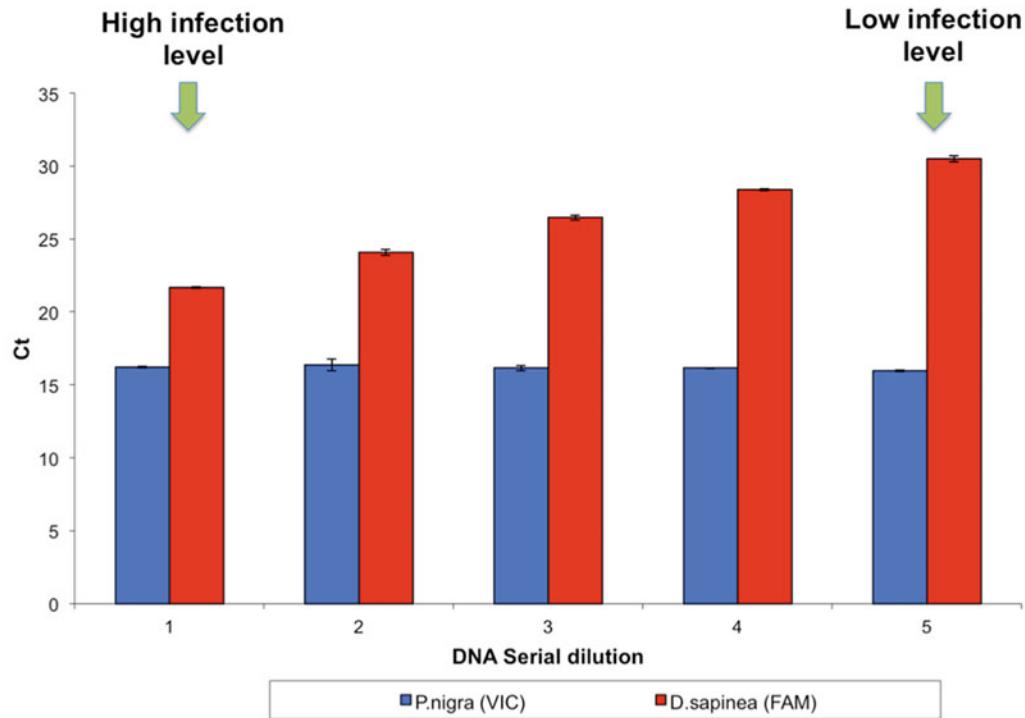
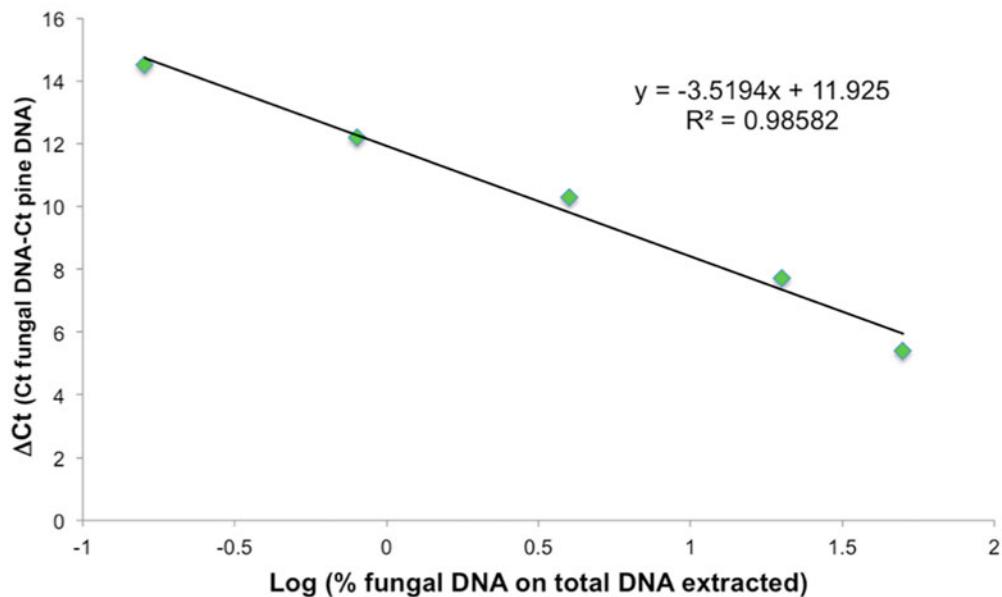


Fig. 3 Quantification of *D. sapinea* DNA from different biological samples: axenic *D. sapinea* cultures (Mycelium); pine symptomatic tissue infected with *D. sapinea* (SymptTiss.); healthy pine tissue with *D. sapinea* latent infections (AsymptTiss.); healthy pine needles (Asymptneedle)

4. The standard curve is generated by using a fivefold serial dilution of *D. sapinea* DNA (ranged from 25 ng to 0.04 ng/ μ l), mixed in 25 ng/ μ l of *P. nigra* DNA.
5. For each run include the suitable control samples and NTC (no-template control).
6. Prepare the real-time PCR reaction in a final volume of 25 μ l by using the following:
 - (a) 300 nM forward primer;
 - (b) 300 nM reverse primer;
 - (c) 200 nM TaqMan probe;
 - (d) 12.5 μ l TaqMan Universal Master Mix;
7. Add 5 μ l template DNA.
8. Run each sample in triplicate according to the following PCR protocol: 50 °C (2 min); 95 °C (10 min); 45 cycles of 95 °C (0.5 min); 60 °C (1 min).
9. For each serial dilution the threshold cycle (C_t) is determined and the ΔC_t value is calculated on the basis of the difference between C_t of plant and fungus (Fig. 4a). The duplex standard curve is reported in a semilogarithmic graph (base 10) by plotting the ΔC_t value versus the dilution factor (expressed as



(a)



(b)

Fig. 4 Duplex qPCR assay by using five fivefold serial dilution of *D. sapinea* DNA (ranged between 25 ng and 0.04 ng/ μ l) mixed in pine DNA with the same concentration (25 ng). (a) The duplex qPCR assay is able to detect the different TaqMan probe (*D. sapinea*-labeled with FAM and *P. nigra*-labeled with VIC) at different infection levels. The value of each dilution is the average \pm S.D. (standard deviation) of three replicates. (b) The duplex standard curve is reported by plotting the ΔC_t value (between C_t fungal DNA and C_t pine DNA) versus the dilution factor (expressed as % fungal DNA on total DNA extracted)

% fungal DNA on total DNA extracted) (Fig. 4b). This plot is used as standard calibrator curve for extrapolating the percentage of fungal DNA in unknown plant samples.

4 Notes

1. Pine shoot samples (ca. 20 cm long) are collected from *Pinus nigra* tree, placed in plastic bags and stored with ice during the transport to the laboratory for subsequent analysis. The distal portion of each shoot is used for DNA analysis.
2. The presence of *D. sapinea* from symptomatic collected plant tissue is confirmed by using traditional approaches as described by Luchi et al. [7].
3. Cut small pieces of distal shoot portion and place (ca. 80 mg) in 2-ml microfuge tubes for subsequent DNA extraction.
4. Plant samples are disrupted and homogenized (1 min at 30 Hz) with the crushing action of the stainless grinding balls. Disassemble the Adapter set (TissueLyser), remove the microtubes and reverse the position of the microtubes within the Adapter Set. Grind the samples (second step) for 1 min at 30 Hz. Remove the microtubes form the Adapter Set and centrifuge for 1 min at 20,000 × *g* add 200 µl lysis buffer P1 and vortex.
5. Do not disturb the insoluble pellet. In this step detergents, proteins, and polysaccharides are removed.
6. In order to allow any residual liquid to drain, invert the microtube on paper towel for 1 min.
7. Repeat this step for two times, and then centrifuge the empty HiBind DNA Mini Column at 10,000 × *g* for 2 min to dry the membrane.
8. Incubate at room temperature for 5 min.
9. Repeat this step for two times.
10. *Diplodia sapinea* strain is grown on 300PT cellophane discs on potato dextrose agar (PDA) in 90-mm Petri dishes and maintained in the dark at 20 °C. After 7 days, mycelium is scraped from the cellophane surface and stored in 2 ml microtubes at –20 °C. Fungal mycelium (ca. 100 mg) is ground by using stainless grinding balls in TissueLyser (see Note 4). DNA extraction was carried out by using E.Z.N.A. Plant DNA kit protocol (see Subheading 3.2).
11. The qPCR assay is validated in a mixture containing pathogen DNA (*D. sapinea*) and plant DNA (*P. nigra*). The sample mimics an extraction from healthy tissue where the amount of plant DNA is higher than pathogen DNA. For these reasons the serial dilution of pathogen (ranged between 25 ng and 0.04 ng/µl) is mixed in host DNA with the same concentration (25 ng).

References

1. Desprez-Loustau ML, Marçais B, Nagleisen LM et al (2006) Interactive effects of drought and pathogens in forest trees. *Ann For Sci* 63:597–612
2. Garrett KA, Forbes GA, Savary S et al (2011) Complexity in climate-change impacts: an analytical framework for effects mediated by plant disease. *Plant Pathol* 60:15–30
3. Ramsfield TD, Bentz BJ, Faccoli M et al (2016) Forest health in a changing world: effects of globalization and climate change on forest insect and pathogen impacts. *Forestry* 89:245–252
4. Shaw MW, Osborne TM (2011) Geographic distribution of plant pathogens in response to climate change. *Plant Pathol* 60:31–43
5. Brodde L, Adamson K, Camarero JJ et al (2018) Diplodia tip blight on its way to the north: drivers of disease emergence in northern Europe. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2018.01818>
6. Stanosz GR, Blodgett JT, Smith DR et al (2001) Water stress and *Sphaeropsis sapinea* as a latent pathogen of red pine seedlings. *New Phytol* 149:531–538
7. Luchi N, Oliveira Longa CM, Danti R et al (2014) *Diplodia sapinea*: the main fungal species involved in the colonization of pine shoots in Italy. *For Pathol* 44:372–381
8. Blodgett JT, Bonello P (2003) The aggressiveness of *Sphaeropsis sapinea* on Austrian pine varies with isolate group and site of infection. *For Pathology* 33:15–19
9. Slippers B, Wingfield MJ (2007) Botryosphaeriaceae as endophytes and latent pathogens of woody plants: diversity, ecology and impact. *Fun Biol Rev* 21:90–106
10. Flowers J, Hartman J, Vaillancourt L (2006) Histology of *Diplodia pinea* in diseased and latently infected *Pinus nigra* shoots. *For Pathol* 36:447–459
11. Luchi N, Pepori AL, Bartolini P et al (2018) Duplex real-time PCR assay for the simultaneous detection of *Caliciopsis pinea* and *Fusarium circinatum* in pine samples. *Appl Microbiol Biotechnol* 102:7135–7146
12. Luchi N, Capretti P, Surico G et al (2005) A real-time quantitative PCR assay for the detection of *Sphaeropsis sapinea* from inoculated *Pinus nigra* shoots. *J Phytopathol* 153:37–42
13. Luchi N, Ghelardini L, Belbahri L et al (2013) Rapid detection of *Ceratocystis platani* inoculum by quantitative real-time PCR. *Appl Env Microbiol* 79:5394–5404
14. Maresi G, Luchi N, Pinzani P et al (2007) Detection of *Diplodia pinea* in asymptomatic pine shoots and its relation to the normalized insolation index. *Forest Pathol* 37:272–280
15. Luchi N, Capretti P, Pinzani P et al (2005) Real-time PCR detection of *Biscogniauxia mediterranea* in symptomless oak tissue. *Lett Appl Microbiol* 41:61–68
16. Luchi N, Capretti P, Vetraino AM et al (2006) Early detection of *Biscogniauxia nummularia* in symptomless European beech (*Fagus sylvatica* L.) by TaqMan™ real-time PCR. *Lett Appl Microbiol* 43:33–38
17. Migliorini D, Ghelardini L, Tondini E et al (2015) The potential of symptomless potted plants for carrying invasive soil-borne plant pathogens. *Div Distrib* 21:1218–1229
18. Migliorini D, Ghelardini L, Luchi N et al (2019) Temporal patterns of airborne *Phytophthora* spp. in a woody plant nursery area detected using real time PCR. *Aerobiologia*. <https://doi.org/10.1007/s10453-018-09551-1>
19. Luchi N, Capretti P, Pazzagli M et al (2016) Powerful qPCR assays for the early detection of latent invaders: interdisciplinary approaches in clinical cancer research and plant pathology. *Appl Microbiol Biotechnol* 100:5189–5204. <https://doi.org/10.1007/s00253-016-7541-5>



Chapter 9

A General Protocol for Accurate Gene Expression Analysis in Plants

Ellen De Keyser, Laurence Desmet, Magali Losschaert, and Jan De Riek

Abstract

Gene expression analysis by means of RT-qPCR is a highly sensitive technique. However, this requires an accurate protocol for the whole procedure from sampling to data analysis. We have optimized this protocol specifically for the analysis of plant tissues. Special attention is paid to RNA quality and integrity and to the appropriate setup of the assays in order to be compliant with the MIQE guidelines. This protocol was already successfully applied in ten different plant species.

Key words RNA integrity, Plant, RT-qPCR, Extraction protocol, Degradation, SPUD, noRT

1 Introduction

The RT-qPCR technique was optimized intensively the past two decades. RT-qPCR is a highly sensitive method of quantification of gene expression, but this requires accuracy at all stages of the analysis. For application in plant gene expression studies, optimization was implemented at all crucial steps from RNA isolation [1] and quality control [2, 3] up to the final quantification methods [4, 5]. MIQE-guidelines (*Minimum Information for Publication of Quantitative Real-time PCR Experiments*; [6]) were set in order to stimulate the scientific community to quantify in an accurate manner and also to provide all essential data when publishing gene expression studies. However, in plant science, still too many papers on gene expression are published with inaccurate quantification [3, 4].

In this chapter, we describe a complete protocol for gene expression analysis using RT-qPCR, starting from sampling the plant material. Not only one has to be careful to avoid degradation during sampling, but it is also crucial to include sufficient biological variation. Grinding plant material with a TissueLyser is preferred, but only possible when sampling volumes are limited and the

harvested material is not too hard. For RNA extraction, a CTAB-based protocol [1] is described. Commercially available RNA extraction kits are often not suited for many plant tissues, since viscous plant compounds tend to clog the filter during the extraction. Moreover, the CTAB-based protocol is less costly. RNA quality and integrity is most crucial for reliable quantification [2, 6]. RNA integrity is determined using a microfluidic capillary electrophoresis system and pitfalls for plant tissues are highlighted [3]. Also the use of the SPUD assay for PCR inhibition [7] is incorporated in the protocol. We also focus on the need to include sufficient noRT (no-reverse transcriptase) controls in the analysis. Multiple, assay-validated reference genes [5] are to be included in a consistent RT-qPCR assay [8, 9]. Data analysis is done using qbase + software (Biogazelle) and gene-specific PCR efficiencies are calculated using the LinReg software [10]. This protocol or parts of it were already successfully implemented on azalea [1, 3], camellia [11], rose [12], strawberry [13], chrysanthemum [14], *Cichorium intybus*, *Raphanus sativus*, *Osteospermum*, *Sinningia speciosa*, and *Viola* (unpublished).

2 Materials

2.1 General

For efficient work in the lab we recommend following standard lab equipment:

1. Centrifuge with cooling option.
2. Calibrated set of micropipettes.
3. PCR consumables such as tubes, strips, white qPCR plates (*see Note 1*).
4. Vortex.
5. PCR thermal cycler and real-time PCR thermal cycler.
6. Ultralow-temperature (ULT) freezer.
7. Liquid nitrogen.

2.2 Sampling

1. 2 ml tubes (*see Note 2*) or small paper bags to collect samples.

2.3 Grinding

2.3.1 Using the Retsch TissueLyser

1. Retsch TissueLyser (Qiagen) with 2 blocks for 24 tubes.

2. 3 mm zirconium beads (three per sample) and tweezers (*see Note 3*).

2.3.2 Using a Mortar and Pestle

1. A single mortar, pestle and spoon per sample (*see Note 3*).

2.4 RNA Extraction

Store all stock solutions at room temperature, unless indicated else.

1. 5 M NaCl: add 73.0 g NaCl to 250 ml of water.
2. Prepare the CTAB Extraction Buffer (50 ml, *see Note 4*) by adding 5 ml 1 M Tris-HCl, pH 8.0 (Ambion), 14 ml 5 M NaCl, 2 ml 0.5 M EDTA (Ambion), 1.5 g of CTAB (hexadecyltrimethylammonium bromide) and 2.5 g PVP40 (polyvinylpyrrolidone, MW 40,000) to 29 ml of nuclease-free water. Leave overnight at room temperature to dissolve.
3. Just before extraction, add 1% (v/v) 2-mercaptoethanol (e.g., 0.5 ml in 50 ml) to the extraction buffer (*see Note 4*). Preheat the buffer at 65 °C.
4. Prepare 75% ethanol by adding 75 ml of ethanol 100% (p.a.) to 25 ml of nuclease-free water. Store at 4 °C.
5. Lithiumchloride Precipitation Solution (7.5 M, Ambion). Dilute completely to a 4 M solution by adding 87.5 ml nuclease-free water to the 100 ml solution.
6. ThermoMixer® (Eppendorf).
7. Test tube rotator.
8. Filter cabinet or fume hood (*see Note 5*).
9. NanoDrop (ND-1000) spectrophotometer (Thermo Scientific).

2.5 DNase Treatment

1. DNA-free™ kit (Ambion).
2. ThermoMixer® (Eppendorf).

2.6 RNA Integrity and Quality Control

1. NanoDrop (ND-1000) spectrophotometer (Thermo Scientific).
2. Experion capillary electrophoresis system (Bio-Rad) and Experion RNA Standard Sense analysis chips (Bio-Rad).
For SPUD analysis (optional):
3. Primers, dual-labeled probe and amplicon (Table 1). Dilute SPUD primers to a concentration of 10 µM and the SPUD probe to 5 µM in nuclease-free water upon arrival.

Table 1
Sequence information of the oligos needed to perform the SPUD assay

SPUD	Sequence (5'-3')
Amplicon (101 bp)	AACTTGGCTTTAATGGACCTCCAATTGAGTGTCACAAGCTA TGGAACACCACGTAAGACATAAACCGGCCACATATGGTGCCATG TAAGGATGAATGT
Primer F	AACTTGGCTTTAATGGACCTCCA
Primer R	ACATTCACTCCTTACATGGCACCA
Probe	FAM-TGCACAAGCTATGGAACACCACGT-Tamra

4. Yeast tRNA (50 ng/ μ l): add 100 μ l yeast tRNA (25 mg/ml, ThermoFisher Scientific) to 50 ml of nuclease-free water.
5. Make a 100-fold serial dilution of the SPUD amplicon by adding 1 μ l from the 5 μ M stock solution to 99 μ l of yeast tRNA (50 ng/ μ l). Again dilute 100-fold by adding 1 μ l of the dilution to 99 μ l of yeast tRNA (50 ng/ μ l) and repeat this another two times until you have diluted the SPUD amplicon 10⁸-fold.
6. Heparin sodium salt from porcine intestinal mucosa (2 U/ μ l, filter sterilized 0.22 μ m, Sigma-Aldrich). Dilute to a stock concentration of 40 U/ml by adding 20 μ l of Heparin to 980 μ l of nuclease-free water. Prior to the SPUD analysis, dilute 10- and 100-fold to 4 U/ml and 0.4 U/ml, respectively.
7. UV cabinets (*see Note 6*).

2.7 cDNA Synthesis

1. iScriptTM cDNA synthesis kit (Bio-Rad).

2.8 qPCR

1. UV cabinets (*see Note 6*).
2. Primers for both target and reference genes are diluted to 10 ng/ μ l.
3. For each gene a PCR mix is made (8 μ l/sample; *see Note 7*) by adding 0.3 μ l of both forward and reverse primer, 5 μ l of SensiFAST SYBR[®] No-ROX Mix (Bioline) and 2.4 μ l of nuclease-free water.

3 Methods

3.1 General Guidelines on Handling RNA

Working with RNA is more demanding than working with DNA, because of the chemical instability of the RNA and the presence of RNases. Since RNases are difficult to inactivate, precautions should be taken to avoid RNase contamination (*see Note 3*).

3.2 Sampling

For gene expression analysis, at least three biological replicates are needed. In case of clonally propagated plants, this can be three individual plants under the same treatment. However, in case of seed propagated plants, genotype differences are present within a variety. In that case, each biological replicate can be a pooled sample of different genotypes. Also in case of very small tissues to be sampled, a pooled sample might be required. In all cases, try to incorporate as much as possible the biological variation that might be present in the dataset.

1. Snap freeze the sample in liquid nitrogen as soon as possible (*see Note 8*). In specific cases, some handling (e.g., cutting out a specific organ or tissue) might be needed. In that case, try to avoid RNA degradation as much as possible (*see Note 3*) and

work as fast as possible. It is impossible to predict how fast RNA is degrading, but in some tissues it might occur very fast.

2. Whenever possible, sample immediately in 2 ml tubes and limit the amount of material to 100 mg. In case of pooling, it will depend on the sampling size whether this is still suitable for grinding in a tube (*see Note 2*).
3. Store samples in a ULT freezer.

3.3 Grinding

3.3.1 Using the Retsch TissueLyser

1. Take the samples from the ULT freezer and put in the pre-cooled blocks (≥ 2 h). Open the tubes carefully (*see Note 9*) and add three zirconium beads using the tweezers. Put the blocks in a box with a 3 cm layer of liquid nitrogen.
2. Grind in the TissueLyser for two times 90 s at 30 Hz (*see Note 10*), rotate and cool the blocks in between. Immerse the blocks in liquid nitrogen again.
3. Centrifuge the samples for 20 s at 4 °C at full speed in small batches to prevent heating of the samples. Open the tubes carefully (*see Note 9*) and put the samples back in the blocks. Transfer the samples to a rack and proceed immediately with the extraction protocol (*see Note 11*).

3.3.2 Using a Mortar and Pestle

1. Put the mortar on a large plastic foil in a box of ice and cool it with liquid nitrogen.
2. Add the sample, which was kept on liquid nitrogen, to the mortar and grind until you get a fine powder (*see Note 10*). Add liquid nitrogen if needed to maintain the sample frozen at all times.
3. Weigh max 100 mg of powder in a 2 ml precooled tube. Store the samples in a rack on liquid nitrogen until all samples are ready for RNA extraction (*see Note 11*).

3.4 Total RNA Extraction

1. Add 700 µl prewarmed (65 °C) CTAB Extraction Buffer to the ground sample. Vortex vigorously for 1 min and then incubate at 65 °C and 1400 rpm for 10 min in the ThermoMixer® to lyse cells completely.
2. Centrifuge at $15,800 \times g$ for 5 min at room temperature and transfer supernatant into a 2 ml tube. Add an equal volume of chloroform p.a. (± 650 µl), vortex vigorously for 1 min, put 5 min in the test tube rotator and centrifuge at $15,800 \times g$ for 10 min at room temperature. The upper phase is transferred to a 1.5 ml tube and centrifuged for 10 min at $15,800 \times g$ at room temperature.
3. Supernatant is transferred to a 1.5 ml tube and followed by the addition of an equal volume (± 450 µl) of 4 M LiCl. Invert the tubes a few times and leave at 4 °C overnight.

4. Centrifuge at $15,800 \times g$ for 20 min at 4 °C and remove supernatant. Wash the pellet gently with 800 µl 75% ethanol (4 °C); try to loosen the pellet from the tube by ticking the tube; centrifuge at $15,800 \times g$ for 15 min at 4 °C and remove supernatant. Repeat this washing step. Perform a short spin to precipitate remaining ethanol droplets and remove by pipetting.
5. Air-dry the pellet for at least 20 min by leaving the tubes open at room temperature. Dissolve the pellet in 50 µl of nuclease-free water.
6. Determine the RNA concentration of each sample using the NanoDrop spectrophotometer (*see Note 12*). Use nuclease-free water as a blank and select for RNA measurement (RNA-40). In case the concentration is below 200 ng/µl, go to the next step immediately. In case of higher concentrations (*see Note 13*).
7. Store the remaining RNA in the ULT freezer prior to further use.

3.5 DNase Treatment

1. Add 5 µl 10× DNase I Buffer (0.1 volume) and 1 µl *rDNase I* (= 2 U) to 50 µl of RNA. Mix and incubate for 30 min at 37 °C in a ThermoMixer® (*see Note 13*).
2. Add 5 µl DNase Inactivation Reagent (0.1 volume) and vortex. Incubate for 2 min at room temperature, while vortexing occasionally. Centrifuge for 1.5 min at $10,000 \times g$ at room temperature and transfer the supernatant carefully to a 1.5 ml tube (*see Note 13*).

3.6 RNA Integrity and Quality Control

1. The RNA yield and purity are measured using the Nanodrop spectrophotometer (*see Subheading 3.4, step 6*). The A_{260/280} ratio is ideally 2, A_{260/230} should be between 2 and 2.5. However, in plant tissue we often see deviating values due to the presence of proteins (A_{260/280}) and polysaccharides and/or polyphenols (A_{260/230}) after extraction. In that case, a SPUD assay can be done to test for PCR inhibition.
2. (Optional) To evaluate for the presence of PCR inhibitors, a TaqMan Probe based SPUD assay can be applied on RNA samples with a deviating absorbance ratio. Prepare the SPUD PCR mix (9 µl/sample; *see Note 7*) by adding 0.5 µl of the diluted SPUD amplicon, 0.48 µl of both forward and reverse SPUD primers (10 µM), 0.2 µl of the SPUD probe (5 µM), 2.34 µl of nuclease-free water, and 5 µl of LightCycler480 Probes Master Mix (Roche). 1 µl of RNA (75 ng) is added to 9 µl of the SPUD PCR Mix in a qPCR plate; each sample is analyzed in duplicate. Also incorporate ten SPUD control samples in which no RNA is added to the SPUD PCR Mix.

As a positive control, 1 μ l of heparin should be added in three different concentrations (0.04, 0.4, and 4 U/ml) to 9 μ l of SPUD PCR Mix (each concentration in duplicate). Seal the plate and perform qPCR analysis in a real-time PCR thermal cycler. In case inhibitors are present in a RNA sample, C_q (Quantification cycle) values of the SPUD amplicon in these samples will be significantly higher compared to the average C_q of the SPUD control samples (typically $>1 C_q$). The C_q value of the heparin positive control samples should decrease when more diluted, with clear inhibition in the highest concentration. RNA samples that appear to be inhibited should be reanalyzed in the SPUD using a tenfold dilution series of the sample. In case diluting can solve the inhibition, one might decide to use a lower concentration of RNA during cDNA synthesis.

3. RNA integrity is determined on a representative random subset of the samples (typically 10%) using a microfluidic capillary electrophoresis system (e.g., Experion, Bio-Rad). Since the Experion software was not developed with plant tissue, one cannot rely on the RNA Quality Indicator (RQI) score (*see Note 14*). RNA integrity is evaluated visually based on the ribosomal RNA ratio (25S rRNA peak higher compared to 18S rRNA peak) in combination with clear ribosomal bands in the virtual gel view. In case chloroplasts were present in the tissue, also small 16S and a 23S rRNA peaks might be present. Degradation of the RNA can also be noticed as a shift in the electropherogram toward the so-called fast region (signal before the ribosomal peaks) and a smear in the virtual gel view (Fig. 1).

3.7 cDNA Synthesis

1. To minimize sample variation, start for all samples from the same amount of RNA (100 fg to 1 μ g total RNA) in a total volume of 10 μ l in PCR strips (*see Note 15*). Prepare samples in double since we intend to use one set for preparation of noRT (no reverse transcriptase) samples (*see Note 16*). Leave samples on ice at all times.
2. Prepare a mix of 4 μ l of 5 \times iScript reaction mix, 5 μ l of nuclease-free water and 1 μ l iScript reverse transcriptase per sample (*see Note 7*). Add 10 μ l of this mix to each RNA sample for cDNA synthesis. Add 10 μ l of nuclease-free water to the RNA samples for noRT.
3. Place the strips in a PCR thermal cycler and run the following program: 5 min 25 °C, 20 min 46 °C, 1 min 95 °C.
4. cDNA is diluted with nuclease-free water depending on the amount of genes (reference + target), so all genes can be tested with the same batch of cDNA (*see Note 17*).
5. Store the strips on ice or at –20 °C until further use.

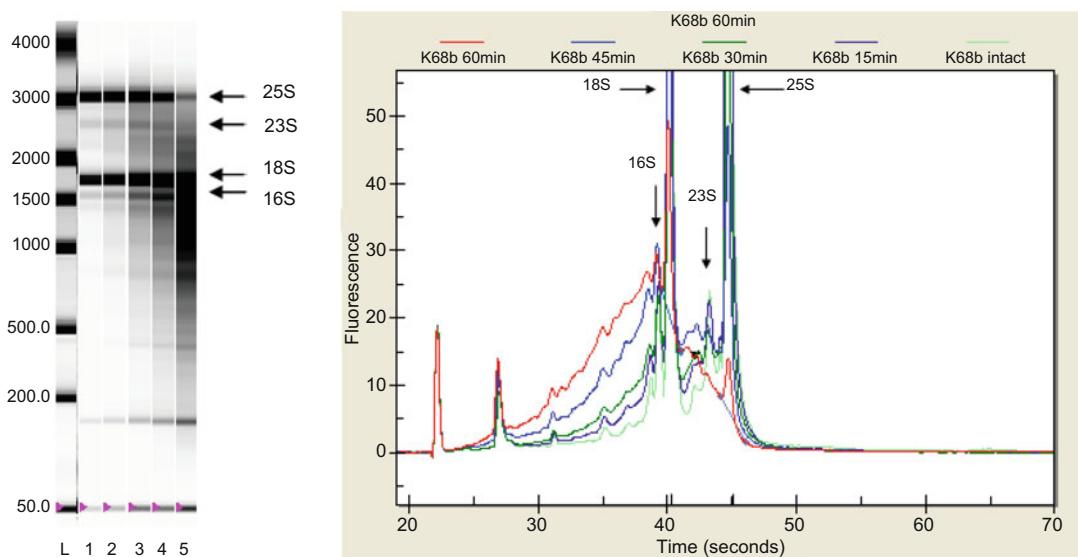


Fig. 1 RNA quality control with the Experion (Bio-Rad). Electropherogram (right) and virtual gel-view (left) of an RNA degradation series that was constructed by heating an RNA sample for 0, 15, 30, 45, and 60 min at 80 °C. The loading marker and small RNA band and cytoplasmic 18S and 25S as well as 16S and 23S chloroplast and mitochondrial ribosomal bands are indicated with arrows. Lanes: (L) size standard, (1) intact RNA, (2) 15 min, (3) 30 min, (4) 45 min, (5) 60 min. Intensity settings can vary between lanes. Reproduced from BioMedCentral [3]

3.8 RT-qPCR

1. Primer development is done using Primer3plus software. In general, 20 bp primers with a Tm of 60 °C and a GC-content between 30% and 80% are developed on the cDNA sequence of the genes of interest (target and reference genes). Ideally, primers are on different exons to avoid amplification of potentially contaminating genomic DNA. Amplicon length should be between 80 bp and 150 bp. Always test primer specificity on a few cDNA samples using the protocol described in **step 3** before using them on your entire dataset (*see Note 18*).
2. For normalization of gene expression, a validated set of unrelated reference genes is needed. Therefore at least ten reference genes (*see Note 19*) should be tested in a GeNorm assay. Analyze these ten genes on at least ten representative samples (*see Note 17*) from your dataset using the RT-qPCR protocol described below, incorporating all treatment, time and tissue variation present; it is better to use over ten samples if a large variation is present (*see Note 20*). Data analysis is done using the qbase+ software (Biogazelle). The software indicates how many and which reference genes are to be incorporated in the normalization factor. These reference genes must then be analyzed on the complete dataset as described in **step 3**.

3. All target and the GeNorm selected reference genes are run on the complete dataset according to the sample maximization method. This means all samples are to be analyzed on the same plate for one gene; a 384-well thermal cycler is most convenient for this purpose (*see Note 21*). It is recommended to include two or three technical replicates to evaluate the C_q variation. noRT samples are analyzed for all genes and at least three NTCs are added to the plate. For a NTC, only the PCR mix is loaded on the plate. For each gene a PCR mix (*see Note 7*) is made and 8 μ l/sample is added to a 384-well PCR plate. Next, 2 μ l cDNA or noRT sample is added to each well. Plates are sealed with an adhesive film, centrifuged (2 min, 1500 $\times g$) to remove droplets and placed in the real-time PCR machine (LightCycler480, Roche). Cycling conditions are 2 min at 95 °C, followed by 45 cycles of 5 s 95 °C, 10 s 60 °C and 20 s 72 °C, with data acquisition at the end of every cycle. Melting curve analysis is performed as follows: 5 s 95 °C, 1 min 65 °C and heating to 97 °C with a ramp rate of 0.06 °C/s. Data acquisition occurs ten times for every °C.

3.9 Data Analysis

1. Data are analyzed first using the software of the real-time PCR machine (*see Note 18*). When possible, the second derivative maximization method [15] is preferred for C_q determination in every run; this is incorporated in the LC480 software. Check for aspecific amplification in both amplification and melting curves. In case the C_q -difference between a sample and the corresponding noRT sample is 5 or less, the sample is discarded from further analysis for this gene. In case this occurs in a reference gene, the sample has to be removed completely since it is impossible to calculate a normalization factor for it.
2. We recommend to use qbase+ software (Biogazelle) for all further data-analysis (*see Note 22*). Preferably, a gene and run-specific amplification efficiency is calculated based on the amplification curves of all samples using LinReg PCR software (*see Note 23*) and these values are added in qbase+. Always check normalization factor stability and replicates. Statistics can be done in qbase+ as well, but we prefer to export the results table to MS Excel and use common statistical packages.
3. Both log-transformed and untransformed gene expression levels can be exported from qbase+. The untransformed values are more suitable for graphical representation, for statistical analysis, log-transformed data are required. For averaging the biological replicates, always use the geometrical mean, since this is less sensitive to outliers.

4 Notes

1. We recommend the use of Eppendorf Safe-Lock Tubes (Eppendorf) as they have optimized sealing properties and are certified PCR clean. White qPCR plates are preferred over transparent plates because data collection occurs more accurately.
2. When sampling immediately in tubes, make sure to use tubes with a round bottom and a SafeLock cap so that they will not easily “pop” open during handling. Do not overfill the tubes, else it is impossible to add the beads for grinding. When sampling takes place over a long time period (hours) and the caps cannot be closed, accumulation of liquid nitrogen in the tube can occur. In those cases, the use of punched caps can be useful. When a sample is taken, close the tube with a separate cap in which a hole is punched with a needle; this will allow the nitrogen to evaporate. When storing the samples in the –80 °C freezer, leave on these punched caps for a few hours. Then you can remove these caps and close the tubes with their proper caps more safely. This should prevent explosion of the tubes in a later stage (*see also Note 9*).
3. Some precautions should be followed while working with RNA. Designate a special area for RNA work only. Always wear a lab coat and gloves when working with RNA to prevent contamination from RNases found on the human body. Change gloves frequently and keep tubes closed whenever possible. Treat surfaces of benches and equipment with commercially available RNase inactivation agents. Use disposable, plastic ware that have been tested and certified RNase-free. Glassware, spatulas, tweezers, etc. should be baked at +180 °C for at least 4 h. All material should be wrapped in aluminum foil to prevent contamination after baking. Purchase reagents that are free of RNases. Reserve separate reagents for RNA work only. All solutions should be made with water that is subject to rigorous quality control testing and is guaranteed to be RNase-free. Alternatively, DEPC-treated water can be used. The treatment typically involves incubating the solution at room temperature with 0.1% DEPC (diethyl pyrocarbonate) for a few hours, usually overnight, followed by autoclaving the solution to eliminate residual DEPC. Be aware that certain reagents such as Tris cannot be DEPC treated.
4. The CTAB Extraction Buffer can be stored for at least 2 weeks at room temperature, so a higher volume can be prepared at once in case of a larger amount of samples to be extracted. However, 2-mercaptoethanol can only be added just before

extraction to the volume of buffer needed at that time. Once 2-mercaptoethanol is added, the CTAB buffer must be used immediately.

5. As during RNA extraction hazardous reagents are used, working in a filter cabinet or fume hood is obliged.
6. Preferably the qPCR runs are set up in a UV cabinet and two separate cabinets are used. In the first one, primers are diluted and PCR mixes are prepared and pipetted into the plate, in the second one the cDNA is added. Always decontaminate during 30 min using UV light before working in the cabinet.
7. The volumes needed per sample are indicated. Volumes must be multiplied by the total amount of samples. Always count in a few extra samples to prepare a mix, since you always loose some of the mix when pipetting.
8. Besides avoiding RNA degradation, flash-freezing the samples is also needed to stop all biological processes in the samples tissue. When this takes too much time, it is possible that certain biological processes are initiated and genes are up- or down-regulated only due to the sampling. This must be avoided in order to obtain reliable transcription profiles.
9. Tubes that were in contact with liquid nitrogen during sampling, grinding, and extraction might be prone to exploding. Always wear safety glasses when handling these samples.
10. Complete disruption of cell walls and plasma membranes is required to release all the RNA contained in the sample. Depending on the tissue, more time can be needed to grind until total disruption in the TissueLyser. Some samples are impossible to grind completely to a powder in the TissueLyser. It can be of help to cut tough samples (e.g., stem tissue) in smaller pieces prior to freezing. A grinding mill (IKA) is not recommended, due to higher risk of cross-contamination and degradation due to insufficient cooling.
11. After grinding (with both methods), samples can also be stored in the ULT freezer prior to proceeding with the RNA extraction step.
12. Pipet 1.5 µl of RNA to a fresh tube or PCR strip for quantification and store the remaining RNA in the ULT freezer prior to further use. In this way, the RNA cannot get contaminated or degraded during measurement.
13. In case the RNA concentration is above 200 ng/µl, a heavy DNase treatment is recommended. Dilute RNA with nuclease-free water to 500 ng/µl if higher. Add 10 µl of 10× DNase I buffer (0.1 volumes) and 1 µl *rDNase I* to 100 µl of (diluted) RNA. Mix and incubate for 30 min at 37 °C, add another 1 µl *rDNase I* followed again by 30 min at 37 °C. Add 20 µl of

DNase Inactivation Reagent (0.2 volumes) and proceed with the protocol as described. DNase Inactivation Reagent could inhibit PCR, so be careful to avoid transferring it to the tube of RNA.

14. The software of the Experion was not developed for plant-specific applications. However, when using the Agilent 2100 Bioanalyzer system, the 25S plant peak is already indicated (no longer 28S). Nevertheless, 16S and 23S ribosomal peaks are not yet described in the software, so one must still be careful when working with chloroplast-containing tissues.
15. Starting from a higher amount of RNA for cDNA synthesis gives better results. However, it is crucial to start from identical concentrations in all samples. If certain samples have clearly lower concentration compared to the rest of the dataset, it is advised to try to repeat the RNA extraction for these samples whenever possible in order to obtain higher concentrations. Sometimes low-concentrated samples can be due to a specific treatment or tissue type and in that case, one might decide to discard these samples from the dataset in order not to compromise the analysis of the majority of the samples.
16. Sometimes only a random subset of 10–20% of the samples is used as noRT samples. Also, some people only test noRT samples with a limited number of genes. However, we have experienced that noRT problems are gene- and sample-dependent, so we strongly advise to use noRT samples for all genes and all samples.
17. A GeNorm analysis requires quite some cDNA. Hence, it is advised to prepare a specific set of cDNA from the RNA samples selected for the GeNorm assay and to use this cDNA solely for the GeNorm analysis. One has to avoid at all times that new cDNA synthesis must be done during the experiment (e.g., when adding some genes to the analysis), since gene expression cannot be compared reliably between different cDNA synthesis of the same RNA; in that case, at least reference genes have to be run again on these samples.
18. When using SYBR Green for quantification, you have the option to perform melting curve analysis. Melting curve analysis can provide valuable information on amplification specificity of the primers. In case of aspecific amplification, a melting peak with a deviating melting temperature can be seen. A peak with a low melting temperature (often around 75 °C) is an indication for the presence of primer dimers. Often these primers dimers are not present in the cDNA samples but are only seen in some noRT and NTC samples; in that case there is no real problem. But in case a primer dimer melting peak is seen in noRT, NTC, and cDNA samples, it might be required to lower the primer

concentration used in the Master Mix or one can just develop new primers. Probe-based quantification instead of SYBR Green is another option to avoid primer dimers and increase specificity of amplification, but labeled probes are far more costly and have to be ordered for each gene. In case of clear amplification in a large part of noRT samples, which indicates the presence of severe genomic DNA contamination, one could decide to perform a heavy DNase treatment (*see Note 13*) on the whole set of RNA samples and repeat the protocol again from this step on.

19. In case no sufficient reference genes are found in literature, we recommend to search for some of the candidate reference genes proposed by [5] using a Blast search against sequence data available in your own species.
20. Selection of the optimal set of reference genes can vary greatly depending on the treatments or sample types under study. There is no such thing as an optimal set of reference genes for a specific plant species. It is crucial to repeat the validation step each time you perform a new experiment with other treatments or tissues. However, if only few changes are made, one can opt to run all previously selected reference genes on the new dataset. When the M and CV-values of the normalization factor are not yet acceptable, it is sometimes sufficient to run the next best reference gene as well and add it to the normalization factor.
21. When using a large dataset, it might be impossible to analyze all samples together, especially when using a 96-well plate. In that case, it is recommended to use interrun calibration for every gene. For that purpose, load at least three samples on all plates that are analyzed with a specific gene; the amplification of these samples will be used to calculate potential interrun variation in qbase+ (Biogazelle). For each gene, different interrun calibrators can be used in case of limited amounts of cDNA. Samples and their corresponding noRT samples are not necessarily run together on the same plate; interrun calibrators are not needed in that case.
22. A short-term license can be bought at www.qbaseplus.be. Else, all calculations should be done according to the method described in [9].
23. Never assume a PCR efficiency of 2, since this is hardly the case. PCR efficiency can also be determined based on standard curves [3]. For that purpose, a dilution series of cDNA (ideally) or plasmid DNA (in case of limited amount of RNA available) can be analyzed for every gene. However, this is time-consuming and plasmid DNA might even overestimate the PCR efficiency slightly.

References

1. Luypaert G, Witters J, Van Huylenbroeck J et al (2017) Induced expression of selected plant defence related genes in pot azalea, *Rhododendron simsii* hybrid. *Euphytica* 213:227
2. Die JV, Roman B (2012) RNA quality assessment: a view from plant qPCR studies. *J Exp Bot* 63:6069–6077
3. De Keyser E, Desmet L, Van Bockstaele E et al (2013) How to perform RT-qPCR accurately in plant species? A case study on flower colour gene expression analysis in an azalea (*Rhododendron simsii* hybrids) mapping population. *BMC Mol Biol* 14:13. <https://doi.org/10.1186/1471-2199-14-13>
4. Guttierrez L, Mauriat M, Guénin S, Pelloux J et al (2008) The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription – polymerase chain reaction (RT-PCR) analysis in plants. *Plant Biotechnol J* 6:609–618
5. Czechowski T, Stitt M, Altmann T et al (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol* 139:5–17
6. Bustin SA, Benes V, Garson JA et al (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55:611–622
7. Nolan T, hands RE, Ogunkolade BW et al (2006) SPUD: a qPCR assay for the detection of inhibitors in nucleic acid preparations. *Anal Biochem* 351:308–310
8. Vandesompele J, De Preter K, Pattyn F et al (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:research0034.1–research0034.11
9. Hellemans J, Mortier G, De Paepe A et al (2007) qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol* 8:R19
10. Ruijter JM, Ramakers C, Hoogaars WMH et al (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res* 37:e45–e45
11. Berruti A, Christiaens A, De Keyser E et al (2015) Cold treatment breaks dormancy but jeopardizes flower quality in *Camellia japonica* L. *Front Plant Sci* 6:983
12. Ouyang L, Leus L, De Keyser E et al (2019) Seasonal changes in cold hardiness and carbohydrate metabolism in four garden rose cultivars. *J Plant Physiol* 232:188–199
13. Merlaen B, De Keyser E, Van Labeke MC (2018) Identification and substrate prediction of new *Fragaria x ananassa* aquaporins and expression in different tissues and during strawberry fruit development. *Hortic Res* 5:20
14. Dierck R, Leus L, Dhooghe E et al (2018) Branching gene expression during chrysanthemum axillary bud outgrowth regulated by strigolactone and auxin transport. *Plant Growth Regul* 86:23–36
15. Luu-The V, Paquet N, Calvo E, Cumps J (2005) Improved real-time RT-PCR method for high-throughput measurements using second derivative calculation and double correction. *BioTechniques* 38:287–293



Chapter 10

Gene Expression Analysis in Bacteria by RT-qPCR

**Danilo J. P. G. Rocha, Thiago L. P. Castro, Eric R. G. R. Aguiar,
and Luis G. C. Pacheco**

Abstract

Reverse-transcription quantitative real-time polymerase chain reaction (RT-qPCR) using fluorescent DNA-binding dyes is now a gold-standard methodology to study bacterial gene expression through relative quantitation of target mRNAs under specific experimental conditions, and recent developments in the technology allow for gene expression analysis in single cells. Nevertheless, several critical steps of the RT-qPCR protocol need to be carefully addressed in order to obtain reliable results, particularly regarding RNA sample quality and appropriate choice of reference genes. Besides, accurate reporting of study conditions is essential, as recommended by the MIQE guidelines. Herein, we provide a practical approach to quantitation of the transcript levels of bacterial genes using RT-qPCR, including a general protocol for obtaining good-quality bacterial RNA and a discussion on the selection and validation of candidate bacterial reference genes for data normalization.

Key words RT-qPCR, Gene expression, Bacteria, RNA extraction, Reference genes

1 Introduction

Since its introduction, nearly 25 years ago, the methodology of reverse-transcription quantitative real-time polymerase chain reaction (RT-qPCR) has been extensively used for bacterial gene expression analysis through relative quantitation of target mRNAs under specific experimental conditions in comparison to appropriate controls [1–4]. The use of this methodology has already contributed over the years to quantitate the expression of virulence genes in various bacterial pathogens during contact with host cells [5], to study the expression of many antimicrobial resistance genes [6], to determine the transcript levels of genes composing different biochemical pathways [7], and to assess bacterial metabolism in cultures and in environmental samples [8], just to name a few applications. Even though PCR was already being used in the mid-1990s as a quantitative methodology in substitution to methods such as Northern blotting and RNase protection assay, the

widespread adoption of real-time RT-qPCR as the method of choice for quantitation of gene expression derived from its advantages over so-called end-point quantification techniques, which relied on post-PCR manipulations and quantitated products after reaching plateau stage of amplification [3, 9]. Real-time detection of amplified products during exponential phase by RT-qPCR significantly improved sensitivity, dynamic range, and accuracy of target mRNA quantitation [9]. To account for gene expression heterogeneity between cells, incremental changes have been introduced in the recent years in the methodology, which permit the quantitation of gene expression in single bacterial cells using an RT-qPCR protocol [10].

Nevertheless, in order to get accurate gene expression analysis by RT-qPCR, critical points need to be addressed that include: (1) appropriate study design; (2) quality control of RNA extraction and storage; (3) control for gDNA contamination; (4) use of validated reference genes; and (5) choice of a proper data analysis workflow. The Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines were proposed in 2009 and describe best practices for performing and reporting quantitative real-time PCR experiments [11]. The adoption of MIQE criteria significantly contributed to the publications of more reliable gene expression studies in recent years [12, 13], and several MIQE-compliant practical guides for performing RT-qPCR experiments can now be found elsewhere [14–17]. Particularly for bacterial gene expression analysis, since publication of the MIQE guidelines, nearly 10 years ago, there was a significant rise in the number of studies dedicated to validation of bacterial reference genes to be used in RT-qPCR normalization [18]. In this chapter, we provide a practical approach to quantitation of the transcript levels of bacterial genes using RT-qPCR, with special focus on the critical steps of bacterial RNA extraction and purification and on the selection and validation of appropriate bacterial reference genes for data normalization.

2 Materials

Use molecular biology grade reagents and nuclease-free microtubes (200 µl and 1.5 ml) and barrier tips (10 µl, 100 µl and 1000 µl) in all steps of RNA extraction, purification, and analysis, and during RT-qPCR setup. Take preventive measures to avoid RNase contamination, as described in [19] (*see Note 1*).

2.1 RNA Extraction and Purification

1. Tris-EDTA buffer: 30 mM Tris-HCl, 1 mM EDTA, pH 8.0; optional: 15 mg/ml lysozyme (*see Note 2*).
2. RNAProtectTM Bacteria Reagent (QIAGEN).

3. TRIzol™ reagent (ThermoFisher).
4. RNeasy® Mini Kit (QIAGEN).
5. SV Total RNA Isolation System kit (Promega).
6. Molecular biology grade chloroform.
7. Absolute ethanol.
8. 0.1% diethylpyrocarbonate (DEPC)-treated water [19].
9. Microcentrifuge.
10. Spectrophotometer μDrop™ Plate (Thermo Scientific).

2.2 Reverse Transcription and qPCR

1. GoScript™ Reverse Transcription System (Promega).
2. GoTaq® qPCR Master Mix kit (Promega) and specific primer pairs.
3. Specific plasticware: barcoded 96-well PCR plates (Bio-Rad).
4. Microseal® ‘B’ PCR Plate Sealing Film (Bio-Rad).
5. Equipment: thermocycler; CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad).

3 Methods

3.1 RNA Isolation and Purification

3.1.1 RNA Isolation

The quick protocol for bacterial RNA extraction and purification is shown in Fig. 1.

1. Place 500 µl of a grown bacterial culture into a new 2 ml microcentrifuge tube and add 1 ml of RNAProtect™ BacteriaReagent (QIAGEN). Vortex well and incubate for 5 min at room temperature.
2. Centrifuge the tube for 10 min at 10,000 × *g*. Remove the supernatant by inversion and proceed to the next step. If necessary, store the stabilized bacterial pellet at –70 °C for up to 1 month.
3. Add 200 µl of the 15 mg/ml lysozyme solution, at room temperature, to the thawed bacterial pellet (*see Note 2*).
4. Add 1 ml of TRIzol™ reagent to the tube and vortex vigorously. Incubate at room temperature for 5 min.
5. Add 200 µl of chloroform to the mixture and vortex vigorously. Centrifuge the microtube for 18 min, at 13,000 × *g* and 4 °C.
6. Carefully transfer 200 µl of the upper aqueous phase to a new microtube, using a micropipette tip (*see Note 3*).
7. Add 700 µl of the RNeasy® Lysis Buffer (QIAGEN) and 500 µl of ethanol (96–100%) to the separated aqueous phase.
8. Mix thoroughly by pipetting up and down and transfer to a new RNeasy® spin column (QIAGEN) with silica-gel

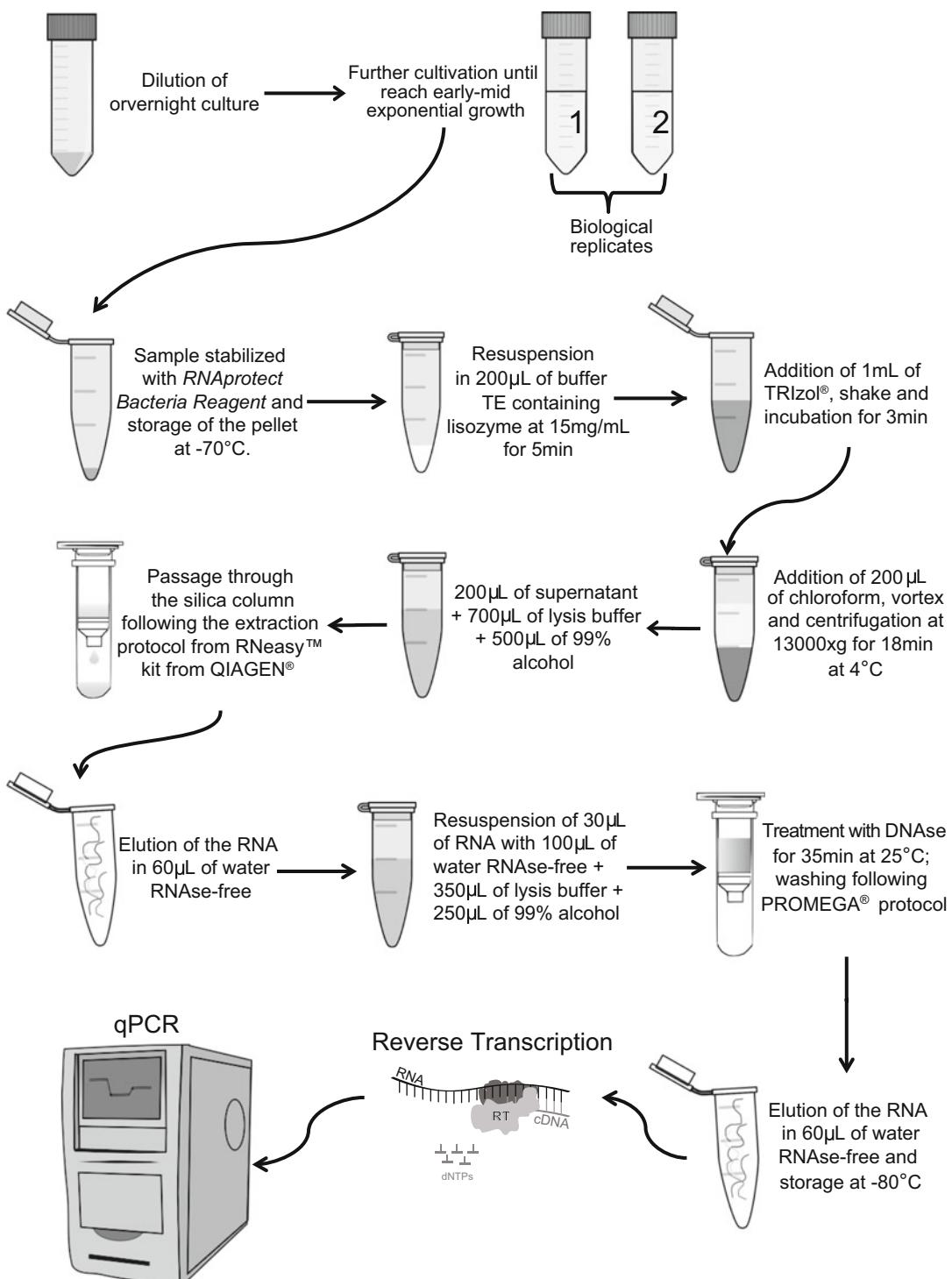


Fig. 1 A general protocol for extracting high-quality bacterial RNA

membrane, sitting in a new 2 ml collection tube. Centrifuge for 30 s at $8000 \times g$.

9. Discard the flowthrough and add 700 μ l of the RNeasy® Wash Buffer (QIAGEN) to the column. Centrifuge for 30 s at $8000 \times g$.
10. Discard the flowthrough and add 500 μ l of the RPE Buffer (QIAGEN) to the column. Centrifuge for 30 s at $8000 \times g$.
11. Repeat **step 10** and place the column in a new 1.5 ml collection tube.
12. Elute the total RNA by placing 30 μ l of RNase-free water onto the center of the silica-gel membrane. Centrifuge for 2 min at maximum speed.
13. Repeat **step 12** using the same 1.5 ml collection tube. Transfer the eluate to a new microtube and store at -70°C .

3.1.2 RNA Cleanup (Optional)

The RNA cleanup procedure is recommended in the presence of any trace of genomic DNA or other contaminants in the total RNA sample. The present protocol combines the use of a silica-gel membrane column with on-column DNase treatment, using the SV Total RNA Isolation System kit (Promega).

1. Add RNase-free water to the RNA eluate to obtain a final volume of 100 μ l.
2. Add 350 μ l of the Lysis Buffer and 250 μ l of ethanol (96–100%).
3. Mix thoroughly by pipetting up and down and transfer approximately 700 μ l to the Spin Column (Promega) sitting in a collection tube. Centrifuge for 30 s at $8000 \times g$.
4. Prepare the DNase incubation mix by carefully combining 40 μ l of the Yellow Core Buffer, 5 μ l of 90 mM MnCl₂, and 5 μ l of RQ1 RNase-Free DNase I in a new microtube, and tap lightly with finger tips (*see Note 4*).
5. Apply the mix prepared in **step 4** directly onto the center of the silica-gel membrane and incubate for 15–25 min at room temperature.
6. Add 200 μ l of DNase Stop Solution to the Spin Column and centrifuge for 1 min at 13,000 $\times g$. Discard the flowthrough.
7. Add 600 μ l of RNA Wash Solution to the Spin Column and centrifuge for 1 min at 13,000 $\times g$. Discard the flowthrough.
8. Add 250 μ l of RNA Wash Solution and centrifuge for 2 min at maximum speed.
9. Place the Spin Column into a new elution microtube and place 60 μ l of RNase-free water onto the center of the silica-gel membrane. Centrifuge for 1 min at 13,000 $\times g$.

3.2 RT-qPCR Setup

3.2.1 Reverse Transcription

The following protocol is used to convert up to 5 µg total RNA in accordance with the GoScript™ Reverse Transcription System (Promega) manufacturer's recommendations.

1. Incubate the total RNA-containing tube for 5 min at 70 °C in a preheated heat block. Certify that the tube lid is tightly closed.
2. Transfer the tube immediately to ice and incubate for 5 min. Alternatively, use a thermal cycler programmed to keep at 4 °C and take the tube when proceeding to the next step. Centrifuge briefly to settle down the sample volume.
3. In a new sterile microtube, prepare a mixture containing the RNA sample, random primers and nuclease-free water to obtain a final volume of 5 µl (*see Note 5*).
4. In a separate sterile microtube, prepare the reverse transcription reaction mix on ice, using the reagents in the following order: GoScript™ 5× Reaction Buffer (4 µl); MgCl₂ (1.5–5 mM); PCR Nucleotide Mix (0.5 mM each dNTP); Recombinant RNasin® Ribonuclease Inhibitor (20 units) (optional); GoScript™ Reverse Transcriptase (1 µl); Nuclease-Free water (to a final volume of 15 µl) (*see Note 6*).
5. Transfer 15 µl of the mixture obtained in **step 4** to the tube prepared in **step 3**, containing the RNA sample and random primers. The final volume of the reaction should be 20 µl.
6. Place the tube in a thermal cycler and run the reaction with the following conditions: primer annealing at 25 °C, for 5 min; extension at 42 °C, for 1 h; reverse transcriptase inactivation at 70 °C, for 15 min (*see Note 7*).

3.2.2 qPCR Preparation

1. Design and print out a plate template that corresponds to the qPCR plate of choice. Use the template to identify wells and plan the distribution of samples across the qPCR plate. Every cDNA template and negative control reaction should be prepared at least in duplicate (*see Note 8*).
2. Thaw the qPCR Master Mix at room temperature. Homogenize gently using vortex and centrifuge briefly, avoiding foam formation and light exposure. Keep the reagents on ice until the use.
3. In a new sterile microtube, prepare the qPCR reaction mix using the reagents in the following order: 1× GoTaq® qPCR Master Mix 2×; 0.2 µM forward primer; 0.2 µM reverse primer; nuclease-free water (to a final volume of 16 µl) (*see Note 9*).
4. Carefully transfer 16 µl of the qPCR reaction master mix to the wells in the plate, in accordance with the plate template prepared in **step 1**. Keep the plate at room temperature.

5. Carefully transfer 4 μ l of the cDNA template to the corresponding wells in the plate. Lastly, use 4 μ l of nuclease-free water to prepare the no-template control (NTC) reactions.
6. Carefully seal the plate using an optical adhesive. Apply force over the spaces between the wells using a clean and sharp plastic object. Make sure the wells are properly sealed before proceeding to the next step.
7. Centrifuge the plate at low speed for 1 min to settle down the reaction volume and eliminate air bubbles.
8. Carefully place the qPCR plate onto the thermal cycler block (*see Note 10*).
9. Setup the cycling program, as follows: 1 cycle of Hot-Start polymerase activation at 95 °C for 2 min; 40 cycles of denaturation at 95 °C for 15 s, and annealing/extension at 60 °C for 60 s; end-point dissociation curve at 60–95 °C. Make sure that the file is properly saved and start the run. The fluorescence detection and melting curve analysis options should be enabled in the thermal cycler (*see Note 11*). Table 1 shows a comparison of software commonly used to manage qPCR reactions and to analysis of qPCR results.

3.3 Assessment of Contamination by Genomic DNA

It is recommended to assess gDNA contamination in the total RNA samples because this may affect the amplification levels of a specific gene transcript, impairing the relative quantification of gene expression. The verification is done by preparing a reverse transcription reaction without the reverse transcriptase enzyme and running qPCR reactions with a set of primers.

1. Separate the same volume (4 μ l) of RNA used for reverse transcription in a new microtube. We recommend doing this in parallel with the preparation of the reverse transcription assays shown on Subheading 3.2.1.
2. In a new microtube, mix the reagents used for reverse transcription but omitting the reverse transcriptase enzyme. In its place, use nuclease-free water. Adjust the volumes to prepare a master mix if necessary.
3. Transfer 16 μ l of the prepared master mix to the tube containing the RNA sample. The final volume should be 20 μ l.
4. Place the microtube in a thermal cycler and run the reaction using the same conditions shown in Subheading 3.2.1.
5. Separate wells in the qPCR plate template for the gDNA contamination assessment. Set up the qPCR reactions as shown in Subheading 3.2.2. Use the product of the no-reverse transcriptase reaction as a template.

Table 1
Commonly used software for management of qPCR reactions and for analysis of qPCR results

Tool	Features	Input	Compatibility	Limitations	Access type	Reference
REST (QIAGEN)	<ul style="list-style-type: none"> - Takes into consideration the different PCR efficiencies of the gene of interest and reference genes where multiple reference genes are used for normalization - Estimates up and down regulation for gene expression studies - Analysis using randomization and bootstrapping techniques - Graphical data output via whisker-box plots - Free 	C_q values	Rotor-Gene® Q and other thermocyclers	<ul style="list-style-type: none"> - Works in computer with Operating system Windows® (32-bit only) - Limited to comparison of two groups, with up to 16 data points in a sample and 16 in a control group, for reference and up to four target genes 	Stand-alone	[27]
CFX Maestro (Bio-Rad)	<ul style="list-style-type: none"> - Allows analysis of the entire process of real-time PCR, including reference gene selection, complex experiment setup, statistical analysis, and data graphing - Charges apply 	Read directly from Bio-Rad real-time PCR machines	Bio-Rad real-time PCR machines	<ul style="list-style-type: none"> - Machine-specific - Work only in Microsoft Windows® operating system 	Stand-alone	[28]

For a comprehensive comparison between most used software tools, please check [29]

6. Check for amplification curves in the wells corresponding to the no-reverse transcriptase control, following the incubation in a qPCR thermal cycler. Contaminant C_q values are acceptable next to the end of cycling and when they are away from the C_q values generated for the cDNA templates.
7. In case the contamination by gDNA is prominent, proceed with the total RNA sample cleanup using the RNA cleanup protocol. Treat the sample with DNase if necessary (*see Note 12*).

3.4 Determination of Primer Efficiencies

The efficiency of amplification of a set of primers in a qPCR reaction depends directly on the sequences of oligonucleotides [11]. The efficiency should be experimentally determined for every set of primers by preparing a tenfold serial dilution of five points for a given cDNA sample (i.e., 10^0 , 10^{-1} , 10^{-2} , 10^{-3} , and 10^{-4}). Prepare the qPCR reactions as described on Subheading 3.2.2. To pipette only cDNA sample and primers in the reaction, use a kit of qPCR Master Mix that already contains a fluorescent DNA-binding dye (e.g., GoTaq® qPCR Master Mix commercialized by Promega). After performing the reactions, detect the quantification cycle (C_q) values generated by the thermal cycler to make a dispersion plot and perform linear regression analysis. Plot C_q values against log of template concentration. The slope value for the regression line should be used in the equation $E = (10^{(-1/\text{slope})} - 1) \times 100$, where “E” is the percentage of amplification efficiency [20]. The amplification efficiency can also be determined using specialized online tools, such as the *qPCR Efficiency Calculator* (available at <https://www.thermofisher.com>).

Follow the steps below to calculate the amplification efficiency for a set of primers:

1. Setup the qPCR reactions using a pair of primers and a cDNA sample of choice. Use a tenfold serial dilution of the cDNA sample for the reactions. Prepare five reaction replicates for every dilution point to improve the efficiency calculation precision.
2. Calculate the average C_q for every cDNA dilution using Microsoft Excel®. Use the C_q averages to generate a dispersion plot with a linear regression curve. The r^2 value should be close to 1 (*see Note 13*).
3. Use the slope value to calculate the amplification efficiency with the equation “ $E = (10^{(-1/\text{slope})} - 1) \times 100$ ”. If percentages are lower than 90.0% the amplification efficiency is poor and new primers should be designed and tested (*see Note 14*).

3.5 Reference Gene Selection and Assessment of Their Expression Stability

3.5.1 Reference Gene Selection

The selection of the organism to study as well as the desired test conditions (e.g., specific treatment and suitable control) is necessary, in order to determine reference genes for performing the RT-qPCR analyses. The reference genes are endogenous control genes and their expression levels should vary as little as possible, according to the organism of choice and experimental conditions tested [11, 13]. Some authors suggested a selection of candidate reference genes for bacteria [18], based on the meta-analysis of gene expression data retrieved from the *Pathosystems Resource Integration Center* (www.patricbrc.org) [21]. For different bacterial phyla, high stability in expression was observed for the genes *gyrB*, *gyrA*, *era*, *secA*, and *dnaG*. These genes encode for the DNA gyrase subunit B, DNA gyrase A, GTP-binding protein Era, translocase subunit SecA and DNA primase, respectively. Although these genes constitute a preliminary set of possible reference genes, other housekeeping genes might provide similar or greater expression stability for the test conditions given the organism of choice [18] (see Note 15).

3.5.2 Assessment of Gene Expression Stability

It is important to determine proper reference genes to make reliable relative expression calculations. To find out the reference genes with the least C_q variation throughout all experimental conditions, proceed with the protocol shown below.

1. Download the NormFinder add-in for Microsoft Excel®, available at: <https://www.moma.dk/normfinder-software>. Alternative software options for analysis of gene expression stability are shown in Table 2.
2. Run the qPCR reactions for the candidate reference genes using the cDNA templates from all experimental conditions. Linearize the C_q data using the equation $Q = E^{\Delta C_q}$ [22], where E is the efficiency of amplification for a specific gene and ΔC_q is the difference between the average C_q obtained for this gene and the lowest C_q value in the dataset. Every candidate reference gene should be evaluated using all experimental conditions. The efficiency amplification factor can be calculated in accordance with the steps shown in Subheading 3.4.
3. Setup a Microsoft Excel® table with the linearized C_q data obtained for each experimental condition and execute the NormFinder add-in.
4. Select the cells containing the data and follow the instructions provided in the appearing window. Execute the algorithm.
5. A new tab will open up with the stability values of the genes analyzed. The least value shown will correspond to the greatest expression stability (see Note 16).

Table 2
Software tools available for analysis of the stabilities of candidate reference genes

Tool	Features	Limitations	Access type	Reference
RefFinder	<ul style="list-style-type: none"> - Uses information of consolidated methods (BestKeeper, NormFinder, geNorm, and dela-ct) to rank and select best genes - Free - Accessible through the website 	<ul style="list-style-type: none"> - Access to internet - Intrinsic limitations from the tools integrated in the pipeline 	Web-based	[30]
BestKeeper	<ul style="list-style-type: none"> - Combines expression of up to ten genes generating an index of stability based on variation of the PCR values - Free (access key under request) - Microsoft Excel®-based software (Macro) 	<ul style="list-style-type: none"> - Limited up to 10 genes in each analysis - Depends on commercial Microsoft Excel® software 	Stand-alone	[31]
NormFinder	<ul style="list-style-type: none"> - Takes advantage of mathematical model of gene expression that enables accurate estimation in any kind of experimental design - Analyzes data derived from any quantitative method (RT-PCR, microarray, etc.) - Free - Microsoft Excel® and R-based versions 	<ul style="list-style-type: none"> - Microsoft Excel® version need commercial software - R version require minimum programming knowledge 	Stand-alone	[32]
Genorm (qbase+)	<ul style="list-style-type: none"> - Gene expression normalization factor is calculated for each sample based on the geometric mean of a user-defined number of reference genes - Charges applied (with 15-day trial) - Previously Microsoft excel®-based (Visual Basic applet), now built in qbase+ application 	<ul style="list-style-type: none"> - Costly 	stand-alone	[33]
RefGenes (Genevestigator®)	<ul style="list-style-type: none"> - Allows to search for genes that are most stable across a chosen set of samples based on microarray data - Contains database with well-curated microarray experiments 	<ul style="list-style-type: none"> - Better performance if its databases have similar experiments (at least 50 microarray 	Stand-alone	[34]

(continued)

Table 2
(continued)

Tool	Features	Limitations	Access type	Reference
	<ul style="list-style-type: none"> - Free - Integrated within GENEINVESTIGATOR® desktop application 	<p>experiments from three independent studies)</p> <ul style="list-style-type: none"> - Candidate genes need to be assessed by others tools (such as above) 		

For detailed comparison between tools, check [35]

**3.6 Case Study:
Reference Gene
Validation for Gene
Expression Analysis in
Escherichia coli Under
Heat Shock Stress**

The following data reporting and analysis format is based on the following references [13, 14, 23]. This could serve as a template for describing RT-qPCR experiments with different bacterial species.

1. Experimental design: *E. coli* cultures were grown to early exponential (EE) or late exponential (LE) phases and split into 10 ml aliquots; two aliquots were maintained at 37 °C and two aliquots were submitted to heat shock, for 10 min. Five “housekeeping” genes where selected from the recent literature as candidate reference genes (see Subheading 3.5). Primer pairs were designed, primer efficiencies calculated, and expression stabilities were evaluated in the conditions of study.
2. Source of samples: *E. coli* strain BL21; overnight 50 ml culture in Luria–Bertani broth, at 37 °C and 200 rpm agitation; in the following day, 1:50 dilution and growth at 37 °C until reach OD_{600nm} = 0.5 or 2.0, for early exponential and late exponential, respectively; for heat shock, incubation at 47 °C, for 10 min; 500 µl of each culture were used for RNA extraction.
3. Method of preservation/stabilization of RNA in cultures: 1 ml RNAProtect Bacteria Reagent (QIAGEN) added to 500 µl sample.
4. Storage time/handling: After 5 min in the stabilization solution, samples were centrifuged for 10 min. at 10,000 × g and were immediately used for RNA extraction (see Subheading 3.1; Fig. 1).
5. Extraction method: Extraction with 1 ml TRIzol™ reagent (ThermoFisher) followed by RNA purification with spin columns and reagents of the RNeasy® Mini Kit (QIAGEN) (see Subheading 3.1; Fig. 1).

Table 3**Spectrophotometric results: concentration and purity of extracted RNA samples**

RNA sample	Before DNase treatment		After DNase treatment and cleanup	
	Conc. (ng/μl)	260/280 ratio	Conc. (ng/μl)	260/280 ratio
EE (control_01)	41.6	2.2	36.1	2.2
EE (control_02)	36.7	1.5	16.1	2.2
EE (heatshock)	38.4	2.2	23.1	2.2
LE (control_01)	85.4	2.2	45.5	2.1
LE (control_02)	98.7	2.2	44.8	2.2
LE (heatshock)	97.2	2.1	47.0	2.2

EE early exponential, LE late exponential

6. RNA:DNA-free: DNase treatment was performed with 5 μl RQ1 RNase-Free DNase I (Promega), for 25 min., 25 °C; RNA cleanup was done as described on Subheading 3.1 and Fig. 1. To evaluate gDNA contamination, a no-reverse transcriptase control, in which the RNA sample was directly used as template in the qPCR reaction, was performed as described on Subheading 3.3 (see Note 17).
7. RNA concentration, purity, and integrity: RNA concentration was measured by spectrophotometry in the Thermo Scientific™ μDrop™ Plate equipment, using replicates of 2 μl samples (see Note 18) (see Table 3).
8. Reverse transcription: Same input RNA for all RT reactions = 200 ng; GoScript™ Reverse Transcription System (Promega); priming with random hexamers; use of RNasin® ribonuclease inhibitor; detailed reaction conditions as described on Subheading 3.2.
9. Primer design: The IDT®PrimerQuest Tool (www.idtdna.com) was used to design primers. PrimerBlast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) and IDT OligoAnalyzer (www.idtdna.com) were used to check primer specificities and secondary structures that may prevent amplification. Amplicons ranging between 75 bp and 150 bp and primer melting temperatures (Tm) between 55 °C and 65 °C were also considered for primer design [24]. Primer sequences, Tm and expected amplicon sizes are given on Table 4.
10. qPCR and primer efficiency calculation: qPCR reactions were performed as carefully described on Subheading 3.2.2. Figure 2 shows an example of end-run dissociation curve, obtained for the reactions targeting the *16S rRNA* transcript. Primer efficiencies were calculated using C_q values obtained for tenfold

Table 4
Primer sequences and general features

Gene	Primer sequence (5'-3')	Tm (°C)	GC%	Amplicon size
<i>dnaG</i>	Forward: GACGAATCCCACGCGTATT	62	52.6	103
	Reverse: TGGAAATTCTTGCCCTGCT	62	47.4	
<i>gyrB_V</i>	Forward: GAAATTCTCCTCCCCAGACCA	60	50.0	80
	Reverse: GCAGTTCGTTCATCTGCTGT	60	50.0	
<i>gyrB_N</i>	Forward: GTCCTGAAAGGGCTGGATG	62	57.9	120
	Reverse: CGAATACCATGTGGTGCAGA	62	50.0	
<i>secA</i>	Forward: GGTAGTCGTAACGATCGCA	61	52.6	79
	Reverse: TTTCCATCTCCGGTCCAT	61	47.4	
<i>16S rRNA</i>	Forward: TCAAGTCATCATGGCCCTTAC	62	47.6	111
	Reverse: CGGACTACGACGCACTTTAT	62	50.0	

Two different primer pairs were designed and tested for the gene *gyrB*, targeting different regions of the gene

serial dilutions of target cDNAs, as described in Subheading 3.4. Average C_q values were plotted against log template concentrations and results of efficiency calculations are shown in Table 5 (see Note 14).

11. Reference gene validation: The primer pairs targeting the four reference genes chosen from literature search [18], whose amplification efficiencies were within the expected range, were used in RT-qPCR reactions with RNA obtained from the control cultures and heat shock samples. Reactions were performed as described in Subheading 3.2 and assessment of gene expression stabilities was done as described in Subheading 3.5. Figure 3 shows the average C_q values obtained with the five primer pairs, in control growth conditions or during heat shock (H). These C_q values and the primer efficiencies were then used in NormFinder to calculate expression stabilities. Stability (S) values obtained were as follows: 0.014 for *dnaG*; 0.014 for *secA*; 0.128 for *gyrB*; and 0.351 for *16S*. A combined analysis with different software (see Table 2) indicated that *dnaG* was the reference gene with the most stable expression in the conditions of this study.

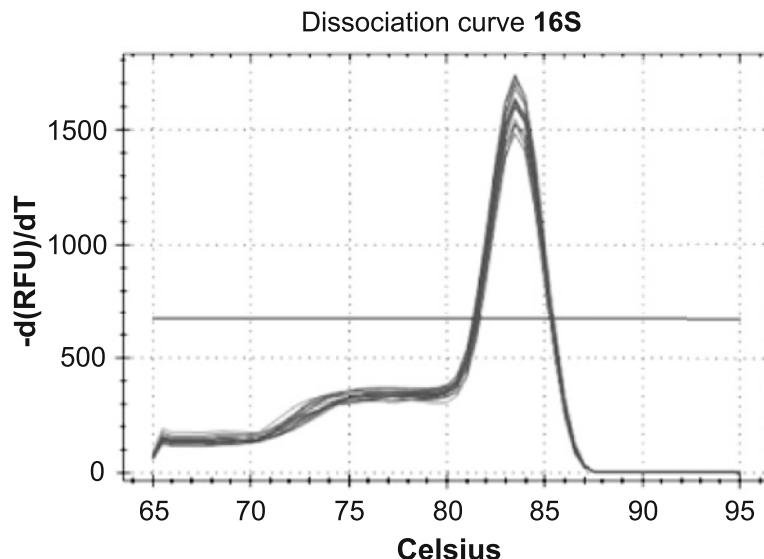


Fig. 2 End-run dissociation curve of RT-qPCR reaction targeting the *E. coli* 16S rRNA

Table 5
Primer efficiency calculation based on tenfold serial dilutions of template cDNA

Primers	Slope	R ²	Efficiency
<i>gyrB</i>	-3.522	0.99	1.92 (92.26%)
<i>gyrBV</i>	-3505	0.99	1.93 (92.89%)
<i>16S</i>	-3.500	0.99	1.93 (93.07%)
<i>dnaG</i>	-3.315	0.99	2.00 (100.29%)
<i>secA</i>	-3.237	0.96	2.04 (103.65%)

Two different primer pairs targeting the *gyrB* transcript rendered similar efficiencies

4 Notes

1. To avoid RNase contamination, wear sterile disposable gloves while touching work surfaces and handling samples, laboratory equipment and reagents. Change your gloves after touching contaminated surfaces. Avoid working in areas with intense people circulation. Decontaminate all work surfaces and materials using 2% sodium hypochlorite solution, 70% ethanol solution, or any commercial RNase decontamination reagent. Use only RNase-free solutions and plasticware (e.g., microcentrifuge tubes and micropipette tips).

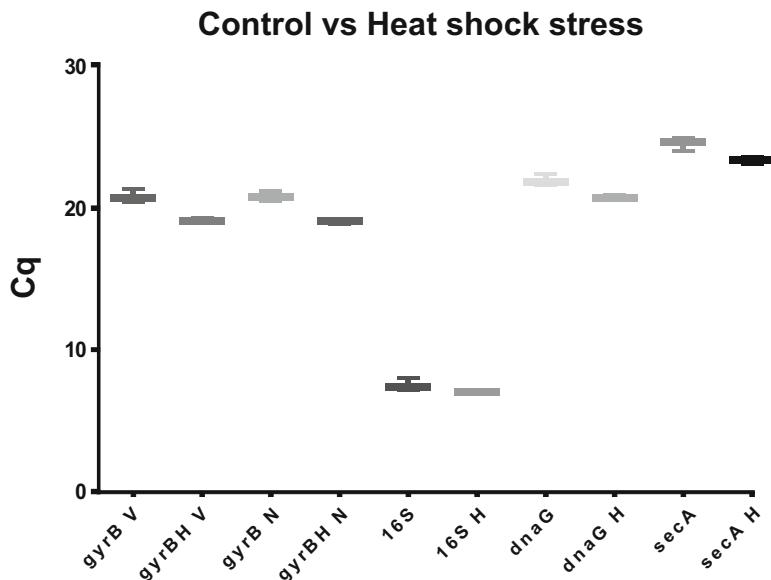


Fig. 3 Average C_q values obtained for the candidate reference genes, in control culture and under heat shock stress

2. Enzymatic lysis with a glycanase (e.g., lysozyme) is recommended for Gram-positive bacteria.
3. Avoid disturbing the DNA between the downer heterogeneous organic and upper aqueous phases.
4. DNases are sensitive to mechanic forces.
5. Mix the components gently with vortex and centrifuge briefly before pipetting. Keep the tube on ice during the preparation. When processing two or more RNA samples, make sure the microtubes are properly identified with permanent marker. Use microtubes compatible with a thermal cycler.
6. Mix the components gently with vortex and centrifuge briefly before pipetting. Make a master mix if two or more reactions are being prepared.
7. The cDNA sample obtained in the reaction is ready to be used in qPCR. Keep the cDNA sample on ice when handling and store at -20°C .
8. Negative controls are prepared to check DNA contamination during the qPCR plate preparation. Prepare at least two control reactions for every set of primers used in the qPCR assay.
9. Prepare a qPCR reaction master mix in accordance with the number of reactions needed for a given set of primers. Include the volume for an extra reaction in the reaction master mix preparation for every ten reactions needed. This will minimize pipetting variation and guarantee the volume necessary to fill all wells.

10. Any SYBR® Green I or FAM™ compatible qPCR system can be used with the GoTaq® qPCR Master Mix. Adjust the reaction volume in accordance with the equipment manufacturer's recommendations.
11. When using the CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad), at the end of the run a “.zpcr” file can be retrieved from the thermal cycler using a USB memory stick. In a computer, use the CFX Manager™ Software (Bio-Rad) to open the file and check amplification results.
12. Some DNase treatment protocols use EDTA-containing buffers. EDTA is a well-known PCR inhibitor and should be properly removed using wash buffers. Be aware that this additional processing may decrease the RNA concentration and increase the risk of contamination of samples.
13. According to [24], an r^2 value higher than 0.980 is desirable for RT-qPCR reactions. It may be necessary to delete points at both ends of the dilution series, as the amplification efficiency of these dilution points will be dependent on the dynamic range of the specific assay.
14. The amplification efficiency should be between 90% and 110% [24]. Amplification efficiency of 100% is obtained when C_q values between consecutive cDNA dilutions are 3.32 apart. In this case, the slope of the regression line plotted for the dilution series experiment should be -3.32 . In some occasions, higher cDNA dilutions yield large C_q variation across the replicates. This might result from the low number of target molecules in the cDNA sample and can be addressed with the use of fivefold serial dilution.
15. The *16S* rRNA gene has been broadly used in RT-qPCR studies as a reference gene. However, this gene frequently shows unstable expression levels, and its transcriptional activity usually far exceeds the transcriptional activity of the mRNA genes since the bacterial genome possesses several *16S* rRNA copies [18, 25, 26]. Minimum variation is preferred in a reference gene, therefore the RT-qPCR amplification profiles should not vary more than 3 C_q across all experimental conditions.
16. A “getting started” option is shown when executing the NormFinder add-in. Select this option to clarify any remaining doubts. Further documentation about the NormFinder can be found in the website of the Aarhus University Hospital of Denmark (<https://moma.dk/normfinder-software>).
17. Use of a no-reverse transcriptase control (NRT) is essential when first extracting RNA, to evaluate gDNA contamination. In the NRT reaction, one might expect no C_q value detected or a C_q significantly higher than in the RT+ reaction. Another negative control commonly included in RT-qPCR assays is the

no template control (NTC), in which no template RNA is used.

18. An $A_{260/280}$ ratio around 2.0 is generally regarded as “pure” for RNA. However, spectrophotometry does not provide information on integrity. RNA integrity could be more accurately evaluated using microfluidics (e.g., Bioanalyzer or Experion™) expressed in “RNA Integrity Number” (RIN) or “RNA quality indicator” (RQI). Desirable RIN/RQI values above 7 and 7.8, respectively [24].

Acknowledgments

Work in our group is supported by grants from Fundação de Amparo à Pesquisa no Estado da Bahia (FAPESB), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

1. Heid CA, Stevens J, Livak KJ et al (1996) Real time quantitative PCR. *Genome Res* 6:986–994
2. Gibson U, Heid CA, Williams PM (1996) A novel method for real time quantitative RT-PCR. *Genome Res* 6(10):995–1001
3. VanGuilder HD, Vrana KE, Freeman WM (2008) Twenty-five years of quantitative PCR for gene expression analysis. *BioTechniques* 44 (5):619–626
4. Martinez AN, Lahiri R, Pittman TL et al (2009) Molecular determination of *Mycobacterium leprae* viability by use of real-time PCR. *J Clin Microbiol* 47:2124–2130
5. Faucher SP, Porwollik S, Dozois CM et al (2006) Transcriptome of *Salmonella enterica* serovar Typhi within macrophages revealed through the selective capture of transcribed sequences. *Proc Natl Acad Sci* 103:1906–1911
6. Dumas JL, van Delden C, Perron K et al (2006) Analysis of antibiotic resistance gene expression in *Pseudomonas aeruginosa* by quantitative real-time-PCR. *FEMS Microbiol Lett* 254:217–225. <https://doi.org/10.1111/j.1574-6968.2005.00008.x>
7. Dror B, Savidor A, Salam BB et al (2019) High levels of CO₂ induce spoilage by *Leuconostoc mesenteroides* by upregulating dextran synthesis genes. *Appl Environ Microbiol* 85. <https://doi.org/10.1128/AEM.00473-18>
8. Graham JE, Wantland NB, Campbell M et al (2011) Characterizing bacterial gene expression in nitrogen cycle metabolism with RT-qPCR. *Methods Enzymol* 496:345–372. <https://doi.org/10.1016/B978-0-12-386489-5.00014-2>
9. Williams PM (2009) The beginnings of real-time PCR. *Clin Chem* 55(4):833–834. <https://doi.org/10.1373/clinchem.2008.122226>
10. Gao W, Zhang W, Meldrum DR (2011) RT-qPCR based quantitative analysis of gene expression in single bacterial cells. *J Microbiol Methods* 85(3):221–227. <https://doi.org/10.1016/j.mimet.2011.03.008>
11. Bustin SA, Benes V, Garson JA et al (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55:611–622
12. Bustin SA, Benes V, Garson J et al (2013) The need for transparency and good practices in the qPCR literature. *Nat Methods* 10:1063
13. Taylor SC, Mrkusich EM (2014) The state of RT-quantitative PCR: firsthand observations of implementation of minimum information for the publication of quantitative real-time PCR experiments (MIQE). *J Mol Microbiol Biotechnol* 24(1):46–52
14. Bustin SA, Beaulieu JF, Huggett J et al (2010) MIQE precis: practical implementation of minimum standard guidelines for fluorescence-

- based quantitative real-time PCR experiments. BMC Mol Biol 11:74. <https://doi.org/10.1186/1471-2199-11-74>
15. Derveaux S, Vandesompele J, Hellemans J (2010) How to do successful gene expression analysis using real-time PCR. Methods 50 (4):227–230. <https://doi.org/10.1016/j.meth.2009.11.001>
 16. Taylor S, Wakem M, Dijkman G et al (2010) A practical approach to RT-qPCR—publishing data that conform to the MIQE guidelines. Methods 50:S1–S5
 17. Obolenskaya MY, Kuklin A, Rodriguez R et al (2016) Practical approach to quantification of mRNA abundance using RT-qPCR, normalization of experimental data and MIQE. Biopolymers & Cell 32(3):161–172
 18. Rocha DJ, Santos CS, Pacheco LG (2015) Bacterial reference genes for gene expression studies by RT-qPCR: survey and analysis. Antonie Van Leeuwenhoek 108(3):685–693. <https://doi.org/10.1007/s10482-015-0524-1>
 19. Green MR, Sambrook J (2012) Molecular cloning: a laboratory manual: three-volume set. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
 20. Rasmussen R (2001) Quantification on the LightCycler. In: Rapid cycle real-time PCR. Springer, Heidelberg, pp 21–34
 21. Wattam AR, Davis JJ, Assaf R et al (2016) Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. Nucleic Acids Res 45:D535–D542
 22. Galisa PS, da Silva HA, Macedo AV et al (2012) Identification and validation of reference genes to study the gene expression in *Gluconacetobacter diazotrophicus* grown in different carbon sources using RT-qPCR. J Microbiol Methods 91:1–7. <https://doi.org/10.1016/j.mimet.2012.07.005>
 23. Taylor S, Buchanan M, Basik M (2011) A MIQE case study—effect of RNA sample quality and reference gene stability on gene expression data. Bio-Rad Bulletin 6245
 24. Taylor S, Wakem M, Dijkman G et al (2015) A practical approach to RT-qPCR—publishing data that conform to the MIQE guidelines. Bio-Rad Laboratories Inc, Hercules, CA. 94547
 25. Carvalho DM, de Sa PH, Castro TL et al (2014) Reference genes for RT-qPCR studies in *Corynebacterium pseudotuberculosis* identified through analysis of RNA-seq data. Antonie Van Leeuwenhoek 106:605–614. <https://doi.org/10.1007/s10482-014-0231-3>
 26. Kozera B, Rapacz M (2013) Reference genes in real-time PCR. J Appl Genet 54(4):391–406
 27. Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 30(9):e36
 28. BIO-RAD (2018) CFX Maestro™ Software. <http://www.bio-rad.com/en-us/product/cfx-maestro-software-for-cfx-real-time-pcr-instruments?ID=OKZP7E15>. Accessed 12 Nov 2018
 29. Pabinger S, Rodiger S, Kriegner A et al (2014) A survey of tools for the analysis of quantitative PCR (qPCR) data. Biomol Detect Quantif 1:23–33. <https://doi.org/10.1016/j.bdq.2014.08.002>
 30. Zhang B (2014) Coton EST database. <http://150.216.56.64/referencegene.php>. Accessed 10 Nov 2018
 31. Pfaffl MW, Tichopad A, Prgomet C et al (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. Biotechnol Lett 26(6):509–515
 32. Andersen CL, Jensen JL, Orntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res 64 (15):5245–5250. <https://doi.org/10.1158/0008-5472.CAN-04-0496>
 33. Vandesompele J, De Preter K, Pattyn F et al (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3(7):RESEARCH0034
 34. Hruz T, Wyss M, Docquier M et al (2011) RefGenes: identification of reliable and condition specific reference genes for RT-qPCR data normalization. BMC Genomics 12:156. <https://doi.org/10.1186/1471-2164-12-156>
 35. De Spiegelaere W, Dern-Wieloch J, Weigel R et al (2015) Reference gene validation for RT-qPCR, a note on different available software packages. PLoS One 10:e0122515. <https://doi.org/10.1371/journal.pone.0122515>



Chapter 11

Detection and Characterization of Circulating Tumor Cells by Quantitative Real-Time PCR

Francesca Salvianti, Filomena Costanza, Gemma Sonnati, and Pamela Pinzani

Abstract

We propose two different approaches involving the use of quantitative real-time PCR for the detection or analysis of circulating tumor cells. In one case cells are indirectly identified through the expression of a marker mRNA, while in the other one cells are enriched by size prior to be submitted to mutational analysis for a specific target. Both methods have been successfully applied to the study of circulating melanoma cells.

Key words CTCs, Melanoma, RT-qPCR, Tyrosinase, qPCR, Somatic mutation, BRAFV600E

1 Introduction

1.1 Circulating Tumor Cells (CTCs)

Circulating tumor cells (CTCs) detected in blood from patients with solid tumors are considered a surrogate of tumor tissue and a promising biomarker for the diagnosis, prognosis and monitoring of the disease. They represent a tool to investigate the tumor heterogeneity and to monitor the evolution of the disease. By a minimally invasive repeatable sampling procedure, they can be used longitudinally as a source of information, especially in advanced stages of cancer when a biopsy is not always feasible and the primary tumor does not reflect the disease status anymore.

The study of CTCs is mainly hampered by their scarcity into the circulation, only a few cells within millions of leucocytes and billions of erythrocytes, which implies the use of highly sensitive and specific detection methods.

New techniques for CTC enrichment and isolation, each one with its own advantages and drawbacks, are being developed and have been extensively reviewed elsewhere [1, 2]. Notwithstanding the extreme variability of methods for CTC detection, they all can be classified into three main approaches:

1. methods relying on the biological properties of the cells, i.e. the identification of specific markers expressed by CTCs. This category comprehends a wide range of different techniques among which those based on the immunomagnetic capture of cells expressing particular proteins;
2. methods exploiting the physical properties of the CTCs, such as size and deformability, to distinguish them from the other blood cells;
3. indirect detection of CTCs through the analysis of tumor- or tissue-specific transcripts by reverse transcription quantitative real-time PCR (RT-qPCR) on whole blood or buffy coat; in this case CTCs cannot be visualized but are directly lysed for RNA extraction.

In recent years microfluidic technologies, based either on physical or biological characteristics of CTCs (or on their combination) have emerged as a new tool for CTC detection and isolation even at the single cell level [3, 4].

Quantitative real-time PCR (qPCR) has different applications in the context of CTC identification and characterization. It has been used for the indirect detection of CTCs in a variety of tumors, such as melanoma [5], breast [6] and colorectal [7] cancer, representing a sensitive method to detect specific molecular markers at the DNA or RNA level in CTC enriched samples. The major advantages of RT-qPCR approaches to CTC detection resides in their low cost, quantitative nature, and relative simplicity, since they do not involve a CTC isolation step nor a specific CTC enrichment procedure to increase the proportion of CTCs with respect to white blood cells, but can be carried out directly on RNA extracted from whole blood (with or without red blood cell lysis) or from the buffy coat requiring a (gradient) centrifugation for collection; on the other hand, they rely only on the specificity of the marker chosen as a target and often illegitimate transcription of tissue and tumor-specific transcripts by leucocytes can determine false positive results and require the establishment of appropriate cut-offs.

Here we present two protocols, the first one for the indirect detection of CTCs by qPCR, the second one for the direct isolation of intact CTCs and subsequent molecular characterization. The methods have been optimized for circulating melanoma cells and applied to melanoma patients but could be transferred, with the appropriate modifications, to other types of tumors.

1.2 Indirect Detection and Quantification of Circulating Melanoma Cells by RT-qPCR for Tyrosinase mRNA

The presence of CTCs in melanoma patients can be determined by an indirect method based on the detection by RT-qPCR of the mRNA for tyrosinase, an enzyme which is involved in the biosynthesis of the melanin, thus specifically expressed by melanocytes and melanoma cells. The protocol has been applied to uveal [8] and cutaneous melanoma [9].

Collection of whole blood is the first step in many molecular assays used to study cellular RNA. However, a major problem in such experiments is the instability of the cellular RNA profile in vitro. This is caused both by rapid RNA degradation and by alteration of the expression of certain genes after the blood is drawn. A method that preserves the RNA expression profile during and after phlebotomy is therefore essential for accurate analysis of gene expression in human whole blood.

Within the European project SPIDIA, aimed at developing evidence-based quality guidelines for the pre-analytical handling of blood samples for RNA molecular testing, two pan-European External Quality Assessments demonstrated that blood collection tubes containing a cellular RNA stabilizer allow reliable gene expression analysis within 48 h from blood collection [10].

As a consequence, we adopted PAXgene Blood RNA Tubes containing a RNA preservative agent, for blood collection.

Alternatively blood can be collected in EDTA tubes and processed within a short time from blood draw: RNA can then be extracted from whole blood by using kits such as RNeasy blood mini kit or from the buffy coat separated by centrifugation (e.g., by the RNeasy Mini kit).

1.3 Direct Isolation and Mutational Analysis of Circulating Melanoma Cells

CTCs can be molecularly characterized upon enrichment or isolation. We propose a filtration technique aimed at CTC enrichment on the basis of cell size, followed by DNA extraction and mutational analysis. Filtration methods rely on the observation that tumor cells have larger dimensions with respect to most white blood cell types.

The ScreenCell MB (ScreenCell, France) is a disposable nuclease-free filtration device composed of a filtration tank and a polycarbonate filter, with randomly distributed 6.5 µm circular pores, capped by a removable nozzle/holder which allows the insertion of a collection tube. Before filtration blood is diluted in a specific buffer to lyse red blood cells [11].

The adopted filtration method based on cell size presents the advantage of an unbiased enrichment of CTCs, with respect to other techniques relying on the expression of markers. Moreover, differently from indirect methods for CTC detection, it allows the observation of the cells on the filter prior to DNA isolation.

2 Materials

2.1 Sample Collection for the Indirect Detection of CTCs by Tyrosinase mRNA

1. PAXgene Blood RNA Tube.
2. EDTA Vacutainer Tube.
3. Micropipettes.
4. Centrifuge for 15 ml tube.

5. Microcentrifuge.
6. Filter tips.

2.2 Isolation of Total RNA from Blood Samples for the Indirect Detection of CTCs by Tyrosinase mRNA

1. PAXgene Blood RNA kit or RNeasy blood mini kit (Qiagen, Germany) or a kit for RNA extraction from the buffy coat separated by centrifugation (e.g., by the RNeasy Mini kit, Qiagen).
2. Vortex mixer.
3. Heating block.
4. Shaker incubator.
5. Ethanol (96–100% purity grade).
6. Pipettes.
7. Crushed ice.

2.3 cDNA Synthesis for the Indirect Detection of CTCs by Tyrosinase mRNA

1. Taqman Reverse Transcription Reagents.
2. Capillary electrophoresis (Agilent technologies).
3. Spectrophotometer (e.g. Nanodrop).
4. Template RNA.
5. RNase-free treated water.
6. RT buffer.
7. MgCl₂.
8. dNTP mix.
9. RNase inhibitor.
10. Reverse Transcriptase.
11. Thermal cycler.

2.4 Reverse Transcription-Quantitative Real-Time PCR (RT-qPCR) for Tyrosinase mRNA

1. qPCR instrument (e.g., 7900HT, Life Technologies).
2. Plasticware specific for the qPCR instrument.
3. Pipettes and pipette tips.
4. Taqman Universal PCR Master Mix (Applied Biosystems).
5. Taqman gene expression assay for tyrosinase (ID: Hs001659756_m1, Life technologies).

2.5 CTC Selection and DNA Extraction from ScreenCell MB Devices for Molecular Biology Analysis

1. ScreenCell MB filtration device (ScreenCell, Paris, France).
2. Nuclease-free Eppendorf tubes.
3. LC dilution buffer.
4. Heating block.
5. Microcentrifuge.
6. QIAmp DNA Micro Kit (Qiagen).

2.6 Purification of Genomic DNA from Circulating Tumor Cells Enriched with the ScreenCell System

1. Ethanol (96–100%).
2. 1.5 ml microcentrifuge tubes.
3. Pipettes and pipette tips (to avoid cross-contamination, we recommend pipette tips with aerosol barriers).
4. Thermomixer, heated orbital incubator, heating block, or water bath capable of incubation at 56 °C and 70 °C.
5. Microcentrifuge with rotor for 1.5 ml and 2 ml tubes.
6. Vortexer.

2.7 Quantitative Real-Time PCR (qPCR) for BRAV600E Mutated DNA Quantification in CTCs

1. qPCR instrument.
2. Plasticware specific for the qPCR instrument.
3. Quantitect® Probe PCR Master Mix (Qiagen).
4. Primers/Probes (for primer/probe sequences *see Subheading 3.2.3*).
5. *BRAF* wild type MCF7 cell line.
6. *BRAF* mutant SKMEL28 cell line.

3 Methods

3.1 Indirect Method for CTC Detection

3.1.1 Sample Collection

We describe a method based on a RT-qPCR assay targeting tyrosinase mRNA for the indirect detection of circulating melanoma cells.

Peripheral blood samples (2.5 ml) are collected in PAXgene Blood RNA tubes (Qiagen) that enable the collection, stabilization, storage, and transportation of human whole blood specimens at room temperature for 3 days. The system requires the use of this tube for blood collection and stabilization, followed by RNA purification using the PAXgene Blood RNA Kit. This technology, besides protecting RNA against degradation, minimizes ex-vivo gene expression changes, thus guaranteeing a high reproducibility of RT-qPCR [12].

3.1.2 RNA Extraction

Refer to PAXgene Blood RNA kit Handbook for a detailed description of the reagents and the manufacturer's protocol (*see Note 1*)

1. Prepare buffer 2 (BR4) by adding 4 volumes of ethanol (96–100% purity grade) as indicated on the bottle to obtain the work solution.
2. Dissolve the solid DNase I in 550 µl of the DNase resuspension buffer (DRB) provided with the kit. DNase I is especially sensitive to physical denaturation. Mixing should only be carried out gently inverting the tube.

3. Ensure that the PAXgene Blood RNA Tubes are incubated for at least 2 h at room temperature after blood collection to ensure complete lysis of blood cells. Incubation of the PAXgene Blood RNA Tube overnight may increase yields. If the PAXgene Blood RNA Tube was stored at 2–8 °C, –20 °C or –70 °C after blood collection, first equilibrate it at room temperature, and then store it at room temperature for 2 h before starting the procedure.
4. Centrifuge for 10 min at 3000–5000 × g using a swing-out rotor.
5. Remove the supernatant by decanting or pipetting. Add 4 ml RNase-free water to the pellet, and close the tube using a fresh secondary BD Hemogard closure (supplied with the kit).
6. Vortex until the pellet is visibly dissolved, and centrifuge for 10 min at 3000–5000 × g . Remove and discard the entire supernatant.
7. Add 350 μl Buffer BR1, and vortex until the pellet is visibly dissolved.
8. Pipet the sample into a 1.5 ml microcentrifuge tube. Add 300 μl Buffer BR2 and 40 μl proteinase K. Mix by vortexing for 5 s, and incubate for 10 min at 55 °C using a shaker-incubator at 400–1400 rpm. After incubation, set the temperature of the shaker-incubator to 65 °C.
9. Pipet the lysate directly into a PAXgene Shredder spin column (lilac) placed in a 2 ml processing tube, and centrifuge for 3 min at maximum speed (but do not exceed 20,000 × g).
10. Carefully transfer the entire supernatant of the flow-through fraction to a fresh 1.5 ml microcentrifuge tube without disturbing the pellet in the processing tube.
11. Add 350 μl ethanol (96–100%). Mix by vortexing, and centrifuge briefly (1–2 s at 500–1000 × g) to remove drops from the inside of the tube lid.
12. Pipet 700 μl sample into the PAXgene RNA spin column placed in a 2 ml processing tube, and centrifuge for 1 min at 8000–20,000 × g . Place the spin column in a new 2 ml processing tube, and discard the old processing tube containing the flow-through.
13. Pipet the remaining sample into the PAXgene RNA spin column, and centrifuge for 1 min at 8000–20,000 × g . Place the spin column in a new 2 ml processing tube, and discard the old processing tube containing the flow-through.
14. Pipet 350 μl Buffer BR3 into the PAXgene RNA spin column. Centrifuge for 1 min at 8000–20,000 × g . Place the spin column in a new 2 ml processing tube, and discard the old processing tube containing the flow-through.

15. Add 10 μ l DNase I stock solution to 70 μ l Buffer RDD in a 1.5 ml microcentrifuge tube. Mix by gently flicking the tube, and centrifuge briefly to collect residual liquid from the sides of the tube.
16. Pipet the DNase I incubation mix (80 μ l) directly onto the PAXgene RNA spin column membrane, and place on the benchtop (20–30 °C) for 15 min.
17. Pipet 350 μ l Buffer BR3 into the PAXgene RNA spin column, and centrifuge for 1 min at 8000–20,000 $\times g$. Place the spin column in a new 2 ml processing tube, and discard the old processing tube containing the flow-through.
18. Pipet 500 μ l Buffer BR4 into the PAXgene RNA spin column, and centrifuge for 1 min at 8000–20,000 $\times g$. Place the spin column in a new 2 ml processing tube, and discard the old processing tube containing the flow-through.
19. Add another 500 μ l Buffer BR4 to the PAXgene RNA spin column. Centrifuge for 3 min at 8000–20,000 $\times g$.
20. Discard the processing tube containing the flow-through, and place the PAXgene RNA spin column in a new 2 ml processing tube. Centrifuge for 1 min at 8000–20,000 $\times g$.
21. Discard the processing tube containing the flow-through. Place the PAXgene RNA spin column in a 1.5 ml microcentrifuge tube, and pipet 40 μ l Buffer BR5 directly onto the PAXgene RNA spin column membrane. Centrifuge for 1 min at 8000–20,000 $\times g$ to elute the RNA.
22. Incubate the eluate for 5 min at 65 °C in the shaker-incubator without shaking. After incubation, chill immediately on ice.
23. If the RNA samples will not be used immediately, store them at –20 °C or –70 °C.

3.1.3 RNA Quality and Quantity

RNA quantity is evaluated by spectrophotometric measurement (Nanodrop, Thermo scientific, USA).

RNA quality is investigated by capillary electrophoresis (Agilent Bioanalyzer 2100, Agilent Technologies, USA). RNA integrity numbers greater than 7 are considered fully suitable for RT-qPCR analysis.

3.1.4 RNA Reverse Transcription

RNA from blood is reverse-transcribed using a commercial kit based on random primers technique (Taqman Reverse Transcription Reagent's, Applied Biosystem, USA), according to the manufacturer's instructions. The reaction is carried out in a final volume of 20 μ l.

The reaction mixture contains the flowing reagents:

- 1× Buffer.
- 5.5 mM MgCl₂.
- 2 mM dNTP.
- 2.5 µmol/l random hexamers.
- 0.4 U/µl RNase inhibitor.
- 1.25 U/µl Reverse transcriptase.

RNA (500 ng) is added to the mix.

The reaction is performed under the following conditions: 25 °C for 10 min, 48 °C for 30 min and 95 °C for 2 min.

3.1.5 RT-qPCR for Tyrosinase mRNA

For each sample measurement, 5 µl cDNA are used in a reaction volume of 25 µl. The reaction mixture contains 12.5 µl of Taqman Universal PCR Master mix 2×(Applied Biosystems) and 1.25 µl of Taqman gene expression assay 20× (ID:Hs001659756_m1).

The Thermal profile of the reaction is the following: denaturation at 95 °C for 10 min followed by 50 cycles at 95 °C for 15 s and 60 °C for 1 min.

All sample are analysed in duplicate.

To calculate the expression of tyrosinase mRNA in each sample, we refer to an external reference curve (*see Note 2* and Fig. 1).

3.2 Direct Isolation and Mutational Analysis of Circulating Melanoma Cells

As an alternative to the previously reported indirect method for CTC detection, we present a filtration technique aimed at CTC enrichment on the basis of cell size, followed by DNA extraction and mutational analysis for the BRAFV600E variant.

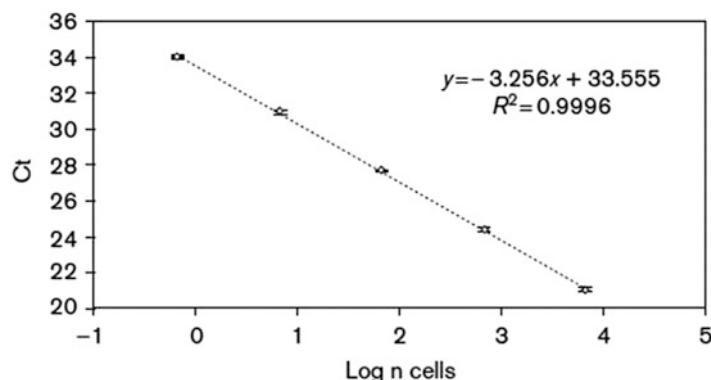


Fig. 1 Standard curve for tyrosinase mRNA quantification

3.2.1 CTC Selection and DNA Extraction from ScreenCell MB Devices for Molecular Biology Analysis

Collect blood samples from patients in EDTA tubes and transport them immediately to the laboratory to be processed within 3 h from drawing. Six ml blood are used in the molecular biology (MB) ScreenCell device. Cells on MB filters are submitted to DNA extraction.

Blood is filtered by the ScreenCell® MB filtration devices (ScreenCell, Paris, France) according to the manufacturer's instructions. ScreenCell® MB device is nuclease-free and devoted to molecular biological studies (*see Note 3*).

1. Before filtration the blood sample is diluted in the ScreenCell® LC dilution buffer.
2. At the end of filtration, the capsule filter is ejected and inserted inside the upper inner part of a nuclease-free Eppendorf® tube. The device allows the extraction of DNA directly from cells retained by the filter.
3. 105 µl lysis buffer (prepared according to ScreenCell manufacturer's instructions) is added into the capsule filter which is then closed with the Eppendorf® tube cap.
4. Following incubation at 56 °C for 10 min, the Eppendorf® tube containing the capsule filter is centrifuged for 1 min at 12,000 × *g* and the capsule filter removed and discarded.

3.2.2 Purification of Genomic DNA from Circulating Tumor Cells Enriched with the ScreenCell System

The flow-through is submitted to DNA extraction using the QIAamp DNA Micro kit (Qiagen) following a modified procedure, as suggested by ScreenCell manufacturer, and used for qPCR analysis.

1. Add 50 µl ethanol (96–100%), close the lid, and mix thoroughly by pulse-vortexing for 15 s. Incubate for 3 min at room temperature (15–25 °C).
2. Briefly centrifuge the 1.5 ml tube to remove drops from the inside of the lid.
3. Carefully transfer the entire lysate into the QIAamp MinElute® column without wetting the rim. Close the lid, and centrifuge at 6000 × *g* (8000 rpm) for 1 min.
4. Place the QIAamp MinElute column in a clean 2 ml collection tube, and discard the collection tube containing the flow-through.

If the lysate has not completely passed through the membrane after centrifugation, centrifuge again at a higher speed until the QIAamp MinElute column is empty.

5. Carefully open the QIAamp MinElute column and add 500 µl Buffer AW1 without wetting the rim. Close the lid and centrifuge at 6000 × *g* (8000 rpm) for 1 min. Place the QIAamp MinElute column in a clean 2 ml collection tube, and discard the collection tube containing the flow-through.

6. Carefully open the QIAamp MinElute column and add 500 µl Buffer AW2 without wetting the rim. Close the lid and centrifuge at $6000 \times g$ (8000 rpm) for 1 min. Place the QIAamp MinElute column in a clean 2 ml collection tube, and discard the collection tube containing the flow-through.

Contact between the QIAamp MinElute column and the flow-through should be avoided. Some centrifuge rotors may vibrate upon deceleration, resulting in the flow-through, which contains ethanol, coming into contact with the QIAamp MinElute column. Careless removal of the QIAamp MinElute column and collection tube from the rotor may also cause flow-through to come into contact with the QIAamp MinElute column.

7. Centrifuge at full speed ($20,000 \times g$; 14,000 rpm) for 3 min to dry the membrane completely.

This step is necessary, since ethanol carryover into the eluate may cause problems in some downstream applications.

8. Place the QIAamp MinElute column in a clean 1.5 ml microcentrifuge tube (not provided) and discard the collection tube containing the flow-through. Carefully open the lid of the QIAamp MinElute column and apply 23 µl AE Buffer to the center of the membrane.

If high pH or EDTA may affect sensitive downstream applications, use water for elution (*see* the QIAamp DNA Micro Handbook).

Important: Ensure that the elution solution is equilibrated to room temperature (15–25 °C). If using small elution volumes (50 µl) dispense the elution solution onto the center of the membrane to ensure complete elution of bound DNA. Remember that the volume of eluate will be up to 5 µl less than the volume of elution solution applied to the column.

9. Close the lid and incubate at room temperature (15–25 °C) for 5 min.
10. Centrifuge at full speed ($20,000 \times g$; 14,000 rpm) for 1 min.
11. Store the purified DNA at –20 °C.

3.2.3 BRAFV600E Mutational Status in CTCs

BRAFV600E mutational status is assessed on DNA extracted by ScreenCell MB enriched samples. We designed a mutation-specific forward primer (5'-AAA ATA GGT GAT TTT GGT CTA GCT ACA GA-3') and a wild type one (5'-AAA ATA GGT GAT TTT GGT CTA GCT ACA GT-3'), while the reverse primer (5'-GAC AAC TGT TCA AAC TGA TGG-3') is common to wild type and mutated sequences. Two dual-labelled LNA probes, one for the mutant allele (5'-6FAM-T[+C]GAGA[+T]TT[+C][+T][+C]TG [+T]AG[+C]T-BHQ1-3', Sigma, USA) and one for the wild type allele (5'-6FAM-CGA[+G]A[+T]TT[+C][+A][+C]TG[+T]AG

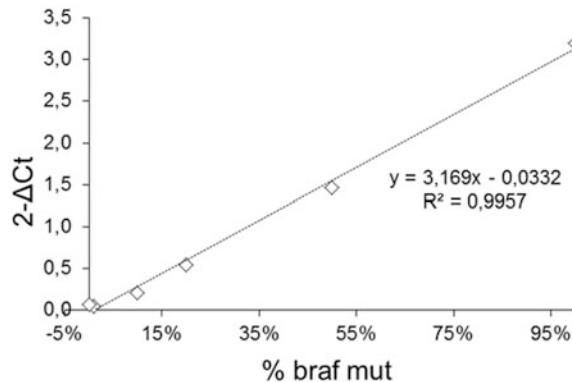


Fig. 2 Example of a standard curve for the quantification of BRAFV600E mutated alleles

[+C]T-BHQ1-3', Sigma, USA) were designed on the complementary strand to that of the BRAFV600E allele specific primer (*see Note 4*).

The procedure is applied to CTC enriched samples. DNA extracted from the ScreenCell MB eluate is used to measure BRAFV600E mutated and wild type allele in two distinct reactions containing the proprietary primers and probe for each variant.

QPCR for both variants is performed by: denaturation at 95 °C for 10 min and 50 cycles of PCR (95 °C for 15 s, 64 °C for 60 s) using Quantitect® Probe PCR Master Mix (Qiagen).

For BRAFV600E standard curve preparation (Fig. 2) (*see Note 5*). Linear regression analysis of the $2^{-\Delta C_t}$ (mut-wt) versus % mutated DNA provides the equation to extrapolate unknown sample results expressed as % of mutated DNA.

4 Notes

1. The PAXgene Blood RNA kit is used for the purification of 2.5 ml of human whole blood collected in PAXgene Blood RNA tubes, according to the manufacturer's instructions. The protocol for the manual extraction of RNA is reported in paragraph Subheading 3.1.2. Alternatively, the same kit can be used for an automated extraction protocol on the QIAcube instrument.
2. Blood from healthy donors is collected in PAXgene Blood RNA tubes and spiked with a given number of cells (as determined by an hemocytometer) of the melanoma cell line SK-MEL-28, to a final concentration range from 4000 to 0.4 cell/ml blood (Fig. 1). RNA is extracted following the same procedure used for the unknown samples. The standard curve is constructed performing RT-qPCR on these samples.

For the standard curve each calibrator is tested in triplicate. The results are expressed as number of SK-MEL-28 cell equivalents/ml blood.

3. Besides ScreenCell MB filters, other types of devices (ScreenCell Cyto) by the same manufacturer are designed to allow the count and cytomorphological characterization of the CTCs enriched from 3 ml blood. We suggest to use these filters in parallel with the ones devoted to DNA extraction in order to obtain both morphological and molecular data on the same sample.
4. The investigation is performed by a modified version of a previously published allele specific real-time qPCR assay [13]. The modified protocol involves the use of a probe for the wild-type allele in order to normalize the results for wild type DNA quantity.
5. The standard curve for BRAFV600E consists of six dilutions (100%, 50%, 20%, 10%, 1% and 0% mutated alleles) obtained by mixing DNA from the mutant SKMEL28 cell line and the wild type MCF7 cell line.

References

1. Chen L, Bode AM, Dong Z (2017) Circulating tumor cells: moving biological insights into detection. *Theranostics* 7:2606–2619. <https://doi.org/10.7150/thno.18588>
2. Alix-Panabieres C, Pantel K (2014) Challenges in circulating tumour cell research. *Nat Rev Cancer* 14:623–631. <https://doi.org/10.1038/nrc3820>
3. Qian W, Zhang Y, Chen W (2015) Capturing cancer: emerging microfluidic technologies for the capture and characterization of circulating tumor cells. *Small* 11:3850–3872. <https://doi.org/10.1002/smll.201403658>
4. Huang L, Bian S, Cheng Y et al (2017) Microfluidics cell sample preparation for analysis: advances in efficient cell enrichment and precise single cell capture. *Biomicrofluidics* 11:011501. <https://doi.org/10.1063/1.4975666>
5. Rodic S, Mihalcioiu C, Saleh RR (2014) Detection methods of circulating tumor cells in cutaneous melanoma: a systematic review. *Crit Rev Oncol Hematol* 91:74–92. <https://doi.org/10.1016/j.critrevonc.2014.01.007>
6. Andergassen U, Kölbl AC, Mahner S et al (2016) Real-time RT-PCR systems for CTC detection from blood samples of breast cancer and gynaecological tumour patients (review). *Oncol Rep* 35:1905–1915. <https://doi.org/10.3892/or.2016.4608>
7. Yang C, Zou K, Zheng L et al (2017) Prognostic and clinicopathological significance of circulating tumor cells detected by RT-PCR in non-metastatic colorectal cancer: a meta-analysis and systematic review. *BMC Cancer* 17:725. <https://doi.org/10.1186/s12885-017-3704-8>
8. Pinzani P, Mazzini C, Salvianti F et al (2010) Tyrosinase mRNA levels in the blood of uveal melanoma patients: correlation with the number of circulating tumor cells and tumor progression. *Melanoma Res* 20:303–310. <https://doi.org/10.1097/CMR.0b013e32833906e3>
9. Salvianti F, Orlando C, Massi D et al (2016) Tumor-related methylated cell-free DNA and circulating tumor cells in melanoma. *Front Mol Biosci* 2:76. <https://doi.org/10.3389/fmbo.2015.00076>
10. Malentacchi F, Pazzaglia M, Simi L et al (2014) SPIDIA-RNA: second external quality assessment for the pre-analytical phase of blood samples used for RNA based analyses. *PLoS One* 9: e112293. <https://doi.org/10.1371/journal.pone.0112293>
11. Desitter I, Guerrouahen BS, Benali-Furet N et al (2011) A new device for rapid isolation

- by size and characterization of rare circulating tumor cells. *Anticancer Res* 31:427–441
12. Rainen L, Oelmueller U, Jurgensen S et al (2002) Stabilization of mRNA expression in whole blood samples. *Clin Chem* 48:1883–1890
13. Pinzani P, Salvanti F, Cascella R et al (2010) Allele specific Taqman-based real-time PCR assay to quantify circulating BRAFV600E mutated DNA in plasma of melanoma patients. *Clin Chim Acta* 411:1319–1324. <https://doi.org/10.1016/j.cca.2010.05.024>



Chapter 12

Molecular Monitoring of Chronic Myeloid Leukemia

Katherine Dominy, Katya Mokretar, Alistair G. Reid,
and Jamshid S. Khorashad

Abstract

Molecular diagnosis and measurement of minimal residual disease (MRD) in patients with chronic myeloid leukemia (CML) is essential for clinical management. In the era of tyrosine kinase inhibitor therapy molecular tests including *BCR-ABL1* transcript monitoring and kinase domain mutation analysis are the main tools used to inform choice of treatment, appropriate dosage and even whether therapy can be safely withdrawn. Quantitation of *BCR-ABL1* oncogene transcript by real-time quantitative PCR (qPCR) is currently the gold-standard method for monitoring as it provides superior sensitivity over karyotyping and fluorescent in situ hybridization (FISH). Here we describe step-by-step methods of RNA conversion to cDNA along with the qPCR protocol which is used in one of the main reference laboratories for this test.

Key words CML, qPCR, *BCR-ABL1*, TaqMan, Multiplex PCR, cDNA synthesis

1 Introduction

Chronic myeloid leukemia (CML) is a pluripotent hematopoietic stem cell malignancy characterized by a balanced reciprocal translocation that involves chromosome 9 and 22 and is designated t(9;22)(q34;q11.0). This translocation fuses the 3' part of *ABL1* gene on chromosome 9 band q34 (9q34) to the 5' part of the *BCR* gene on chromosome 22 [1, 2] (Fig. 1). The resulting BCR-ABL1 fusion protein, a deregulated tyrosine kinase, is central to the pathogenesis of CML and also to Philadelphia positive acute lymphoblastic leukemia (ALL). The majority of CML patients are diagnosed in chronic phase (CML-CP) which has a median duration of 3–4 years without proper treatment [3]. Untreated, the disease inevitably progresses into an accelerated phase (CML-AP) and then into blast phase or blast crisis (CML-BP) [4]. The improved survival resulting from the use of tyrosine kinase inhibitor (TKI) therapy, in addition to the increased life expectancy of the general population in the West, anticipates an increase in prevalence of CML from 70,000 in 2010 to reach a plateau at 35 times that in

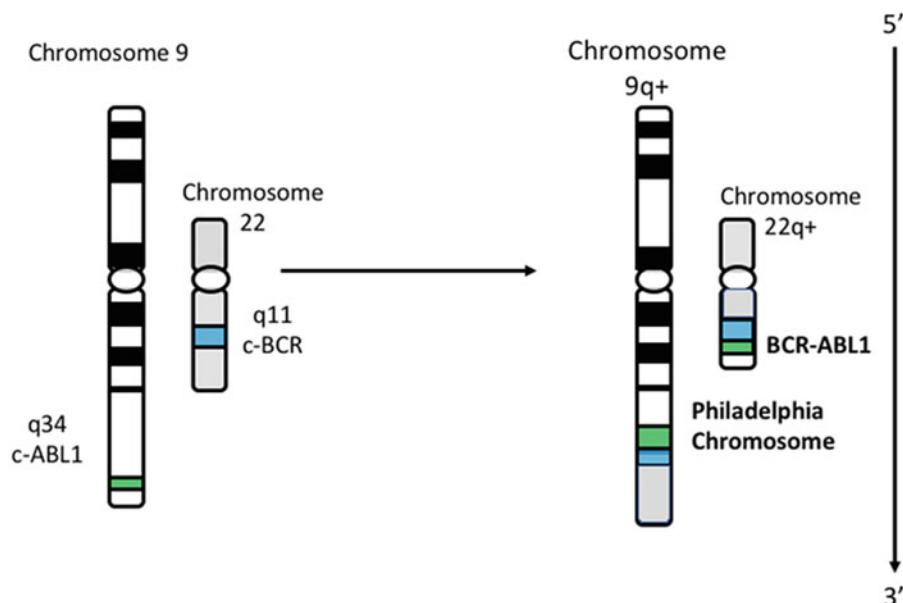


Fig. 1 Reciprocal translocation between *c-ABL1* on chromosome 9 and *c-BCR* on chromosome 22

2050 [5]. Imatinib is the standard first-line therapy for CML patients diagnosed in CP [6]. To overcome resistance or intolerance to imatinib, which is observed in 30% of cases, second-generation (dasatinib, nilotinib, bosutinib) and third-generation (ponatinib) TKIs are used. Allogeneic bone marrow transplantation (BMT) is an option for patients who have failed multiple TKI therapies [7].

1.1 Molecular Diagnosis and Monitoring

Reverse transcription polymerase chain reaction (RT-PCR) is the method of choice for molecular diagnosis of CML [8]. Multiplex PCR, involving the use of multiple primers, is suitable for primary detection and characterization of the *BCR-ABL1* transcript type [9]. Classification of the *BCR-ABL1* transcript type is followed by its measurement by real-time quantitative PCR (qPCR) to determine the leukemic load at baseline. Monitoring CML patients on TKI therapy involves assessment of three laboratory parameters which, in order of increasing sensitivity, are: (1) hematological response, involving degree of normalization of blood counts; (2) cytogenetic response, referring to the percentage of Ph + cells in the bone marrow; and (3) molecular response determined by the level of expressed *BCR-ABL1* transcript in a blood or bone marrow sample [10]. The sensitivity of qPCR enables detection of a single leukemic cell in approximately 10^5 to 10^6 normal cells [11, 12]. qPCR is performed by comparing the transcript level of *BCR-ABL1* to that of a reference gene such as *ABL1* or *GUSB*. The resulting value is expressed as a percentage of *BCR-ABL1* transcripts relative to the level of control transcript.

1.2 Breakpoints and Transcript Types

The *ABL1* gene consists of 11 exons with two alternative first exons termed, in order of transcriptional direction, 1b and 1a. Chromosomal breakpoints in the *ABL1* gene are distributed over a 300 kb region at the 5' end of the gene [13] extending from upstream of exon 1b to downstream of exon 1a between the two exons. Sequences from exons 1a and/or 1b are spliced out of the primary hybrid transcript resulting in mRNA where 5' *BCR* sequences are fused to *ABL1* at exon 2. Rarely, hybrid transcripts in which *BCR* sequences are juxtaposed to *ABL1* exon 3, rather than exon 2, may arise [14, 15].

In contrast, breakpoints within the *BCR* gene are confined to one of three breakpoint cluster regions. In over 95% of CML cases, the break occurs within a 5.8 kb sequence spanning exons 12 to 16, known as the major breakpoint cluster region (*M-bcr*) [16]. Subsequently, transcripts are formed that fuse e13 (40%) or e14 (55%) from *BCR* with a2 from *ABL1* [17]. The e13a2 mRNA can also result from an e14a2 fusion, since e14 is subjected to alternative splicing [18]. Thus, in around 5% of cases both transcripts are concomitantly detected, as a product of the same clone [14]. The resultant mRNA from an *M-bcr* breakpoint is translated to a 210 kDa fusion protein (p210^{*Bcr-Abl*}). *M-bcr* breakpoints can also be detected in one-third of Ph positive acute lymphoblastic leukemia (ALL) in adults and a small proportion on Ph positive ALL in childhood. In rare cases of CML (~1%) and 50–80% of Ph positive ALL, the break occurs within a 55 kb intronic sequence between *BCR* exon 1 and exon 2 termed the minor breakpoint cluster region (*m-bcr*). This generates a hybrid transcript with a junction between exon 1 of *BCR* and exon a2 of *ABL1* (e1a2), producing a 190 kDa chimeric protein (p190^{*Bcr-Abl*}) [19]. A third rare breakpoint region is located within intron 19 of *BCR* (μ -*bcr*) and results in fusion of *BCR* exon 19 to *ABL1* exon 2 (e19a2). The derived transcript is translated to a 230 kDa protein (p230^{*Bcr-Abl*}) [20]. Even more rarely than the latter, a tiny minority of CML patients acquire fusion genes that juxtapose *BCR* exon 6 or *BCR* exon 8 to *ABL1* exon 2 (e6a2 or e8a2) (Fig. 2).

1.3 Testing for *BCR-ABL1*

The detection of *BCR-ABL1* is diagnostic of CML, therefore all patients suspected of CML or myeloproliferative disorders should be tested for this fusion gene by multiplex RT-PCR. Molecular analysis has greater sensitivity than cytogenetic analysis and can identify rare cryptic *BCR-ABL1*-causing chromosomal rearrangements that may otherwise go undetected. Following diagnosis, regular minimal residual disease (MRD) monitoring by qPCR assists in predicting patients at risk of relapse, following TKI therapy or BMT. *BCR-ABL1* gene transcripts can be detected and quantified in peripheral blood or bone marrow samples.

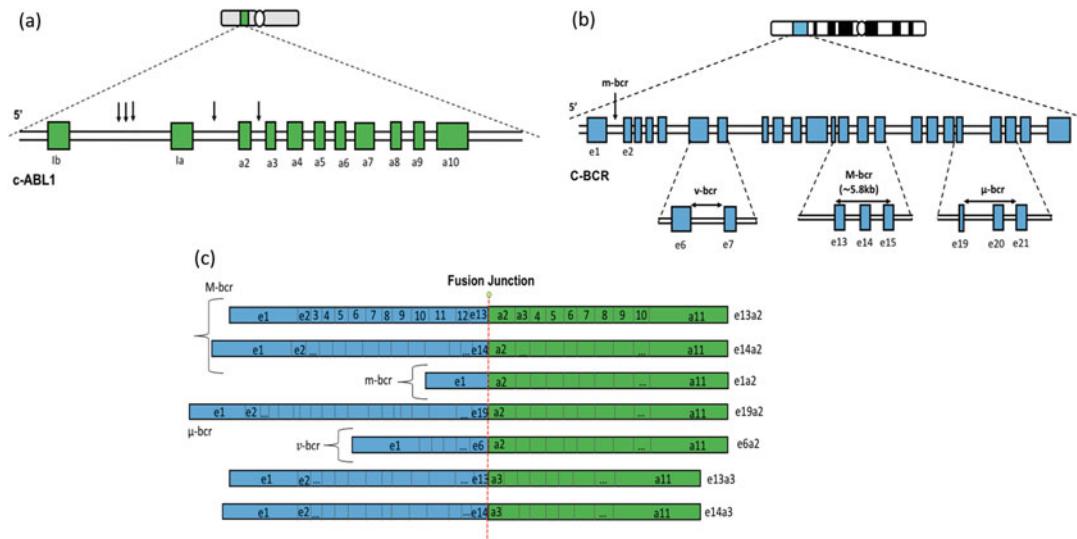


Fig. 2 Schematic representation of (a) *c-ABL1* gene containing 11 exons with breakpoints indicated (b) *c-BCR* comprising 23 exons with major (M-bcr) and rare (m-bcr, μ -bcr, and ν -bcr) breakpoints indicated. (c) *BCR-ABL1* fusion transcripts derived from M-bcr, m-bcr, μ -bcr, and ν -bcr

1.4 Multiplex PCR

It is important to know the exact *BCR-ABL1* breakpoint type prior to carrying out the qPCR reaction (which is specific to transcript type), so that accurate analysis and monitoring can subsequently be performed. This is determined by standard multiplexed PCR of cDNA from the initial patient sample before treatment, when *BCR-ABL1* levels are at their highest. In contrast to conventional PCR, multiplex PCR includes more than two primers, allowing the identification of all the most frequently described *BCR-ABL1* transcript types by the size of the PCR amplicon (Fig. 3). Multiplex assays are designed to detect the major breakpoint sub-types e13a2 (310 bp), e14a2 (385 bp) and e1a2 (481 bp). Rarer transcripts including e8a2 (1318 bp), e6a2 (1120 bp), e19a2 (925 bp) and e13/14a3 (~211 bp) can also be detected. Primers to amplify the *BCR* gene (808 bp) are included as a positive control for the reaction (Fig. 4) and samples where there is no amplification of the *BCR* gene are considered to have failed. Details for this assay are given in Subheading 3.3.

1.5 Real-Time Quantitative Polymerase Chain Reaction (qPCR)

Real-time quantitative PCR (qPCR) is a fluorescence-based amplification assay that allows for the rapid, sensitive, and specific detection and quantification of nucleic acid and is used routinely in the management of CML patients on therapy [21–23]. The detection of very low levels of nucleic acid is essential for the monitoring of *BCR-ABL1* due to the effectiveness of the treatments available in most patients, combined with a small but significant risk of relapse. Persistent low level or undetectable transcript may also assist in identifying patients who may benefit from therapy withdrawal.

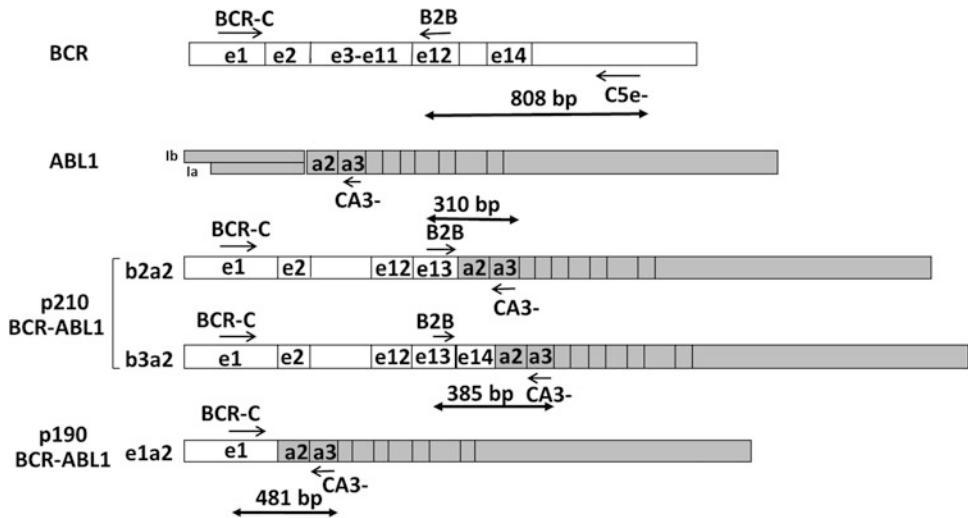


Fig. 3 Amplicons generated in the *BCR-ABL1* multiplex assay

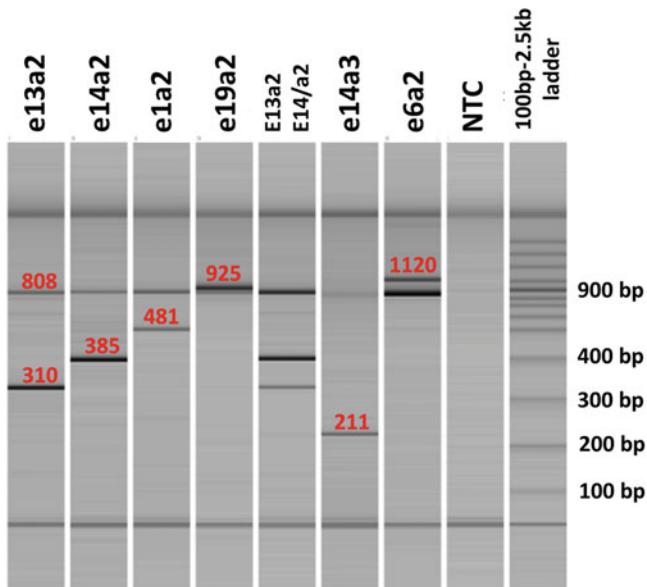


Fig. 4 Multiplex PCR fragment analysis control samples. Multiplex assays are designed to detect the major breakpoint subtypes—e13a2 (310 bp), e14a2 (385 bp), and e1a2 (481 bp). Rarer transcripts including e8a2 (1318 bp), e6a2 (1120 bp), e19a2 (925 bp,) and e13/14a3 (~211 bp) can also be detected. Primers to amplify the *BCR* gene (808 bp) are included as a positive control for the reaction and samples where there is no amplification of the *BCR* gene are considered to have failed

qPCR simultaneously detects target template and quantifies it by monitoring the increase in fluorescent signal associated with product formation during each PCR cycle. The accumulating amplicons are detected by including a sequence-specific probe labelled with fluorochromes in addition to the PCR primers. Monitoring is performed during the exponential phase of PCR amplification, when the number of amplicons detected is directly proportional to the initial number of targets present. Real-time detection of accumulating amplicons eliminates post PCR manipulation and the risk of contamination.

Two reporter systems are commonly used for qPCR analysis; the SYBR Green system [21] and the hydrolysis probe (TaqMan™) system [22, 23]. SYBR Green binds nonspecifically to any double-stranded DNA and once bound, will emit fluorescence. During the PCR reaction, as more double stranded product is produced, there is a proportional increase in fluorescent signal. These assays must be carefully optimized with highly specific primers to avoid nonspecific amplification products which would also fluoresce. To ensure the specificity of the reaction, a melt curve analysis is performed at the end of the qPCR where the double stranded product is gradually heated and the change in fluorescence measured. As the product denatures, the SYBR Green becomes unbound and no longer emits fluorescence. The change in fluorescence is measured and plotted against temperature. There should be one unique peak to indicate the presence of only one PCR product. Different products will have different melting temperatures and will therefore produce their own temperature specific melt curve. This system is not specific enough for use in the *BCR-ABL1* diagnostic setting and is often used as a research tool when developing new assays.

TaqMan™ assays avoid the problem of non-specific fluorescence by using a fluorescently labelled probe that hybridizes to a region within the target amplicon sequence. The probe has a 5' fluorescent label and a 3' quencher molecule in close proximity, preventing fluorescence by virtue of fluorescent resonance energy transfer. At the PCR annealing stage, primers and probe bind to the cDNA. During the next stage of template extension, the Taq polymerase will cleave the 5' fluorophore from the probe, thus releasing it from the quencher and allowing it to fluoresce. Amplification is thereby measured by accumulation of the fluorophore during each extension stage, with the intensity of the fluorescence proportional to the number of amplicons, and the accumulation is related to the quantity of the target, that is, *BCR-ABL1*, in the starting sample.

The number of transcripts is calculated by comparison with standard curves generated using serial dilutions of *BCR-ABL1* cDNA at known concentrations, which are included in each run. Certified reference material is available for the *BCR-ABL1* test [24] which consists of plasmids at known concentrations that

contain both the *BCR-ABL1* and *ABL1* genes. The standards cover a copy number range from 10^1 to 10^6 copies per μl . qPCR data is also normalized against a reference gene (usually *ABL1*, *BCR* or *GUS*) to take into account the differing concentration of RNA in the initial clinical sample [25].

Multiple qPCR reactions can be performed simultaneously in the same well when the probes have different color fluorescent labels. This is an advantage of using TaqMan™ probes over the nonspecific SYBR Green system. A duplex qPCR has become the standard method of testing for the major *BCR-ABL1* transcript [26].

2 Materials

2.1 Sample Preparation, RNA Extraction, and cDNA Synthesis

There is a wide range of RNA extraction methods used across different labs. Commercial kits can be paired with automated systems such as the QIASymphony RNA Kit and QIASymphony from Qiagen, used in our laboratory.

- M-MLV Reverse Transcriptase (200 U/ μl) (Invitrogen).
- 5× reaction buffer (Invitrogen).
- Dithiothreitol (DTT) (0.1 M).
- dNTPs (25 mM).
- Random hexamers.
- RNase-free water.
- Heat block.
- Ice.
- 2 ml screw top tube.

- Multiplex PCR kit (we used QIAGEN kit).
- 0.2 ml PCR tubes.
- Control cDNA obtained by extracting RNA from the following cell lines and converting to cDNA:
 - HL-60 (as a negative control for *BCR-ABL1* amplification).
 - BV173 (as a positive control for e13a2).
 - K562 (as a positive control for e14a2).
 - SD-1 (as a positive control for e1a2).
 - AR230 (as a positive control for e19a2).
 - Cell lines can be purchased from DSMZ and ATCC cell line banks.
- RNase-free water.
- Agarose gel tank or automated system (e.g., QIAxcel).

2.3 qPCR

- Real-time quantitative PCR machine.(e.g., Applied Biosystems 7900HT Fast real-time PCR system).
- 96-well semiskirted PCR plates.
- Optically clear adhesive film.
- PCR Workstation.
- TaqMan fast advanced mastermix (Life Technologies).
- ERM-AD623 BCR-ABL pDNA CALIBRANT [24].
 - The 6 plasmid solutions are available from the European Commission’s Joint Research Centre (JRC).
 - <https://crm.jrc.ec.europa.eu/p/q/bcr-abl/ERM-AD623-BCR-ABL-pDNA-CALIBRANT/ERM-AD623>.
- Primers and probes

	Primers:	Probes:
BCR-ABL:		
e13a2/e14a2	ENF-501 and ENF-561	FAM-ENP541MGB
e1a2	ENF-402 and ENF-561	FAM-ENP541MGB
e19a2	E19F1 and ENF-561	FAM-ENP541MGB
e13a3/e14a3	ENF501F2 and ENR1063	FAM- ENP1043
e6a2	E6F2 and ENF-561	FAM-ENP541MGB
e8a2	E8F2 and ENF-561	FAM-ENP541MGB
ABL:		
	ENF1003 and ENR1063	VIC-ABL1043MGB

For sequences, see Table 1 for Major transcript and Table 2 for rare transcripts.

- HL-60 cDNA as Negative *BCR-ABL1* Control [27].
- K562 cDNA as positive *BCR-ABL1*control.
- High (10% *BCR-ABL1*) and low (0.1% *BCR-ABL1*) controls.

The high and low QCs are prepared from dilution of K562 cDNA in HL-60 cDNA. New stocks of cDNA are tested to determine the copy number of *ABL1* and *BCR-ABL1*. Cell line cDNA is then diluted to approximately 5×10^3 *ABL1* copy number and retested against the standard curve. With HL-60 and K562 at the same copy number, 1 part K562 and 9 parts HL-60 can be combined to make the high QC. The high QC can then be used to create the low QC by combining one part high QC with 99 parts HL-60 cDNA.

Table 1
Preparation of major transcript mastermix for duplex reaction. The primer/probe mix (770 µl) should be combined with 1100 µl TaqMan fast advanced Mastermix

Primer/probe	Sequence (5'-3')	Volume (µl)	Final concentration in 20 µl qPCR reaction
<i>BCR-ABL1</i>			
ENF-501 (80 µM stock)	TCCGCTGACCATCAAYAAGGA	8.25	300 nm
ENF-561 (80 µM stock)	CACTCAGACCCCTGAGGGCTCAA	8.25	300 nm
FAM-ENP541MGB ⁷ (100 µM stock)	FAM-CCCTTCAGCGGCCAGT-MGB	2.2	100 nm
<i>ABL1</i>			
ENF1003 (80 µM stock)	TGGAGATAACACTCTAACGATAACTAAAGGT	4.125	150 nm
ENR1063 (80 µM stock)	GATGTAGTTGCTTGGGGACCCA	4.125	150 nm
VIC-ABL1043MGB ⁷ (100 µM stock)	VIC-CATTTCAGGGTTGGGCTTC-MGB	4.4	200 nm
Water		738.65	
Total		770	

Primer sequences have been previously published [26]

Table 2
Primer–probe mixture reagents used for rare transcript singleplex PCR

Primer/Probe	Sequence (5'-3')	µl	Final concentration ^a
<i>BCR-ABL1e1a2 transcript</i>			
ENF-402 (80 µM stock)	CTGGCCCAACGATGGCGA	18	300 nM
ENF-561 (80 µM stock)	CACTCAGACCCTGAGGCTCAA	18	300 nM
FAM- ENP541MGB (100 µM stock)	FAM-CCCTTCAGCGGCCAGT-MGB	9.6	100 nM
TE buffer		194.4	
<i>BCR-ABL1e19a2 transcript</i>			
E19F1 (80 µM stock)	GGAGGAGGTGGGCATCTACCG	14.4	300 nM
ENF-561 (80 µM stock)	CACTCAGACCCTGAGGCTCAA	18	300 nM
FAM- ENP541MGB (100 µM stock)	FAM-CCCTTCAGCGGCCAGT-MGB	9.6	100 nM
TE buffer		198	
<i>BCR-ABL1e13a3/e14a3 transcript</i>			
ENF501F2 (80 µM stock)	TTCCGCTGACCATCAAYAAGG	14.4	300 nM
ENR1063 (80 µM stock)	GATGTAGTTGCTTGGGACCCA	14.4	300 nM
FAM- ENP1043 (100 µM stock)	FAM-CCATTGGTTGGGCTTCACACCATT -TAMRA	4.8	100 nM
TE buffer		206.4	
<i>BCR-ABL1e6a2 transcript</i>			
E6F2 (80 µM stock)	CAAAGATGCCAAGGATCCAACGACCAAG	14.4	300 nM
ENF-561 (80 µM stock)	CACTCAGACCCTGAGGCTCAA	18	300 nM
FAM- ENP541MGB (100 µM stock)	FAM-CCCTTCAGCGGCCAGT-MGB	9.6	100 nM
TE buffer		198	

(continued)

Table 2
(continued)

Primer/Probe	Sequence (5'-3')	µl	Final concentration ^a
<i>BCR-ABL1e8a2 transcript</i>			
E8F2 (80 µM stock)	ACGGCAGTCCATGACGGTGAAGAAG	14.4	300 nM
ENF-561 (80 µM stock)	CACTCAGACCCTGAGGCTCAA	18	300 nM
FAM- ENP541MGB (100 µM stock)	FAM-CCCTTCAGCGGCCAGT-MGB	9.6	100 nM
TE buffer		198	
Total for each primer mix		240	

^aIn 20 µl qPCR reaction

3 Methods

Appropriate personal protective equipment should be used during all lab work. Gloves must be worn when handling samples and reagents to avoid both contamination and the degradation of RNA by RNases naturally occurring on hands. Always use tips and tubes that have been tested and certified RNase-free.

3.1 Extraction of RNA

1. 10–20 ml whole blood or bone marrow is required in EDTA tubes, less than 72 h old and stored at room temperature. Local procedures for RNA extraction can be followed. There a number of systems and kits available for RNA extraction and the exact method will depend on local protocols and equipment available.
2. Red blood cells must be removed before an extraction of RNA can be carried out. White blood cells can then be lysed using a suitable method such as a Tissue disruptor or TissueLyser.
3. An automated system may be available for RNA extraction for example the QIASymphony system can be used with the QIA-Symphony RNA kit to obtain 55 µl of prepared RNA for downstream applications.

3.2 Generation of cDNA

The initial stage of qPCR analysis involves converting the mRNA to a single-stranded complementary DNA (cDNA). This is achieved by the reverse transcription enzyme murine leukemia virus reverse transcriptase (M-MLV RT) and primed with random hexamers (see

Table 3
Reagents for cDNA synthesis mastermix

Reagent	Volume per reaction (μl)
5× reaction buffer	20
Dithiothreitol (DTT) (0.1 M)	10
dNTPs (25 mM)	
dATP	0.5
dCTP	0.5
dGTP	0.5
dTTP	0.5
Random primers (3 $\mu\text{g}/\mu\text{l}$)	0.2
Nuclease-free water	9.2
RNasin (40 U/ μl)	1.2
M-MLV (200 U/ μl)	2.4
Total	45

Note 1) with the entire RNA sample converted to cDNA during the process (*see Note 2*).

1. Incubate 55 μl eluted RNA at 65 °C for 10 min. This is to ensure the RNA is linear and to denature any secondary structures. RNA is not routinely quantified in our laboratory as the samples are diagnostic and therefore processed regardless of RNA quantity.
2. During the incubation prepare a mastermix of reagents as listed in Table 3.

The mastermix can be prepared in advance (without RNasin or M-MLV) and stored at –20 °C until required. The mix can be defrosted and the RNasin and M-MLV added whilst the RNA is denaturing (**step 1**).

3. Remove samples from the heat block and place on ice for 30 s. This rapidly cools the RNA and prevents the formation of secondary structures. Pulse-spin the tubes to collect the sample at the bottom of the tube.
4. Add 45 μl of the mastermix to each 55 μl RNA sample. Mix by pipetting. Incubate at 37 °C for 2 h. The extended incubation achieves a better efficiency of synthesis.
5. Inactivate the reverse transcription enzyme by incubating the samples at 65 °C for 10 min. The cDNA is now ready for downstream applications.

Table 4
Preparation of multiplex PCR for 1, 24, and 96 samples

Reagent	Volume per reaction	Volume for 24 samples (including additional 10% for pipetting error)	Volume for 96 samples (including additional 10% for pipetting error)
Multiplex mix	25 µl	660	2640
10× primer mix ^a	5 µl	132	528
H ₂ O	18 µl	475.2	1900.8
Total volume	48 µl		

^aThe 10× primer mix is made by mixing 46 µl of H₂O and 1 µl working solution (10 µM) of each primer (BCR-C, B2B, C5e-, CA3-) for a total volume of 50 µl. Primer sequences and binding sites are shown in Fig. 3.

3.3 Multiplex PCR for Identification of BCR-ABL1 Transcript Type

Assays should be prepared in a designated pre-PCR room in a PCR hood which has been decontaminated with UV light for 15 min prior to starting the assay set-up. Multiplex PCR mix is ready-made by Qiagen (*see Note 3*).

The following controls must be included with each assay:

Non Template Control (NTC); BV173 cell line cDNA (a positive control for the e13a2 transcript); K562 cell line cDNA (a positive control for the e14a2 transcript); SD1 cell line cDNA (a positive control for the e1a2 transcript) and AR230 cell line cDNA (a positive control for the e19a2 transcript). Control cDNAs are prepared to an *ABL1* copy number of approximately 5×10^3 as measured by qPCR. Precise concentration is not essential for this qualitative assay.

1. Thaw the multiplex PCR Kit tube at room temperature. Prepare a 0.2 ml PCR strip tube for each sample and each control.
2. Prepare a mastermix as described in Table 4.
3. Aliquot 48 µl of the multiplex mix into each labelled 0.2 ml PCR tube. Add 2 µl of patient cDNA directly into the corresponding PCR tube.
4. In a separate room add 2 µl of each positive control cDNA to the corresponding PCR tube. Ensure PCR components are mixed well by pipetting. Tubes can be spun to collect contents to the bottom of the tube.
5. Transfer to a thermal cycler and run the following cycling parameters:

95 °C	15 min	
94 °C	30 s	35 cycles
63 °C	90 s	
72 °C	90 s	
72 °C	10 min	
4 °C	Hold	

- PCR products must be analyzed in a separated post-PCR room. Fragment size analysis can be performed on a standard agarose gel or on an automated system such as the QIAxcel. Band size will determine the breakpoint type for each sample and can be compared to the positive control (Fig. 4).

3.4 Duplex qPCR Test for BCR-ABL1

A duplex qPCR test (*see Note 4*) is performed to simultaneously detect the major *BCR-ABL1* transcript *e13a2/e14a2* and the *ABL1* reference gene. Assays should be prepared in a designated pre-PCR room in a PCR hood which has been decontaminated with UV light for 15 min prior to starting the assay set-up.

- Prepare a mastermix of the primer/probe as detailed in Tables 3 and 4, which is sufficient for 96 samples (one plate). A larger volume of mastermix can be made and 770 µl aliquots can be stored at –20 °C. The reference (*see Note 5*) (housekeeping) gene used in this example is *ABL1* (*see Note 6*). Just prior to performing the assay the 770 µl probe mix can be combined with 1100 µl TaqMan Fast Advanced Mastermix.
- For a 96 well plate, assign three wells for each of the six reference plasmid concentrations (10^1 – 10^6 copies per µl) (Standards 1–6). Include a no template control and a *BCR-ABL1* negative control such as cDNA from an HL-60 cell line [27]. For quality control monitoring known samples of low and high *BCR-ABL1* concentrations should also be run on the plate. Each sample should be assayed in triplicate. Record which samples are in which well using a grid such as the one below:

	1	2	3	4	5	6	7	8	9	10	11	12
A	Standard 1		HL-60		Sample 7		Sample 15					
B	Standard 2		NTC		Sample 8		Sample 16					
C	Standard 3		Sample 1		Sample 9		Sample 17					
D	Standard 4		Sample 2		Sample 10		Sample 18					
E	Standard 5		Sample 3		Sample 11		Sample 19					
F	Standard 6		Sample 4		Sample 12		Sample 20					

(continued)

G	High QC	Sample 5	Sample 13	Sample 21
H	Low QC	Sample 6	Sample 14	Sample 22

3. Pipette 17 µl mastermix into each well. Add 3 µl cDNA to each sample and HL-60 well. Cover with adhesive film and spin the contents to the bottom of the plate.
4. In a separate room, carefully uncover the plate and add 3 µl of each standard (to each of the three wells). Also add 3 µl of the high and low QC samples. Cover with optically clear adhesive film and spin the plate to collect the contents to the bottom of the tube.
5. The plate can be transferred onto the qPCR instrument and run with the following cycling parameters:

Stage	Temperature	Time (seconds)	Cycles
Hold	95 °C	20	1
Denature	95 °C	3	45
Annealing and extension	60 °C	45	

Fast settings should be used with a ramp rate of 100%, no auto increment, sample volume: 20 µl, reporter dye: FAM (*BCR-ABL1*) and VIC (*ABL1*), quencher: MGB (see Note 7), passive reference: ROX (see Note 8).

6. The number of cycles taken for the fluorescence to cross a threshold value of ten times the standard deviation of baseline emission is used for quantitative measurement. This value is termed the quantification cycle threshold (C_q) (see Note 9) and it is inversely proportional to the amount of starting material of the target. The numbers of transcripts of the target are read off the standard curve generated using serial dilutions of plasmid. The level of expression of the target gene is reported as percentage ratio compared to the reference gene to obtain a normalized value for the gene of interest independent of the integrity of the RNA and efficiency of the reverse transcription reaction.

3.5 Rare Transcripts

As mentioned in the introduction there are different types of *BCR-ABL1* fusion genes depending on where the break points occur in both the *BCR* and *ABL1* genes. To quantify the rare *BCR-ABL1* transcripts two separate reactions are run: first, a singleplex reaction for assessing the copy number of the specific *BCR-ABL1* transcript, using primers specific for the different transcript types; the second reaction is the duplex reaction as described above, from which only the value for the reference gene (*ABL1*) copy numbers is taken and used to calculate the ratio.

To minimize interassay variation and for convenience the Taq-Man primer–probe mixture is prepared in a ready to use aliquot of mix containing probe and primers at final concentrations shown in Tables 2 and 4.

1. Before starting the reaction calculate how much mixture is needed, depending on the number of reactions you are going to set up following the volumes below:

Reagent	Final volumes per well
Fast advanced mastermix	10 µl
Primer/probe mix	1 µl
dH ₂ O	6 µl
cDNA	3 µl
Total	20 µl

2. In a 96-well plate, assign three wells for each of the six reference plasmid concentrations. Include a no template control and a *BCR-ABL1* negative control such as cDNA from an HL-60 cell line [28]. Each patient’s sample should be assayed in triplicate. For quality control monitoring known samples of low and high *BCR-ABL1* concentrations should also be run on the plate. (To create the high and low controls we use a mixture of HL-60 cell line and SD-1 cell line for e1a2 and AR230 cell line for e19a2 transcripts. Record which sample belongs to which well.)
3. Pipette 17 µl mastermix into each well. Add 3 µl cDNA to each sample and HL-60 well. Cover with adhesive film and spin the contents to the bottom of the plate.
4. In a separate room, carefully uncover the plate and add 3 µl of each standard to each of the 3 wells. Also add 3 µl of the high and low QC samples. Cover with optically clear adhesive film and spin the plate to collect the contents to the bottom of the tube.
5. The plate can be transferred onto the real-time PCR instrument and run with the following cycling parameters:

Stage	Temperature	Time (s)	Cycles
Hold	95 °C	20	1
Denature	95 °C	1	45
Annealing and extension	60 °C	20	

6. When the run is finished, evaluate each standard curve and remove outliers (see Note 10). Proceed with quality assessment

Table 5
Parameters for quality assessment of the qPCR

Parameter	Acceptable range		
<i>Standard curve</i>			
Slope	–3.2 to –3.6		
R^2	≥0.98		
Efficiency	100% ± 10%		
<i>ABL1</i> intercept	≥38.5 to ≤40.2		
<i>BCR-ABL1</i> intercept	≥36.8 to ≤38.7		
<i>ABL1</i> plasmid 10 ²	≥31.8 to ≤33.40		
<i>ABL1</i> plasmid 10 ⁶	≥18.5 to ≤19.8		
<i>BCR-ABL1</i> plasmid 10 ²	≥30.8 to ≤31.8		
<i>BCR-ABL1</i> plasmid 10 ⁶	≥16.8 to ≤18.20		
<i>Samples</i>			
<i>ABL1</i> copy number	≥10 ⁴		
Difference between highest and lowest C_q	$C_q \leq 30$ $C_q 31–33$ $C_q 34–37$	$\leq 0.5 C_q$ $\leq 1.0 C_q$ $\leq 1.5 C_q$	

3.6 Quality Assessment

Each qPCR reaction is subject of quality assessment before accepting and reporting the results. The following parameters must be achieved for the qPCR to pass quality control assessment for the run. One of the triplicates for each sample can be removed if needed to allow these criteria to be met (Table 5). All parameters from this table are recorded and monitored according to the Westgard rules (see Note 11).

A monthly quality record is used to compare all QC parameters for the last 6 months. Thus, any slight changes in the performance of the reaction are detected and eliminated to ensure the stability of the test performance in time.

As mentioned previously, each run should include also the following controls:

Non template control (NTC): this control should have 0 values for *BCR-ABL1* and for *ABL1* tests. If it shows positive values it should be assessed by the lead scientist. If the C_q value is ≥40 cycles, the result is accepted, if the C_q value is ≤39.5, the run is considered failed due to contamination.

BCR-ABL1 negative control: HL-60 cell line is usually used as a negative control for the *BCR-ABL1* oncogene. The quantified *ABL1* copy number for this negative control by the Duplex assay

Table 6
Interpretation of control data

NTC (-) NAC (-)	Pass
NTC (-) NAC (+)	Contamination of the cDNA synthesis step. Consistent positive NAC for a newly designed assay after exclusion of the contamination suggests the assay is not optimized to specifically amplify RNA and there is a requirement for redesigning the assay
NTC (+) NAC (+)	Contamination of the duplex mastermix

should be $\geq 10,000$ copies and it should be negative for *BCR-ABL1* transcript. Any positive *BCR-ABL1* value indicates contamination. In the case of contamination, the run should be repeated using a new vial of HL-60 cDNA if the investigation suggests the source of contamination is HL-60 cDNA (such as observing negative *BCR-ABL1* for any other sample in the same run).

Controls produced from the mixes used during the cDNA synthesis are also run on the duplex reaction and should show 0 for both *BCR-ABL1* and *ABL1* transcripts. The no amplification control (NAC) is used to assess the contamination of external cDNA during the cDNA synthesis. This control has all the components of the cDNA synthesis except RNA. The control for the duplex mix is the reaction without cDNA. The application of NAC is important to show the assay specifically amplified the mRNA product without amplifying any DNA template and also shows the cDNA synthesis process was not contaminated with external *BCR-ABL1* cDNA. The no template control (NTC) is for assessment of the duplex mix contamination. Positive NTC suggests contamination of the duplex mastermix. Interpretation is summarized in Table 6.

Each run also contains internal quality control (ICQ) assessment material: QC *High* (ratio 10%)/*Low* (ratio 0.1%) samples. The values for these are also monitored on a run based and assessed against Westgard rules. The high and load QC are prepared from dilution of K562 cells in HL-60 cells.

4 Notes

1. OligodTs could be used to ensure that only mRNA is reverse transcribed and not the more abundant ribosomal RNA (rRNA); however, this can give a bias toward the 3' end of a PCR product, particularly in samples where there is some degradation of the RNA.

2. The maximum recommended amount of RNA in a reverse transcription reaction is 2 µg. The RNA is not routinely quantified and the input amount can be normalized by means of the *ABL1* reference gene.
3. The QIAGEN Multiplex PCR Master Mix includes Hot-Start Taq DNA Polymerase and a unique PCR buffer containing the novel synthetic Factor MP. Together with optimized salt concentrations, this additive stabilizes specifically bound primers and enables efficient extension of all primers in the reaction without the need for optimization. Q-Solution, a novel additive that enables efficient amplification of “difficult” (e.g., GC-rich) templates, is also supplied.
4. There can be difficulties in performing multiplexed qPCR assays as opposed to assaying each transcript individually. Different targets are using the same components in the reaction and have to compete for resources, potentially resulting in detection of the more abundant templates while less abundant ones fail to be detected. However, once fully optimized and validated, the system can dramatically reduce time and reagent costs for reactions.
5. The term housekeeping has historically been used but the accepted term is now reference gene. The MIQE guidelines: Minimum Information for publication of Quantitative Real-Time PCR Experiments [28] provide useful information on what is required in a qPCR experiment and clearly define the terms that should be used in the context of qPCR.
6. Reference genes used by different laboratories vary but, in this protocol, *ABL1* is used. Accepted alternatives include *BCR* and *GUS*.
7. To enhance the TaqMan™ system, minor groove binder (MGB) probes have been developed [29] which stabilize the probe, allowing shorter sequences of 13–20 bases to be used and reducing the overall size of the amplicon required. The MGB group is chemically attached to the 3' end of the TaqMan probe and gives better quenching of the reporter. The probes can be shorter than conventional TaqMan probes; therefore, they offer a greater flexibility for probe design.
8. Passive reference dyes are usually used in qPCR reactions to normalize for non-PCR related fluorescence signal difference. This normalizes for well-to-well optical variations on the instrument. Since the passive reference does not participate in the PCR reaction, the passive reference dye signal is stable during the PCR reaction. This gives a reference internal to the reaction to which the reporter dye signal (typically denoted as R, for Reporter) can be normalized.

Normalization usually involves dividing the emission intensity of the reporter dye by the emission intensity of the passive reference (P), to gain a ratio denoted as R_n (normalized reporter) for a particular reaction site: $R_n = (R/P)$. An R_n value is calculated for every cycle and every reporter and is normally plotted as an available view of the qPCR amplification data.

9. The terms threshold cycle (C_t), crossing point (C_p), and take-off point (TOP) have all been used in the literature but quantification cycle (C_q) is recommended in the MIQE guidelines [28].
10. An outlier is a replicate value which differs from the other values by $\geq 0.5 C_q$. For example, in a sample with C_q values of 26.1, 26.2, and 27, the 27 would be removed because it is $> 0.5 C_q$ different from the other values. For C_q values ≥ 30 , a higher difference in C_q values is acceptable (Table 5).
11. Westgard Rules are multirule QC rules to help to determine whether or not an analytical run is in-control or out-of-control. It uses a combination of decision criteria, usually including five different control rules, to judge the acceptability of an analytical run [30].

References

1. Nowell PC, Hungerford DA (1960) Chromosome studies on normal and leukemic human leukocytes. *J Natl Cancer Inst* 25:85–109
2. Rowley JD (1973) Chromosomal patterns in myelocytic leukemia. *N Engl J Med* 289:220–221. <https://doi.org/10.1056/NEJM197307262890421>
3. Khorashad JS, Deininger MW (2011) Selection of therapy: rational decisions based on molecular events. *Hematol Oncol Clin North Am* 25:1009–1023., vi. <https://doi.org/10.1016/j.hoc.2011.09.006>
4. Kantarjian HM, Keating MJ, Talpaz M et al (1987) Chronic myelogenous leukemia in blast crisis. Analysis of 242 patients. *Am J Med* 83:445–454
5. Huang X, Cortes J, Kantarjian H (2012) Estimations of the increasing prevalence and plateau prevalence of chronic myeloid leukemia in the era of tyrosine kinase inhibitor therapy. *Cancer* 118:3123–3127. <https://doi.org/10.1002/cncr.26679>
6. Druker BJ, Guilhot F, O'Brien SG et al (2006) Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *N Engl J Med* 355:2408–2417. <https://doi.org/10.1056/NEJMoa062867>
7. Eissa H, Gooley TA, Sorror ML et al (2011) Allogeneic hematopoietic cell transplantation for chronic myelomonocytic leukemia: relapse-free survival is determined by karyotype and comorbidities. *Biol Blood Marrow Transplant* 17:908–915. <https://doi.org/10.1016/j.bbmt.2010.09.018>
8. Melo JV, Gordon DE, Cross NC et al (1993) The ABL-BCR fusion gene is expressed in chronic myeloid leukemia. *Blood* 81:158–165
9. Cross NC, Melo JV, Feng L et al (1994) An optimized multiplex polymerase chain reaction (PCR) for detection of BCR-ABL fusion mRNAs in haematological disorders. *Leukemia* 8:186–189
10. Baccarani M, Deininger MW, Rosti G et al (2013) European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. *Blood* 122:872–884. <https://doi.org/10.1182/blood-2013-05-501569>
11. Cross NC, White HE, Colomer D et al (2015) Laboratory recommendations for scoring deep molecular responses following treatment for chronic myeloid leukemia. *Leukemia*

- 29:999–1003. <https://doi.org/10.1038/leu.2015.29>
12. Cross NC, Hochhaus A, Muller MC (2015) Molecular monitoring of chronic myeloid leukemia: principles and interlaboratory standardization. *Ann Hematol* 94(Suppl 2): S219–S225. <https://doi.org/10.1007/s00277-015-2315-1>
 13. Deininger MW, Goldman JM, Melo JV (2000) The molecular biology of chronic myeloid leukemia. *Blood* 96:3343–3356
 14. Roman J, Jimenez A, Barrios M et al (2001) E1A3 as a unique, naturally occurring BCR-ABL transcript in an indolent case of chronic myeloid leukaemia. *Br J Haematol* 114:635–637
 15. Jinawath N, Norris-Kirby A, Smith BD et al (2009) A rare e14a3 (b3a3) BCR-ABL fusion transcript in chronic myeloid leukemia: diagnostic challenges in clinical laboratory practice. *J Mol Diagn* 11:359–363. <https://doi.org/10.2353/jmoldx.2009.090008>
 16. Laurent E, Talpaz M, Kantarjian H et al (2001) The BCR gene and Philadelphia chromosome-positive leukemogenesis. *Cancer Res* 61:2343–2355
 17. Claudio S, Apperley JF, Gale RP et al (2017) E14a2 BCR-ABL1 transcript is associated with a higher rate of treatment-free remission in individuals with chronic myeloid leukemia after stopping tyrosine kinase inhibitor therapy. *Haematologica* 102:e297–e299. <https://doi.org/10.3324/haematol.2017.168740>
 18. Hanfstein B, Lauseker M, Hehlmann R et al (2014) Distinct characteristics of e13a2 versus e14a2 BCR-ABL1 driven chronic myeloid leukemia under first-line therapy with imatinib. *Haematologica* 99:1441–1447. <https://doi.org/10.3324/haematol.2013.096537>
 19. Melo JV (1996) The diversity of BCR-ABL fusion proteins and their relationship to leukemia phenotype. *Blood* 88:2375–2384
 20. Li S, Ilaria RL Jr, Million RP et al (1999) The P190, P210, and P230 forms of the BCR/ABL oncogene induce a similar chronic myeloid leukemia-like syndrome in mice but have different lymphoid leukemogenic activity. *J Exp Med* 189:1399–1412
 21. Wittwer CT, Herrmann MG, Moss AA et al (1997) Continuous fluorescence monitoring of rapid cycle DNA amplification. *BioTechniques* 22:130–131. 134–138
 22. Holland PM, Abramson RD, Watson R et al (1991) Detection of specific polymerase chain reaction product by utilizing the 5'-3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc Natl Acad Sci U S A* 88:7276–7280
 23. Livak KJ, Flood SJ, Marmaro J et al (1995) Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Methods Appl* 4:357–362
 24. White H, Deprez L, Corbisier P et al (2015) A certified plasmid reference material for the standardisation of BCR-ABL1 mRNA quantification by real-time quantitative PCR. *Leukemia* 29:369–376. <https://doi.org/10.1038/leu.2014.217>
 25. Wang YL, Lee JW, Ceserman E et al (2006) Molecular monitoring of chronic myelogenous leukemia: identification of the most suitable internal control gene for real-time quantification of BCR-ABL transcripts. *J Mol Diagn* 8:231–239. <https://doi.org/10.2353/jmoldx.2006.040404>
 26. Gerrard G, Mudge K, Foskett P et al (2012) Fast-mode duplex qPCR for BCR-ABL1 molecular monitoring: innovation, automation, and harmonization. *Am J Hematol* 87:717–720. <https://doi.org/10.1002/ajh.23212>
 27. Collins SJ (1987) The HL-60 promyelocytic leukemia cell line: proliferation, differentiation, and cellular oncogene expression. *Blood* 70:1233–1244
 28. Bustin SA, Benes V, Garson JA et al (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55:611–622. <https://doi.org/10.1373/clinchem.2008.112797>
 29. Kutyavin IV, Afonina IA, Mills A et al (2000) 3'-minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures. *Nucleic Acids Res* 28:655–661
 30. Westgard JO, Barry PL, Hunt MR et al (1981) A multi-rule Shewhart chart for quality control in clinical chemistry. *Clin Chem* 27:493–501



Chapter 13

Normalization in Human Glioma Tissue

**Ana Paula Santin Bertoni, Isabele Cristiana Iser,
Rafael Paschoal de Campos, and Márcia Rosangela Wink**

Abstract

For tissues obtained from glioma samples with/without nonneoplastic brain there is no consensus for universal reference gene but there are some potential genes that might have good stability, under certain conditions. Considering all points described in this work, the care with tissue collection, until gene amplification, directly impacts on the reliable characterization of its mRNA levels. Moreover, it is clear the importance of selecting the most appropriate reference genes for each experimental situation, to allow the accurate normalization of target genes, especially for genes that are subtly regulated.

Key words Reference gene, Quantitative PCR, Glioma, Preanalytical handling

1 Introduction

Gliomas are the most frequent tumors that affect the central nervous system. These tumors are classified in accordance with three types (astrocytomas, oligodendrogiomas, and mixed oligoastrocytomas) and four grades [1, 2]. Malignant gliomas (grades III and IV) are the most common tumors among the malignant primary brain tumors, corresponding to 70% of new cases diagnosed in adults in the USA each year. Glioblastomas (GBM) are grade IV gliomas which represent 60–70% of malignant gliomas and are characterized by high morbidity and mortality rates. The prognosis of patients diagnosed with malignant gliomas is very poor, especially for those with GBM whose average survival is of only 12–15 months [3]. The treatment of gliomas is largely dependent on their histological classification but usually comprise surgical resection, radiation, and/or chemotherapy with alkylating agents such as temozolomide. However, despite the intense efforts of scientists to improve treatment efficiency, little progress has been made in the last years in terms of patient survival. Poor prognosis has been attributed, among many other factors, to the acquisition of resistance to chemotherapy, resulting in the failure of conventional therapies [4].

During tumorigenesis, the steadily regulation of gene expression, even by up and down regulation, leads tumor cells to develop advantageous abilities over their healthy neighbors. They change the microenvironment via modifications in their secretome and other soluble mediators, which facilitate tumor progression and communication with resident and recruited cells (e.g., microglia, stromal and/or immune cells). In addition, these changes in gene expression can lead to activation of several metabolic pathways and posttranslational modification of proteins, which can culminate in resistance to treatment.

To work around this issue, researchers are investigating prognostic biomarkers to predict the outcome and response to treatment to each individual patient, providing a personalized therapy. The knowledge of the individual molecular profile has value in the clinical practice to guide patient stratification for personalized therapy, as well as adding predictive value on diagnostics or controlling clinicopathological variables of each patient [5]. Also, the molecule profile reflects the status of the intracellular processes. Therefore, showing how cells may respond to drugs, which is important as a therapeutic monitoring tool. Currently, there is no molecular panel regularly incorporated into clinical procedures for glioma. But the identification of clinically relevant subtypes of glioblastoma by Verhaak et al. using data from The Cancer Genome Atlas Research (TCGA) has brought valuable predictors that may eventually be incorporated into clinical practice. Through an integrated genomic analysis, this study was able to classify tumors in four main subtypes (proneural, neural, classical and mesenchymal) that present different survival outcomes and therapy responses [6], highlighting the importance of molecular studies.

The molecular characterization of each individual patient begins by the isolation of ribonucleic acid (RNA) from tumor cells or healthy brain tissues. The RNA is a single-stranded copy of a gene which is mostly composed of coding sequences. After isolation, RNA is transcribed into complementary DNA (cDNA) by a reverse transcriptase. Its quantification is measured by a reverse transcriptase-coupled quantitative polymerase chain reaction (RT-qPCR), which is reproducible, specific and sensitive enough to detect very low levels of RNA, even considering a single cell. For the quantification of the expression levels of target genes it is necessary the use of appropriate reference genes (also called housekeeping genes). They normalize pipetting errors among samples and/or correct the intra-assay and interassay variance in the efficiency of reverse transcription and PCR itself.

Currently, it is well-known that an adequate reference gene should maintain stable expression under different experimental conditions. Thus, for each experimental parameter, validation should be performed to find the best reference gene. For example, the adequate reference gene for a certain glioma type may not be

suitable to normalize samples of GBM and also, within GBM subtypes the adequate reference gene can vary. This variability can also be observed in the same brain tissue, depending on its origin or on its preservation, for example, whether a biopsy is frozen in dry ice or fixed in formalin. The use of inappropriate gene for normalization can generate misinterpretations of data and erroneous results. Variability in biological factors (age, sex, lifestyle, BMI, previous disease, or treatment) may affect the stability of reference gene and for that, these factors can be taken into consideration when comparing the candidates to reference gene.

Normalization using adequate reference gene is even more critical for genes that are slightly regulated in tumorigenesis. This “slightly” up or down regulation on gene expression does not mean that these genes will have low influence in the multistage process of malignant cellular transformation. On the other hand, genes that are up or downregulated over two times may not be affected, due to the difference in magnitude of expression between these target genes and the reference gene.

Undoubtedly, the set of information obtained from samples of the tumor region, comprising the surrounding margins and non-tumor (normal) adjacent tissue are essential to understand the process of tumor progression, as well as to correct and reduce the individual background. Clearly, this is the best scenario to obtain the molecular pattern of target genes, even considering that adjacent “normal” tissue surrounding tumor can present alterations known as “pretumor stage alterations.” Histopathological diagnosis of glioma is dependent on the tissue obtained from resection surgery or small biopsy and/or autopsy specimens, which then, become sources to evaluate the gene expression profile of gliomas. When molecular analysis is conducted using autopsy samples, all conditions involved (such as hypoxia, coma or seizures) and post-mortem delay until tissue collection should be documented and considered in the interpretation of data in molecular studies.

Preanalytical handling in sample collection

Preanalytical handling is the process that occur during collection, preservation and storage of the tissue sample. Preanalytical variables are factors in handling that can affect the sample stability and integrity, which can lead to misinterpretation of gene expression data, inaccurate results, and/or systematic biases [7].

Fresh-frozen samples from surgical resections or small biopsy represent the most accurate pattern to reveal the actual genomic status. However, fresh brain tissue is very difficult to obtain, since tumor tissue extracted in the surgery is scarce and necrotic. Besides, the priority is to get the pathological diagnosis or immunohistochemical analysis of tumors. Although formalin-fixed, paraffin-embedded (FFPE) is a standard method for pathology analysis, this method is not the most ideal for gene expression analysis. Formalin, a widely used fixative, may induce base alteration,

chemical rearrangement and modifications in the structure of nucleic acids, generating highly degraded RNA, thereby impairing gene expression analysis. In addition, factors such as the duration and time until fixation, fixative type, storage conditions, and others may contribute to oscillations in gene expression levels by inducing RNA degradation [8].

An alternative to overcome these problems would be to replace formalin without her fixatives (such as Bouin's solution, Carnoy's fixative, acetone, or alcohol) to improve the quantity and quality of FFPE samples for molecular studies. Moreover, it may be required to do specific adaptations of standardized protocols to obtain better results. Despite these barriers, RNA derived from FFPE tissue is interesting to use in research. Mainly because (1) nucleic acids can be obtained from years-old FFPE samples; (2) large files of FFPE can be easily generated and accessed; and (3) it allows an accurate separation of areas of interest by macro or microdissection [9].

In spite of attempts to improve the quality of RNA obtained from FFPE tissue samples, the low quality of RNA originated makes it even more important to choose a proper reference gene. This will avoid misleading conclusions and interpretation of data with possible undesirable clinical consequences.

To preserve and provide the faithfully patterns of gene profile on tumorous or healthy brain tissues, the preservation must occur at the time of resection surgery or biopsy collection. The tissues should be fast handling. The RNA present in specimens is preserved by freezing them at cold temperatures or using chemical fixatives and other solutions that potentially inactivate RNases. Before preservation, the brain tissue should be quickly cleaned with commercially available solutions containing RNase inhibitors to remove blood and connective tissues. The preservation of tissue on liquids that do not have any RNA stabilization property, such as phosphate buffered saline (PBS) or other saline solutions, could affected the levels and integrity of RNA.

Disregarding biases on preservation of brain tissue, no significant detrimental effect was found in samples from fresh tissue kept stored at -80°C , neither in RNA integrity nor quantitative yield, even for prolonged times (1–17 years) [10, 11]. FFPE tissues are stable at room temperature and suitable for long-term storage, without influence in RNA integrity or quantitative yield [12], which is more influenced by extraction method. To maintain RNA integrity, tumor tissues should be fractionated and frozen in different tubes. Data indicate that RNA extracted from tumor tissues that are submitted to freeze-thaw cycles more than three times are unsuitable for qPCR amplification due to the high rate of RNA degradation [13]. Because of the many RNA alterations caused by FFPE process, researchers are joining efforts to optimize RNA extraction and gene expression analysis protocols when

working with these samples. Both, quality and quantity of RNA are dependent on the method of RNA extraction.

Various methods for the isolation of nucleic acids from FFPE samples are commercially available and are widely used. However, a number of limitations may be inherent to all of them, including total processing time, involvement of toxic chemicals, such as xylene for manual deparaffinization, and high tissue consumption.

It is important that researchers observe the different variables involved in the RNA extraction process and evaluate different factors that might interfere in the amplification of genes from the samples.

mRNAs are present in low abundance due to an equilibrium between rates of synthesis or decay of mRNA (process named transcription). This is necessary to carry on the fast and dynamic cellular changes to both, intrinsic and extrinsic stimuli, otherwise, the cell homeostasis can be compromised. The main intrinsic factors are them RNA sequence itself, which predicts its lifetime, and the concentration that is determinant and considered inversely proportional to its stability [14]. As extrinsic factors, the presence of microRNAs (MiRs) can affect mRNA levels through a decrease in mRNA abundance. In the brain tissue, mRNAs with 3'-untranslated region (UTR) binding sites to miR-124, miR-29, miR-9, and miR-128 have been associated with reduced mRNA stability, whereas the presence of RNA-binding sites to certain protein families, such as RBFOX and ZFP36, is associated with increased mRNA stability [15].

In the reverse-transcription reaction of RNA to complementary DNA (cDNA), total RNA or messenger RNA (mRNA) can be selected according to reverse transcription (RT)-priming conditions. The most common approaches use either random primers, which synthesize cDNA from total RNA, or oligo-dT which binds to poly-A tails and therefore, majorly, targets eukaryotic mRNA [16]. Both conditions have advantages and disadvantages. Thus, empirical data is required to choose the most suitable option for each experiment. Oligo-dT priming usually results in higher product yields due to its specificity for polyadenylated RNA. For this reason, it may reduce the threshold cycle for detection, which is particularly interesting for less abundant transcripts [17]. However, it is important to consider that oligo-dT priming might overrepresent the 3' end of transcripts, depending on the extension capabilities of each reverse transcriptase enzyme [18]. In consequence, the distance between the poly-A tail and PCR primers annealing site needs to be carefully considered when designing primers. Also, oligo-dT priming is not recommendable for FFPE samples as a result of RNA fragmentation and dislocation of poly-A tails [19]. On the other hand, random hexamer primers bind to several regions of RNA molecules. Thus, they are less susceptible to misrepresenting transcripts due to the distance of PCR annealing sites

and poly-A tails, target RNA length and integrity [18, 20]. Nonetheless, it is important to consider that ribosomal RNA accounts for around 90% of total RNA. Hence, product yield of low expression genes may not be adequate for proper measurements, when using random primers [16, 20].

2 Reference Genes in Human Glioma Tissue

It is already a consensus that the selection of appropriate reference genes for clinical samples is essential to gene expression analysis. However, many studies make use of different reference genes without appropriate validation. Reports have identified that the “classic” genes used in normalization seem to be unsuitable due to the possible variation of stability on different biological samples, cell types, development stages, pathological/healthy conditions and under different experimental procedures [21]. It is important to keep in mind that for accurate quantification, it is crucial to choose a reference gene whose transcription level is similar as those of target genes. In addition, the normalization of qPCR data in GBM has mostly been performed with a single reference gene. However, the use of a combination of at least two validated reference genes may be necessary to prevent misinterpretation of gene expression data.

2.1 Candidates

Reference Gene for Glioma Samples: Where to Start from?

Considering the publications which have analyzed reference gene expression in human glioma tissue samples (Table 1), there is no consensus regarding the best gene or combination of reference genes that can be used to compare different glioma classifications or tumor samples, with adjacent nontumor brain tissue. Thereby, it is evident the importance of validation of adequate reference genes prior to every study that evaluates gene expression.

The selection of the panel to test candidate reference genes should consider their different functions on cell/tissue metabolism to avoid coregulation. Obviously, it is essential to exclude genes that are *upregulated* or *downregulated* as well as genes that show susceptibility to accumulate mutations. As example, B2M showed significant frequency of loss-of-function mutations in human glioblastoma samples and mouse model of GBM, which suggest a possible driver role in glioma genesis in both species [28]. Three studies using human gliomas samples (summarized on Table 1) [23, 24, 26], and another of our group, using rat glioma C6 cell line [29] corroborate with this data suggesting that B2M is inappropriate to use as reference gene in samples of human glioma tissues or in C6 cells.

To evaluate potential candidate genes for glioma samples, we analyzed the stability of 52 reference genes selected by Sharan et al. [30] on 613 gliomas samples from The Cancer Genome Atlas

Table 1
Literature review of publications about reference gene in human glioma tissue samples

Author; year	n; classification	Origin tissue; preservation; storage	RNA extraction method; primer type	Candidates	Reference gene selected
Röhn et al. (2018) [22]	10 astrocytoma WHO grade II; with CT _X 10 astrocytoma WHO grade II; without CT _X 10 astrocytoma WHO grade III; with CT _X 10 astrocytoma WHO grade II; without CT _X 10 GBM WHO grade IV; with CT _X 10 GBM WHO grade IV; without CT _X	Surgery; snap-frozen in liquid nitrogen; –80 °C	RNAeasy (Qiagen); QuantiTect® RT (Qiagen)	RPL13A, GAPDH, SDHA, POLR2A, ACTB, and TBP	SDHA and ACTB
Grube et al. (2015) [23]	3 diffuse astrocytoma 8 anaplastic astrocytoma 8 GBM 8 non-neoplastic brain tissue	ND; ND; –80 °C	Qiazol Reagent and the RNAseasy Mini Kit (Qiagen); ND	B2M, GAPDH, HMBS, RPL13A, CYCL1, TBP, Glioma × normal: GAPDH and RPL13A	
Aithal and Rajeswari (2015) [24]	10 GBM WHO grade IV 2 normal brain	Biopsy; FFPE; ND Autopsy; FFPE; ND	Trizol; random primers	HPRT, GAPDH, TBP, B2M, RPL13A, and RN18S1	TBP and RPL13A
Gresner et al. (2011) [25]	7 astrocytomas WHO grade II 10 anaplastic astrocytomas WHO grade III 15 glioblastoma WHO grade IV	ND	ND	GAPDH, HPRT1, POLR2A, RPLP0, ACTB, and H3F	GAPDH alone or RPLP0 and H3F combination
Kreth et al. (2010) [26]	9 astrocytoma WHO grade II 9 astrocytoma WHO grade III 9 GBM 9 non-neoplastic brain	Biopsy; RNA lysis buffer; ND	RNAqueous Micro Kit (Ambion); Eberwine method	ACTB, ALAS, B2M, β-Globin, GAPDH, GAPDH, GUSB, HPRT1, TBP, IPO8, PBDG, PGKL, PPLA, RPL13A, and 18S rRNA	GAPDH, IPO8, RPL13A, SDHA, and TBP
Valente et al. (2009) [27]	30 GBM 9 non-neoplastic white matter	Surgery; snap-frozen in liquid nitrogen; microdissection	Trizol; oligo(dT)	ACTB, GAPDH, GUSB, HMBS, HPRT1, TBP, and 18S rRNA	TBP and HPRT1

CT_X chemotherapy, ND not described, GBM glioblastoma

(TCGA), based on their grade as described by Ceccarelli et al. [31]. The expression for genes RN18S1, H3F and PBGB was not available in TCGA dataset and CYC1 and YWHAZ were included based on the literature review described in Table 1.

The coefficient of variation (CV), defined as standard deviation normalized to expression mean (log-transformed, base-2), was used to analyze the variability of expression of each gene and are described on Table 2. The genes that presented CV below 25% are considered stable.

Table 2
Stability of candidates of reference gene for Grade II–III and Grade IV gliomas

Grade II (<i>n</i> = 216)		Grade III (<i>n</i> = 241)		Grade IV (<i>n</i> = 156)	
Gene	C.V. (%)	Gene	C.V. (%)	Gene	C.V. (%)
RAB5C	17.30	RAB5C	17.72	EIF2B1	17.20
EIF2B1	20.48	EIF2B1	18.76	CTBP1	20.44
CTBP1	20.83	MRPL19	20.34	MRPL19	20.55
GOLGA1	22.54	GOLGA1	20.52	RAB5C	21.24
TBP	22.79	CTBP1	22.41	ALAS1	21.66
MRPL19	22.82	ALAS1	23.70	TBP	23.50
PPIA	23.00	IPO8	24.15	PUM1	25.53
CASC3	24.37	TBP	24.96	CASC3	26.65
ESD	25.77	ESD	25.32	SDHA	27.78
HSP90AB1	27.02	PUM1	27.25	ACTB	28.56
PUM1	27.13	PPIA	28.00	IPO8	28.56
POP4	28.35	CASC3	28.32	SFRS4	30.56
PSMB6	28.38	HSP90AB1	28.43	PSMB6	31.27
ACTB	28.39	PSMB6	30.57	GOLGA1	32.03
SDHA	29.35	SDHA	30.79	ESD	32.06
G6PD	30.61	SFRS4	31.51	PES1	33.27
ALAS1	31.63	POP4	32.55	ELF1	34.22
PMM1	32.57	POLR2A	34.35	POP4	34.51
SFRS4	32.88	YWHAZ	34.91	POLR2A	34.59
POLR2A	33.55	ACTB	35.29	YWHAZ	35.96
YWHAZ	34.60	G6PD	36.44	G6PD	37.24
GAPDH	34.69	NDUFA1	37.03	HMBS	38.15
GUSB	34.75	PES1	37.19	PMM1	38.68

(continued)

Table 2
(continued)

Grade II (<i>n</i> = 216)		Grade III (<i>n</i> = 241)		Grade IV (<i>n</i> = 156)	
Gene	C.V. (%)	Gene	C.V. (%)	Gene	C.V. (%)
HMBS	35.82	PES1	38.11	UBC	38.73
CYC1	37.03	HMBS	38.36	PES1	42.33
NDUFA1	38.04	UBC	38.48	PPIA	42.53
PES1	38.15	PMM1	38.97	GAPDH	43.75
EEF1A1	38.88	EEF1A1	40.51	CYC1	45.07
UBC	39.11	ELF1	41.24	RPL19	45.58
ELF1	39.30	CYC1	42.98	EEF1A1	48.96
RPL19	39.87	RPL19	45.31	GUSB	49.47
PES1	40.34	GAPDH	48.41	RPL4	49.67
CETN2	41.81	GUSB	52.85	RPL4	49.67
IPO8	44.23	RPS13	54.12	NDUFA1	50.39
TMBIM4	50.13	RPL4	59.09	RPLP0	56.00
RPS13	50.36	RPL4	59.09	ADA	57.26
TPT1	51.68	TPT1	59.70	B2M	58.50
RPLP0	51.75	RPS18	60.55	HPRT1	58.54
RPL4	52.86	TUBA1B	61.73	TMBIM4	63.04
RPL4	52.86	B2M	65.06	YAP1	64.51
RPL37A	54.32	RPS14	65.64	RPS13	64.70
RPL29	57.69	RPL13A	65.96	RPL13A	65.74
RPL13A	58.50	RPLP0	66.31	TUBA1B	65.95
RPS14	59.59	RPL41	74.80	CETN2	66.57
TUBA1B	59.88	YAP1	75.06	RPL29	69.46
RPS18	64.72	CETN2	75.81	TPT1	70.20
B2M	64.82	TMBIM4	75.99	RPL41	72.79
RPL41	69.93	RPL29	78.41	RPL37A	72.96
YAP1	73.69	HPRT1	80.09	RPS18	73.80
ADA	75.95	ADA	81.04	RPS14	77.50
PRKG1	99.90	PRKG1	83.50	PRKG1	87.50
HPRT1	100.22	RPL37A	85.35	HSP90AB1	87.71

Based on inference described above, we suggest analyzing the stability of EIF2B1 (*Eukaryotic Translation Initiation Factor 2B Subunit Alpha*), CTBP1 (*C-Terminal Binding Protein 1*), and MRPL19 (*Mitochondrial Ribosomal Protein L9*) to normalize target genes on samples of different glioma grades or GBM molecular subtypes. Based on our experience, we also suggest evaluating the stability of the TBP (TATA box-Binding Protein) gene, since it has shown to be stable in several experimental conditions [29]. B2M [28] and RAB5C [32] genes are unsuitable due to their regulation on glioma tumors (see Note 1).

2.2 Conditions for Reference Gene Quantification on qPCR Experiments: Which Parameters Should Be Considered?

Briefly, real-time quantitative PCR (qPCR) allows to indirectly measure the amount of copies of cDNA/DNA synthesized by a thermostable DNA polymerase, during the repeating cycles of template denaturation, primer annealing and primer extension, through fluorescent dyes. There is a substantial number of different qPCR master mixes and instruments available. Thus, individual specificities and instructions should be followed to obtain optimal performance. However, regardless of qPCR instruments or master mixes chosen, there are valuable recommendations known as “*The MIQE Guidelines*” for proper standardization and reproducibility of the qPCR technique that should be followed [33] (see Note 2).

2.3 qPCR Master Mixes

To improve the reproducibility, sensibility, and specificity of qPCR, the use of preformulated qPCR master mixes is mightily recommended. It is mainly because the fluorescence emission of dyes is dependent on solution factors, such as pH and salt concentration. Also, minimal variations, certainly, provide different absolute C_t (cycle threshold) values. The best choices of preformulated qPCR master mixes are those that contain a Passive Reference Dye (e.g., ROX (*carboxy-X-rhodamine*) or CXR (*chest X-ray*)), which allow to compensate fluorescent fluctuations produced by variations in volume or concentration in each well. Excitation source and optics configurations may vary among different types of real-time PCR thermal cyclers. Therefore, the concentration of passive reference dye should be adjusted in each system.

2.4 Primer Design

Ideally, qPCR primers should be designed with the help of online available tools (e.g., PrimerBlast, Primer3, IDT Scitoools). They can be set to provide parameters that need to be considered, when selecting a target region to amplify, such as length, annealing temperature, specificity of transcript variants and other information (summarized recently by Green et al. [34]). In general, the best primer pairs are predicted at the top of the option list. After design, melting temperatures (T_m) and secondary structures as dimers and hairpins should be evaluated by tools such as OligoAnalyser. Regardless, primer specificity should be confirmed by the

visualization of a single peak melting curve and a single band of the expected molecular size, after gel electrophoresis of the qPCR amplicons.

Two parameters of primer design should be careful considered as they are strongly associated with amplification efficiency:

1. Amplicon size: shorter amplicons (less than 80 bp) are hardly distinguishable from primer dimers, whereas larger amplicons (more than 150 bp) result in rapid fluorescence saturation since they can bind more molecules of fluorescent dyes. When working with poor-quality RNA, design amplicons shorter than 100 bp;
2. %CG content: percentages of cytosine (C) and guanine (G) bases higher than 50% should be avoided, since they may prevent double-strand denaturation and therefore reduce amplification efficiency.

2.5 Amplification Efficiency (E)

Remarkably, inadequate primer design, accompanied by nonoptimal reaction conditions, affects primer-template annealing and is the main driver for low amplification efficiency. On the other hand, higher efficiencies ($>110\%$) reflect the contamination of samples with PCR inhibitors, such as phenol, ethanol, or proteins.

To calculate the efficiency of a qPCR method, a standard curve must be experimentally determined, for each primer pair. For that, it is necessary to prepare a serial dilution of standards and analyze each dilution point, in five to sixfold, under the same conditions that will be used to amplify test samples. Typically, the expression of reference and target genes in experimental studies show a wide dynamic range and for that, pooled samples can be used to construct the standards. Importantly, the dilution series should cover and expand the maximum and minimum range of C_t values by at least 20%. Also, the C_t value of dilution points and unknown samples should be between cycles 15 and 35. This is because the *software* usually automatically set the baseline (background signal from the earliest change in fluorescence signal) from cycles 3 to 15. Most often, after 35 cycles the reaction enters plateau and thus is not reliable. A series of tenfold and twofold dilutions represent a good choice for genes with high and low expression, respectively. To transfer liquid volumes across the dilution points, work with at least 5 μ l to minimize pipetting errors.

Importantly, the amplification efficiency should be analyzed for each run and adjusted for different experimental condition. If the number of molecules of amplicon double during each replication cycle, the amplification efficiency represents 100%, which is determined based on the slope of this graph, as $E = 10^{(-1/\text{slope})}$. It is acceptable to have efficiencies between 90% and 110% (slope -3.2 to -3.5) and also, the R^2 value, which reflects linearity of reaction, should be >0.98 , in order to predict with a good confidence the

quantification of unknown samples. The precision of a qPCR should be considered adequate, when the standard deviation of C_t values between replicates is ≤ 0.167 .

2.6 Template

The adequate concentration of cDNA should be determined empirically. It should yield a threshold cycle between the range of 20 and 30. In general, cDNA templates from 1 µg to 10 pg cover a range to amplify genes with very low to very high expression, respectively. Excess template can inhibit PCR by binding all the available primers. It is inappropriate to estimate the concentration of cDNA via spectrophotometry, because the final cDNA reaction also contains remnants of RNA, primers and unincorporated dNTPs, which will absorb at 260 nm. For this purpose, cDNA can be purified and quantified by fluorimetry using specific dyes that binds to double-stranded nucleic acids (cDNA).

3 Data Analysis

Relative quantification using a standard curve is the indicated method to analyze the stability and/or changes in mRNA levels of reference or/and target genes, when comparing different tissue samples (tumor vs. nontumor) or different experimental conditions (before vs. after treatment), or even independent samples from different patients (glioma samples). When comparing at least two groups, the comparative C_t method (also named as $\Delta\Delta C_t$), which reports the fold change of samples, in relation to a control, might also be used. However, it is important to note that if the stability of the reference gene is not adequate, it might hide the differences in gene expression between experimental groups, as C_t values from reference genes are first subtracted from C_t values of target genes (ΔC_t). Then, the difference between conditions is compared ($\Delta\Delta C_t$). It is important not to mistake the $\Delta\Delta C_t$ data analysis method, with the ΔC_t stability analysis method for reference genes, which is later described.

The standard curve described for assessing the amplification efficiency can be used to determine the amounts of cDNA in experimental samples. Based on the plotted curve, the software of the qPCR equipment, automatically generates a linear regression of the C_t values (y -axis), versus the log of each known concentration in the dilution series (x -axis). Then, based on this equation, it is possible to directly relate the expression of reference and/or target genes with their respective C_t values. Because of this direct correlation, all cautions written above, regarding standard curve preparation, and performance analysis (efficiency, slope, and R^2) are essential to avoid and correct possible biases. It is also important that all samples that will be compared have equal concentrations. The stability of the diluted samples and standards is severely

affected by successive freeze–thaw cycles and for this reason should be aliquoted, stored at -80°C and thawed only once before use. Indeed, accurate pipetting of samples and standards is required to provide reliable results.

3.1 Analysis of Reference Gene Stability

3.1.1 GeNorm

It calculates the pairwise variation (V) between all candidates across samples and then it creates a stability score (M) and cyclically removes the less stable candidate, until the most stable pair is obtained [35]. This tool uses relative expression data. M values lower than 1.0 reflect suitable reference genes. Since the M value is based on gene ratios, the two most stable reference genes for a subset of samples cannot be ranked and thus, GeNorm is only appropriated to evaluate the global stability across all samples analyzed.

3.1.2 NormFinder

NormFinder also uses pairwise correlations but considers intragroup and intergroup variability of the log-transformed expression [36]. A cutoff value of 0.15 is recommended for the pairwise variation. This tool is appropriated to assess whether stability is homogeneous in all groups analyzed (e.g., stability on different glioma grades or molecular subtypes, comparison between glioma vs. nontumor brain or female vs. male).

3.1.3 Comparative ΔC_t

This method evaluates the stability of candidate reference gene according to the standard deviation across samples of the ΔC_t values obtained from all possible pairs of candidate genes [37]. The pair of genes with the lowest standard deviation is considered the most stable.

3.1.4 BestKeeper

This algorithm ranks the most stable gene (maximum of 10 genes and 50 samples) according to an index (Pearson's correlation coefficient) composed by the values of C_t , fold-change, standard deviation, and coefficient of variation [38]. The candidates that show values of index closer to 1.0 are the most stable genes and genes with index up to 1.5 are unsuitable to be used as reference gene.

3.1.5 RefFinder

RefFinder is an user-friendly web-based interface which integrates the rankings results obtained from GeNorm, Normfinder, BestKeeper, and comparative ΔC_t method, creating a “recommended comprehensive ranking” for the candidate genes tested (<http://150.216.56.64/referencegene.php>; accessed on 12/22/18). This program does not allow to separate the samples into groups and raw C_t values must be input.

3.1.6 Coefficient of Variation (CV%)

This parameter is useful to evaluate stability across samples. It is defined by the standard deviation divided by the mean (can be expressed as a percentage) and a 25% variation is acceptable.

3.1.7 Maximum Fold Change (MFC)

The MFC is the ratio of the maximum and minimum C_t value observed for a given gene. For this, it is inappropriate to analyze samples with high heterogeneity due to outlier bias. Candidates with $MFC < 1.5$ are considered stable and $MFC < 2$ are considered unsuitable.

4 Notes

1. Since RT-qPCR represents the gold standard to validate gene expression changes from cDNA microarrays, the stability of these suggested genes as reference genes should be evaluated in any conditions tested. Also, the NCBI Unigene website allows to search for candidates of reference genes and can give information about levels of expression, possible isoforms, as well as its expression under tumor and/or healthy conditions.
2. In our lab, we validate and assess gene expression using master mixes with the SYBR Green intercalating dye, and for this, some precautions, described in this chapter, are crucial to ensure the validity and reliability of mRNA expression analyses.

References

1. Cohen AL, Colman H (2015) Glioma biology and molecular markers. *Cancer Treat Res* 163:15–30
2. Hanif F, Muzaffar K, Perveen K, Malhi SM, Simjee SU (2017) Glioblastoma multiforme: a review of its epidemiology and pathogenesis through clinical presentation and treatment. *Asian Pac J Cancer Prev* 18(1):3–9
3. Wen PY, Kesari S (2008) Malignant gliomas in adults. *N Engl J Med* 359(5):492–507
4. Qazi MA, Vora P, Venugopal C, Sidhu SS, Moffat J, Swanton C et al (2017) Intratumoral heterogeneity: pathways to treatment resistance and relapse in human glioblastoma. *Ann Oncol* 28(7):1448–1456
5. Jansen MP, Fockens JA, van Staveren IL, Dirkzwager-Kiel MM, Ritstier K, Looij MP et al (2005) Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling. *J Clin Oncol Off J Am Soc Clin Oncol* 23(4):732–740
6. Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD et al (2010) Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* 17(1):98–110
7. Ellervik C, Vaught J (2015) Preanalytical variables affecting the integrity of human biospecimens in biobanking. *Clin Chem* 61(7):914–934
8. Hennig G, Gehrmann M, Stropp U, Brauch H, Fritz P, Eichelbaum M et al (2010) Automated extraction of DNA and RNA from a single formalin-fixed paraffin-embedded tissue section for analysis of both single-nucleotide polymorphisms and mRNA expression. *Clin Chem* 56(12):1845–1853
9. Gouveia GR, Ferreira SC, Ferreira JE, Siqueira SA, Pereira J (2014) Comparison of two methods of RNA extraction from formalin-fixed paraffin-embedded tissue specimens. *Biomed Res Int* 2014:151724
10. Birdsill AC, Walker DG, Lue L, Sue LI, Beach TG (2011) Postmortem interval effect on RNA and gene expression in human brain tissue. *Cell Tissue Bank* 12(4):311–318

11. Choi S, Ray HE, Lai SH, Alwood JS, Globus RK (2016) Preservation of multiple mammalian tissues to maximize science return from ground based and spaceflight experiments. *PLoS One* 11(12):e0167391
12. Cronin M, Pho M, Dutta D, Stephans JC, Shak S, Kiefer MC et al (2004) Measurement of gene expression in archival paraffin-embedded tissues: development and performance of a 92-gene reverse transcriptase-polymerase chain reaction assay. *Am J Pathol* 164(1):35–42
13. Yu K, Xing J, Zhang J, Zhao R, Zhang Y, Zhao L (2017) Effect of multiple cycles of freeze-thawing on the RNA quality of lung cancer tissues. *Cell Tissue Bank* 18(3):433–440
14. Nouaille S, Mondeil S, Finoux AL, Moulis C, Girbal L, Cocaign-Bousquet M (2017) The stability of an mRNA is influenced by its concentration: a potential physical mechanism to regulate gene expression. *Nucleic Acids Res* 45(20):11711–11724
15. Alkallas R, Fish L, Goodarzi H, Najafabadi HS (2017) Inference of RNA decay rate from transcriptional profiling highlights the regulatory programs of Alzheimer's disease. *Nat Commun* 8(1):909
16. Weber CF, Kuske CR (2012) Comparative assessment of fungal cellobiohydrolase I richness and composition in cDNA generated using oligo(dT) primers or random hexamers. *J Microbiol Methods* 88(2):224–228
17. Lekanne Deprez RH, Fijnvandraat AC, Ruijter JM, Moorman AF (2002) Sensitivity and accuracy of quantitative real-time polymerase chain reaction using SYBR green I depends on cDNA synthesis conditions. *Anal Biochem* 307(1):63–69
18. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5(7):621–628
19. Zeka F, Vanderheyden K, De Smet E, Cuvelier CA, Mestdagh P, Vandesompele J (2016) Straightforward and sensitive RT-qPCR based gene expression analysis of FFPE samples. *Sci Rep* 6:21418
20. Stahlberg A, Hakansson J, Xian X, Semb H, Kubista M (2004) Properties of the reverse transcription reaction in mRNA quantification. *Clin Chem* 50(3):509–515
21. Huggett J, Dheda K, Bustin S, Zumla A (2005) Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun* 6(4):279–284
22. Rohn G, Koch A, Krischek B, Stavrinou P, Goldbrunner R, Timmer M (2018) ACTB and SDHA are suitable endogenous reference genes for gene expression studies in human astrocytomas using quantitative RT-PCR. *Technol Cancer Res Treat* 17:1533033818802318
23. Grube S, Gottig T, Freitag D, Ewald C, Kalff R, Walter J (2015) Selection of suitable reference genes for expression analysis in human glioma using RT-qPCR. *J Neuro-Oncol* 123(1):35–42
24. Aithal MG, Rajeswari N (2015) Validation of housekeeping genes for gene expression analysis in glioblastoma using quantitative real-time polymerase chain reaction. *Brain Tumor Res Treat* 3(1):24–29
25. Gresner SM, Golanska E, Kulczycka-Wojdala D, Jaskolski DJ, Papierz W, Liberski PP (2011) Selection of reference genes for gene expression studies in astrocytomas. *Anal Biochem* 408(1):163–165
26. Kreth S, Heyn J, Grau S, Kretzschmar HA, Egensperger R, Kreth FW (2010) Identification of valid endogenous control genes for determining gene expression in human glioma. *Neuro Oncol* 12(6):570–579
27. Valente V, Teixeira SA, Neder L, Okamoto OK, Oba-Shinjo SM, Marie SK, Scrideli CA, Paco-Larson ML, Carlotti CG Jr (2009) Selection of suitable housekeeping genes for expression analysis in glioblastoma using quantitative RT-PCR. *BMC Mol Biol* 10:17
28. Chow RD, Guzman CD, Wang G, Schmidt F, Youngblood MW, Ye L et al (2017) AAV-mediated direct *in vivo* CRISPR screen identifies functional suppressors in glioblastoma. *Nat Neurosci* 20(10):1329–1341
29. Iser IC, de Campos RP, Bertoni AP, Wink MR (2015) Identification of valid endogenous control genes for determining gene expression in C6 glioma cell line treated with conditioned medium from adipose-derived stem cell. *Biomed Pharmacother* 75:75–82
30. Sharan RN, Vaiphei ST, Nongrum S, Keppen J, Ksoo M (2015) Consensus reference gene(s) for gene expression studies in human cancers: end of the tunnel visible? *Cell Oncol* 38(6):419–431
31. Ceccarelli M, Barthel FP, Malta TM, Sabedot TS, Salama SR, Murray BA et al (2016) Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell* 164(3):550–563
32. Mallawaaratchy DM, Hallal S, Russell B, Ly L, Ebrahimkhani S, Wei H et al (2017) Comprehensive proteome profiling of glioblastoma-derived extracellular vesicles identifies markers

- for more aggressive disease. *J Neurooncol* 131 (2):233–244
33. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M et al (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55(4):611–622
34. Green MR, Sambrook J (2018) Optimizing primer and probe concentrations for use in real-time polymerase chain reaction (PCR) assays. *Cold Spring Harb Protoc* 2018(10): pdb.prot095018
35. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A et al (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3(7): RESEARCH0034
36. Andersen CL, Jensen JL, Orntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 64 (15):5245–5250
37. Silver N, Best S, Jiang J, Thein SL (2006) Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol Biol* 7:33
38. Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: Best-Keeper—excel-based tool using pair-wise correlations. *Biotechnol Lett* 26(6):509–515



Chapter 14

qPCR Applications for the Determination of the Biological Age

Mauro Castagnetta, Ulrich Pfeffer, Aldo Chiesa, Elena Gennaro, Massimiliano Cecconi, Domenico Coviello, and Nicoletta Sacchi

Abstract

Individual age is a phenotypic trait that provides useful information not only in forensic investigations but also in the aging research which is becoming an urgent call due to the dramatic growth of the aging population worldwide.

TaqMan quantification PCR (qPCR) can be successfully applied to biological age estimation, using method defined in Zubakov et al. (*Curr Biol* 20:R970-R971, 2010). Since levels of signal joint T-cell receptor rearrangement excision circle (sjTREC) in human lymphocytes are known to decrease with age increasing, the qPCR of sjTREC represents a simple and relatively reproducible technique which offers highly accurate age estimation results.

Key words qPCR, sjTREC, Lymphocytes, Aging

1 Introduction

Estimation of biological age of an unknown age individual can be useful in many fields, like legal affairs or humanitarian relief.

More recently the concept of biological age as a measure of human health in individuals with the same chronological age has gained popularity.

In keeping with the unprecedented growth rate of the world's aging population, there is a clear need for a better understanding of the biological aging process and the determinants of healthy aging. Toward this aim, a quest for markers that track the state of biological aging and ideally lend insights to the underlying mechanisms has been embarked upon.

Already existing methods often rely on identification of bone features or teeth; using a non invasive DNA-based molecular method can overcome subjective interpretation and, in our case, can be useful to correlate analytical data to age estimation.

T lymphocytes use specific receptors, called T-cell receptors (TCRs), to recognize foreign antigens. In order to create a broad array of TCRs, each immature T lymphocyte undergoes unique somatic rearrangements in its TCR loci during intrathymic development.

During this process, the DNA sequences in the TCR loci are deleted and circularized into episomal DNA molecules, also called *signal joint TCR excision circles* (sjTRECs).

Moreover it has been shown the existence of an inverse logarithmic correlation between sjTREC number and human age, probably reflecting the lifelong thymus involution [1].

In particular, the δ Rec- ψ J α sjTREC, which is present in about 70% of lymphocytes, is the main target for this qPCR approach.

In this method, firstly we extracted DNA from human blood, using a broadly commercially available QIAGEN mini kit, and obtained samples at an average volume of 50 μ l, concentration of 100–300 ng/ μ l.

Successively we set up a qPCR using specific TaqMan probes to quantify sjTRECs using a single copy Albumin gene to account for the amount of initial DNA, on a ROCHE LightCycler 480 II to carry out the PCR reactions.

Finally, we estimated ΔC_t between sjTREC and albumin intended as the difference between the PCR cycle number when sjTREC probe signal exceed the threshold minus the same cycle number referred to albumin probe's signal.

This absolute number, is directly linked to the quantity of sjTREC molecules physically present in the initial sample, so can be used as a single predictor in a linear regression model, explaining a statistically significant fraction of the total age variance.

2 Materials

2.1 Sample Assessment and Quality Control

To estimate an individual's biological age through this approach it is necessary to extract nuclear DNA from lymphocytes. To obtain this it is possible to use any DNA extraction kit or method capable of giving about a final yield of 20–50 ng/ μ l, and a minimum volume of 20 μ l, with acceptable purity ($A_{260}/280 > 2$ $A_{260}/230 > 2.2$) (e.g., QIAamp blood mini kit—QIAGEN).

If samples are not freshly extracted but stored in good condition (at least +4 °C) they can be used too. In this case they should be tested for quality and quantity using NanoDrop ND-3300 Spectrophotometer (ThermoScientific) or similar instrument.

The quality requirements for this technique are the same applied in any other qPCR application.

2.2 QPCR Setup and TaqMan Assay

1. ROCHE Light Cycler 480 Instrument II.

2. LightCycler® 480 96-well Block Kit.
3. LightCycler® 480 Multiwell Plate 96.
4. Custom TaqMan probes. TAMRA probes are dual-labeled probes used for real-time PCR applications using TaqMan chemistry. TaqMan TAMRA Probes feature a 5' fluorescent reporter dye (FAM, VIC, or TET) and 3' fluorescent quencher (TAMRA dye). All TaqMan TAMRA probes are HPLC—quality (*see Note 1*).

TaqMan Probe sjTREC, sequence 5'-(FAM) CACGGT-GATGCATAGGCACCTGC-3' (TAMRA).
5. TaqMan Probe sjTREC primers 5'-TCGTGAGAACGGTGAA TGAAG-3' and 5'-CCATGCTGACACCTCTGGTT-3'.
6. 2× TaqMan® Fast Advanced Master Mix (Applied Biosystems).
7. Albumin probe 6FAM-CCT gTCATg CCC ACA CAA ATC TCT CC—BBQ.
8. Albumin primers gCTgTC ATC TCT TgTgggCTg T and AAA CTC ATgggAgCTgCTgg T T.
9. Nuclease-free water.

2.3 Software Tools for Data Analysis

Any statistic software capable of managing several data point can be used.

Microsoft Excel or OpenOffice can also be used effectively. If using Microsoft Excel, it is better to install the Analysis ToolPak add-on.

3 Methods

3.1 DNA Extraction and Testing

1. To prepare sample for testing, extract DNA sample following the kit manufacturer's instructions. The minimum elution volume advisable is 20 µl.
2. To assess the quality of sample, use NanoDrop or similar instrument to verify that the DNA obtained has a minimum concentration of 50 ng/µl and the A260/280 and A260/230 ratio are respectively >2 and >2.2

3.2 PCR Protocol

Prepare all reactions in ice.

1. To simplify PCR preparation, it is advisable to prepare aliquots of each sample with the same concentration and pipet in the 96-well plates using a multichannel pipettor.
2. Thoroughly mix TaqMan® Fast Advanced Master Mix.
3. Thaw frozen samples and frozen TaqMan® Assays on ice. Resuspend by vortexing, then briefly centrifuge.

4. Calculate the total volume required for each component: volume for 1 reaction × total number of reactions.
5. Mix gently. Do not vortex. Centrifuge briefly and then prepare the PCR reaction plate.

Component	Volume (μ l) for 1 reaction	
	96-well plate	Final concentration
TaqMan® Fast Advanced Master Mix (2×)	10.0	1×
TaqMan® Assay primer/probe (20×)	1.0	1×
DNA template	2.0	0.001–100 ng/well
Nuclease-free water	7.0	
Total volume per reaction	20.0	

6. Before preparing reactions, it is advisable to set up the necessary protocol on LightCycler 480.

PCR run start with incubation at 50 °C for 2 min
95 °C for 10 min
45 cycles of denaturation at 95 °C for 15 s
Annealing/elongation at 60 °C for 30 s

7. PCR plates should be kept refrigerated until ready to be inserted in the machine.
8. All reactions should be performed in duplicates or better in triplicates, and PCR experiments should be repeated in case the C_t difference between replicates exceeded 1.5 (see Note 2).
9. Use the ROCHE LightCycler 480 II software to obtain C_t data and export in a suitable format for Microsoft Excel or Open-Office, like .csv format.

3.3 Data Analysis

1. Differently from typical qPCR applications, to estimate biological age is not necessary to calculate ddC_t . In fact, in this type of test there are not samples considered as internal controls, but we need to have a subpopulation of normal “control” samples necessary to calculate regression linear regression factor.

Samples composing this population should be relatively unaffected by pathologies interfering with lymphocytes

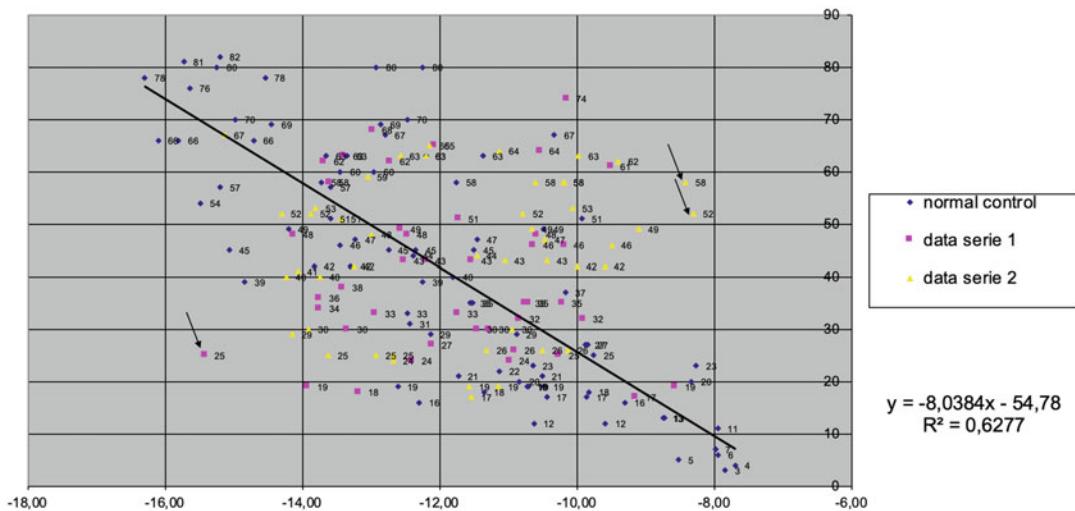


Fig. 1 Quantity of sjTRECs are indicated on x-axis and the age on y-axis. Age of each data point is indicated on the side

metabolism, and their age should vary from young age to old, the most uniform as possible (*see Note 3*).

It is advisable to choose available samples using three or four different age ranges (example: 0–20, 20–40, 40–60, 60–80 years).

Data can be grouped by sex, since even if no specific differences between male and female in sjTREC metabolism have been detected, males have on average lower sjTREC copy numbers than females [2].

2. Using Microsoft Excel or similar software, import .csv data into a new worksheet using available tools and match for the same sample paste its correspondent registry age data. For each sample, you must have C_t and age expressed in years. It is advisable to use age data referred at the moment of DNA extraction, in case the samples are obtained from biobank or similar structure.
3. It is advisable to calculate average values of replicate wells, excluding wells that exceed 1.5 C_t , and use average values for the following runs.
4. To estimate quantitative age prediction the algorithm used is linear regression, where age is considered as the dependent variable and dC_t values as the predictor. The prediction accuracy is evaluated using R^2 , which is the proportion of the age variability related to dC_t .

The formula used is:

$$\rho_{XY} = \sigma_{XY}/\sigma_X \sigma_Y$$

where X and Y are respectively sjTRECs abundance data and registry age, σ_X standard deviation of X and σ_Y standard deviation of Y , and σ_{XY} is covariance between X and Y .

ρ_{XY} should be then squared to obtain R^2 (determination coefficient) varying between 0, there is no correlation between data, and 1 meaning perfect correlation (*see Note 4*).

5. Through Microsoft Excel, it is possible to use the function LINEST to calculate linear regression coefficient.

It is also possible, using “SCATTER” graph format, to display graphically data to analyze for biological age, adding linear trend line and regression equation in graph option.

From a graphical point of view, we obtain a function described as

$$Y = \alpha + \beta x$$

where α is y -intercept and β the slope; if β tends to 1 there is a positive correlation between data, otherwise, if β tends to -1 , we have a negative correlation.

In Fig. 1 are shown example data. In this case, we have a series of data considered as normal control (blue dots), and two subpopulation (data series 1 and data series 2).

In this case, the regression line is calculated from normal control data and is $y = -8.0384x - 54.78$, with a $R^2 = 0.6277$.

Data show a negative correlation between quantity of sjTRECs and age (Y), with a good majority of data points grouped on the regression line, and some points relatively out of it, in particular:

Point “58” and “52” are samples showing a biological age better than their registry age, in fact they are comparable with the correspondent samples of 20–25 years; at the opposite, point “25” shows a biological age worse than its registry age, and comparable with samples in the 60–70–80 years area.

4 Notes

1. TaqMan probes for sjTRECs and albumin have to be designed and ordered at the TaqMan website.
2. Like any other qPCR applications, primers used in this method should be calibrated before use using qPCR manufacturer’s procedure.
3. It is necessary to recollect registry age data of all the samples used. If this information is not available, samples cannot be used.

4. The effectiveness of linear regression model implies a *high* number of samples. The higher the number of normal controls composing the population, the better.

References

1. Douek DC, McFarland RD, Keiser PH et al (1998) Changes in thymic function with age and during the treatment of HIV infection. *Nature* 396:690–695. <https://doi.org/10.1038/25374>
2. Zubakov D, Liu F, van Zelm MC et al (2010) Estimating human age from T-cell DNA rearrangements. *Curr Biol* 20:R970–R971. <https://doi.org/10.1016/j.cub.2010.10.022>



Chapter 15

QuantStudio™ 12K Flex OpenArray® System as a Tool for High-Throughput Genotyping and Gene Expression Analysis

Chiara Broccanello, Letizia Gerace, and Piergiorgio Stevanato

Abstract

Real time technology provides great advancements over PCR-based methods for a broad range of applications. With the increased availability of sequencing information, there is a need for the development and application of high-throughput real time PCR genotyping and gene expression methods that significantly broaden the current screening capabilities. Thermo Fisher Scientific (USA) has released a platform (QuantStudio™ 12K Flex system coupled with OpenArray® technology) with key elements required for high-throughput SNP genotyping and gene expression analysis. This allows for a rapid screening of large numbers of TaqMan® assays (up to 256) in many samples (up to 480) per run. This advanced real-time method involves the use of an array composed of 3,000 through-holes running on the QuantStudio™ 12K with OpenArray® block. The aim of this chapter is to outline the OpenArray® approach while providing a comprehensive in-depth review of the scientific literature on this topic. In agreement with a large number of independent studies, we conclude that the use of OpenArray® technology is a rapid and accurate method for high-throughput and large-scale systems biology studies with high specificity and sensitivity.

Key words TaqMan® assay, OpenArray® plate, SNP genotyping, Association analysis, Gene and miRNA expression profiling

1 Introduction

PCR-based techniques and in particular real-time PCR have become essential in diverse research areas such as agriculture, pharmacology, biomedicine, virology, microbiology, and biotechnology. Real-time technology provides great advancements over conventional PCR-based methods for a broad range of applications. Gene quantification, SNP genotyping, methylated DNA, and microRNA analysis are some of the frequently used applications [1]. A real-time PCR instrument is composed of a thermal cycler system coupled with a fluorescence detection system that detects the increasing quantity of amplified product at any given cycle in which a variable number of samples and reaction volume can be



Fig. 1 TaqMan® OpenArray® Plate

accommodated [2]. In recent years the development of high-throughput RT-PCR instruments has made it possible to minimize consumed reagents and costs while still delivering high performance and the reproducibility required for accurate results. The QuantStudio™ 12K Flex OpenArray® (Thermo Fisher Scientific, CA, USA) is a high-throughput system that enables profiling, confirmation and screening of many research applications such as gene expression profiling and sample genotyping. QuantStudio™ 12K Flex OpenArray® involves the use of arrays composed of 3072-through-holes (Fig. 1). These through-holes are externally hydrophobic and internally hydrophilic and are able to hold 33 nl of reagents via capillary action. The array is organized into 48 subarrays each composed of 64 through-holes (Fig. 2). The OpenArray® block can accommodate up to four array plates per run, in order to significantly increase the throughput to over 43,000 expression profiles, or more than 110,000 genotypes in a day. Inherent in the OpenArray® technology is the possibility that users can design the OpenArray® plates with the number of targets and samples that fit their experiments. Once the most suitable array format is chosen according to the type of experiment, the number of samples, and targets, an online tool allows the operator to configure and order the array plates. The OpenArray® plates contain custom or pre-designed TaqMan® GeneExpression or SNP Genotyping assays pre-loaded into the plate through-holes. TaqMan® Assays can also be designed and tested for robustness in previous validation experiments in order to select the best ones to insert into the arrays. The instrument detects FAM and VIC fluorescence signals of one or two probes for gene expression and miRNA or genotyping application.

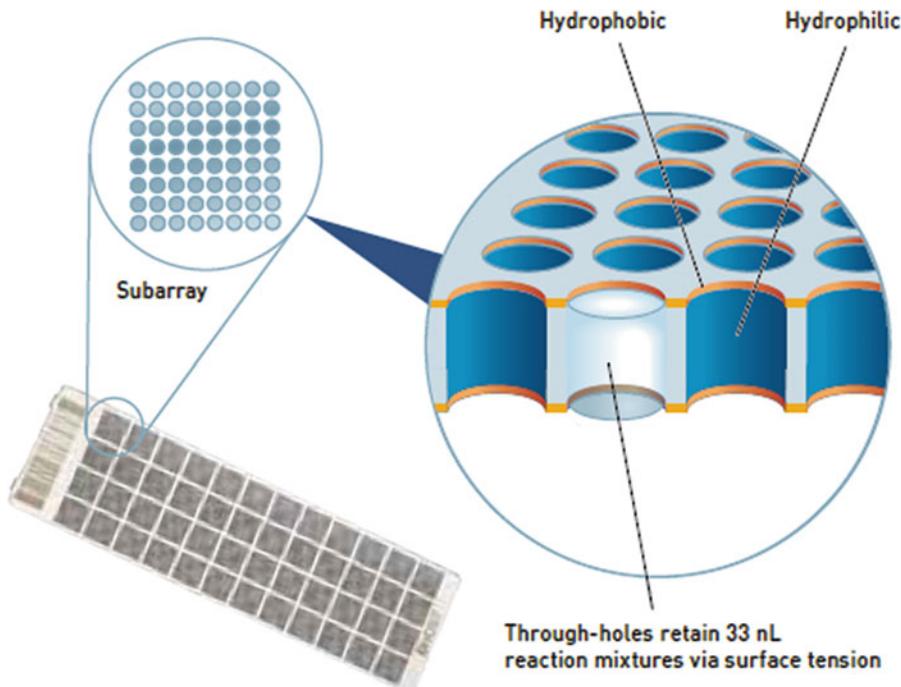


Fig. 2 TaqMan® OpenArray® Plate anatomy

2 Materials

2.1 Prepare Samples

1. MicroAmp® Optical 96-Well Reaction Plate (Thermo Fisher).
2. MicroAmp® Clear Adhesive Film (Thermo Fisher).
3. 10 µL of genomic DNA (or plasmid DNA or cDNA).
4. TaqMan® OpenArray® Genotyping Master Mix or TaqMan® OpenArray® Real-Time Master Mix.

2.2 Prepare 384-Well Sample Plates

1. OpenArray® 384-Well Sample Plates.
2. QuantStudio™ 12K Flex OpenArray® 384-Well Plate Seals.
3. Fine-tip marker.
4. OpenArray® Sample Tracker Software.

2.3 Prepare the QuantStudio™ 12K Flex OpenArray® Plates

1. QuantStudio™ 12K Flex TaqMan® OpenArray® plates.
2. QuantStudio™ OpenArray® AccuFill™ System.
3. QuantStudio™ 12K Flex OpenArray® Accessories Starter Kit. The accessories kit contains: QuantStudio™ 12K Flex OpenArray® Lids (six lids), QuantStudio™ 12K Flex OpenArray® Plugs (six plugs), QuantStudio™ 12K Flex OpenArray® Carriers (two carriers), QuantStudio™ 12K Flex OpenArray® Immersion Fluid (six syringes), QuantStudio™ 12K Flex

OpenArray® Immersion Fluid Tip, OpenArray® AccuFill™ System Loader Tips (one box of 384 tips), OpenArray® 384-Well Sample Plates (ten plates), QuantStudio™ 12K Flex OpenArray® 384- Well Plate Seals (ten seals).

4. QuantStudio™ OpenArray® Plate Press.
5. Foil seals.
6. Ethanol.
7. Razor blade.
8. Powder-free gloves.
9. Tweezers or forceps.

2.4 Perform the Instrument Run

1. QuantStudio™ 12K Flex with OpenArray® Block.
2. QuantStudio™ 12K Flex Software.

2.5 Analyze the Experiment Results

1. Thermo Fisher Cloud or desktop secondary analysis software (TaqMan® Genotyper, ExpressionSuite).

3 Methods

3.1 DNA Quantity and Quality

gDNA extraction (or cDNA preparation) is a crucial step for Open-Array® analysis. Samples need to be quantified and normalized at a concentration of 10–50 ng/µL with A_{260/230} ratio between 1.7 and 1.9 and A_{260/280} ratio between 1.7 and 1.9.

3.2 Track the Samples

The first step required to set up a run is the preparation of a 96-well plate and the corresponding 384-well plate. The OpenArray® Sample Tracker Software allows the tracking of the samples from the 96-well plate to the OpenArray® plate. Once selected the assay layout (genotyping or gene expression and the number of SNP or gene targets) samples names can be upload manually or through a *.csv file into the software. The OpenArray® Sample Tracker Software automatically maps the sample locations from the 96-well reaction plates to the appropriate locations in the 384-well sample plates and TaqMan® OpenArray® plates. Plate maps need to be exported in a *.csv file. Use this *.csv file to import setup information into the QuantStudio™ 12K Flex Software.

3.3 Transfer the Samples

Following the OpenArray® Sample Tracker Software information, load the 96-well plates. Each well is composed by 2.5 µL of gDNA or cDNA plus 2.5 of TaqMan® OpenArray® Genotyping Master Mix or TaqMan® OpenArray® Real-Time Master Mix. Transfer 5 µL of each PCR sample from 96-well plates to the 384-well sample plate and cover with a foil seal.

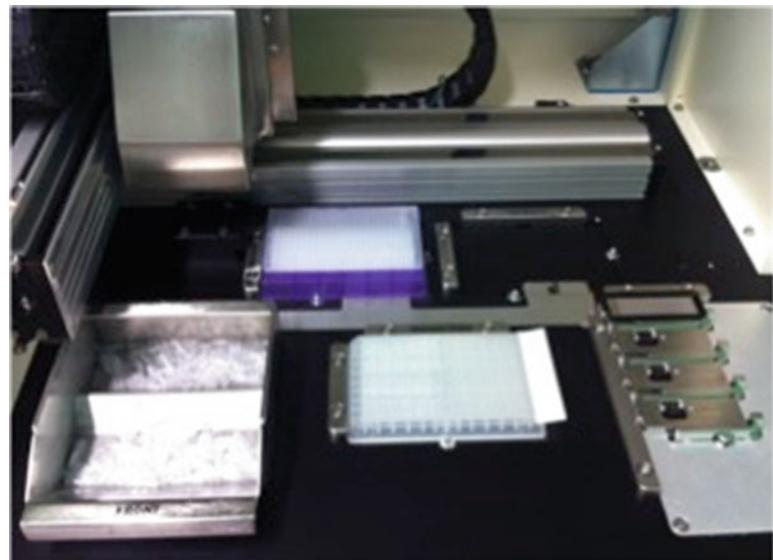


Fig. 3 QuantStudio™ OpenArray® AccuFill™ System deck, with a 384-well plate, a 384 AccuFill tips box and one OpenArray® plate

3.4 Prepare the QuantStudio™ 12K Flex OpenArray® Plate

The 384-well plate, one box of 384 AccuFill™ tips and an OpenArray® plate are placed on the AccuFill™ deck into their specific holders (Fig. 3). The QuantStudio™ OpenArray® AccuFill™ System automatically transfers samples and Master Mix from the 384-well plate to the OpenArray® plate.

The loaded OpenArray® plate must be sealed with 90 s with the OpenArray® Case Lid, filled with OpenArray® Immersion Fluid through the loading port and finally sealed by inserting the OpenArray® Plug into the loading port. The OpenArray® plate is now ready for the amplification run.

3.5 Perform the Instrument Run

The QuantStudio™ 12K Flex can be operated directly by the instrument touchscreen or by the QuantStudio™ 12K Flex Software on an instrument PC. To run set up requires two files: a .csv file, generated by the Sample Tracker software, that maps sample information (names and sample positions), a .tpf or .spf file, respectively for GeneExpression or Genotyping applications, (provided by Thermo Fisher Scientific), that maps the targets on the OpenArray® plate and provides the thermal cycling program to the instrument.

3.6 Data Analysis

This technique has been successfully applied on sugar beet to assess the genetic diversity among five cytoplasmic male steriles (CMSs), five pollinators, and five commercial varieties, for a total of 150 individuals (representing 15 genotypes) [3]. We used ten OpenArray® plates, each of them allowing the screening of 192 SNPs on 16 samples. The majority (95%) of the SNPs were polymorphic across

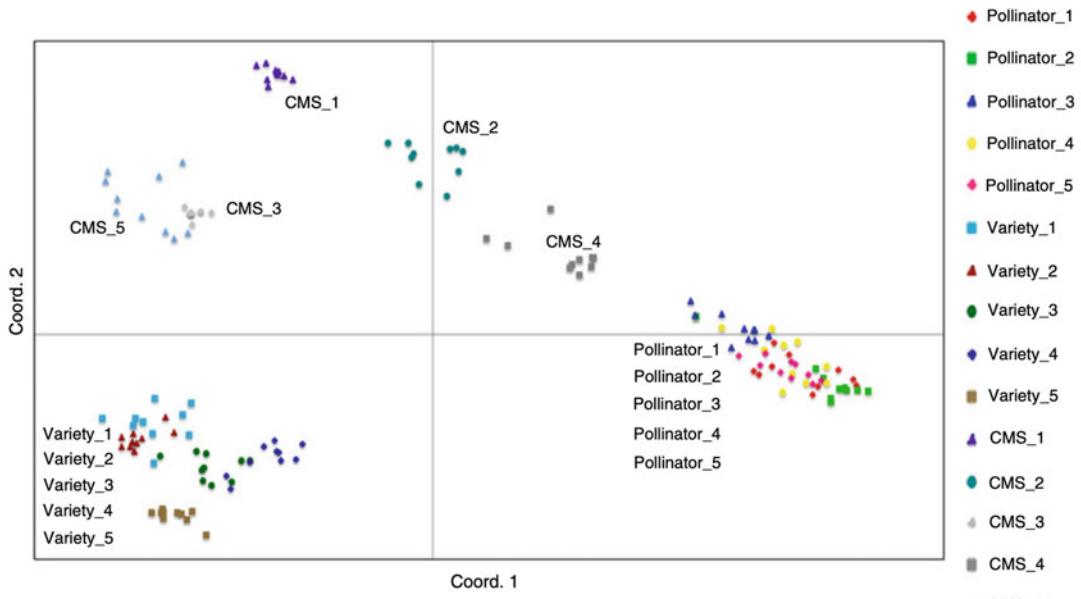


Fig. 4 Two-dimensional PCoA based on 192 SNPs of sugar beet genotypes (CMSs, Pollinators, Varieties). Each dot represents one individual. The first two principal coordinates of PCoA accounted for 58% of the total variation

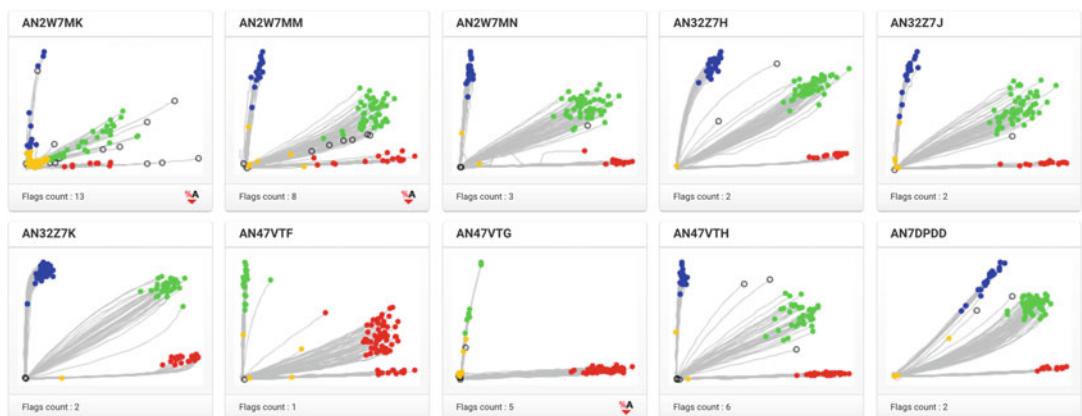


Fig. 5 Allelic discrimination plots of 10 SNPs in sugar beet. Each dot represents the genotype of one sample at a specific SNP. Blue dots represent homozygote samples for Allele1, red dots represent homozygote samples for Allele 2, and green dots represent heterozygote samples. Yellow dots are nonamplifying samples and circles are undetermined genotypes

sugar beet genotypes, producing high-quality signals with a rate of undetermined results accounting for only 0.45%. With the obtained dataset, we calculated the genetic distance and heterozygosity across the 15 genotypes. Data were visually represented through a principal component analysis as shown in Fig. 4. The same 192 SNPs panel was used to establish the presence or absence

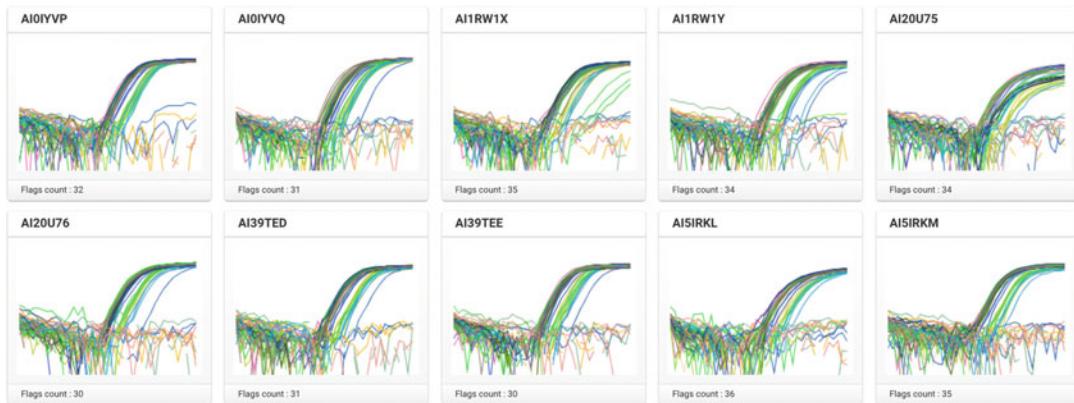


Fig. 6 Amplification plots for 10 targets (genes) in sugar beet

of common characteristics or phylogenetic relationships between two rhizomania resistance sources, Rz1 and Rizor, in sugar beet. An example of allelic discrimination plots obtained with OpenArray® technology is showed in Fig. 5 (see Note 1).

Noteworthy, OpenArray® technology can be successfully used for association analysis. We identified and mapped on chromosome 5, through the screening of 384 SNPs, the first nematode tolerance gene (*HsBrm-1*) from *Beta vulgaris* ssp. *maritima*. Additional association studies have been carried out to find SNP markers associated with, root length [4] and bolting tendency [5] in sugar beet. Finally, OpenArray® technology has been applied in genomic selection of sugar beet [6, 7].

Gene expression analysis (see Note 2) has been done in order evaluate the nutritional status of sugar beet under different nutrient availability conditions. We designed OpenArray® plates for analysing the gene expression of 53 sugar beet gene targets putatively involved in sulfate nutrition [8]. The gene expression level was evaluated in deprived and supplied leaves and roots. Thanks to the OpenArray® technology we were able to identify that relative expression of 53 genes in leaves was significantly higher compared to the expression of the same gene panel in roots. Moreover we highlighted the presence of potential biomarkers involved in sulfate nutrition. Representative expression profiles of these genes are showed in Fig. 6. These 53 sugar beet genes has now been recognised as a signature for monitoring the biostimulant effects of two microalgae extracts (*Chlorella vulgaris* and *Scenedesmus quadricauda*) [9]. Our results highlighted that relative gene expression is generally higher in plants treated with *S. quadricauda* than with *C. vulgaris*. We used the OpenArray® technology to select SNP markers associate to relevant agronomic and productive traits in apple [10] and in *Jatropha curcas* breeding [11] (see Note 3).

4 Notes

1. In recent years, arrays are increasingly used for research in the medical science. Particularly, genotyping application has been involved in the study of Parkinson disease. Ross et al. in 2018 [12] developed an array to study 22 SNPs associated to essential tremor. The extreme versatility of arrays makes them an efficient system for assessing blood types instead of the use of hemagglutination. Denomme and Schanen in 2015 [13] developed arrays for mass-scale donor red cell genotyping to be used as a screening tool.
2. Medical research has also widely adopted the OpenArray® technology for gene expression analysis. Cook et al. in 2018 [14] developed and validated a gene expression profile to evaluate mechanisms underlying melanoma tumor biology and improve the prediction of metastasis formation. Arrays can also be used for miRNA screening. For example, a chip was developed to investigate the expression of miRNA to discriminate between malignant pleural mesothelioma patients (MPM) and subjects with past asbestos exposure (PAE) [15]. Thermo Fisher Scientific made available some easy to use OpenArray® plates containing human predesigned assays, such as “Cancer Panel,” “Signal Transduction Panel,” “Stem Cell Panel,” “Inflammation Panel,” “Kinome Panel,” “Advanced Micro-RNA Panel,” and “Pharmacogenomics Express Panel.” Recently, Gnani et al. in 2017 [16] used the “Cancer Panel” to analyzed 624 cancer genes to study the genetic expression during hepatocarcinogenesis, while ten housekeeping genes related to leiomyoma has been found through the use of a predesigned TaqMan® human endogenous control panel [17]. The “MicroRNA Panel” was used by Hudsun et al. in 2013 [18] to discover an overexpression of miR-10a and miR-375 in medullary thyroid carcinoma. An interesting comparison of miRNA analysis using two different platforms, OpenArray® and Dinamic Array, has been done by Farr et al. in 2015 [19]. The author found that the most reproducible results, with less interrun and intrarun variations, were obtained by the Open array platform.
3. Analogous approaches were also adopted in animal genetics to identify useful SNPs associated with variations at regulatory regions of the milk protein genes in the Sardinian sheep [20]. A panel of SNP markers was also tested by Viale et al. in 2017 [21] using OpenArray® technology to fingerprinting chicken breeds. OpenArray® technology has proved to be very effective also for individual identification and paternity

test. The use of SNPs for individual identification and paternity test, on a high-throughput real-time platform, has become easier and more accurate than STR assays. Pomeroy et al. in 2009 [22], for forensic analysis, developed a method for individual identification using only seven SNPs.

References

1. Lamas A, Franco CM, Regal P et al (2016) High-throughput platforms in real-time PCR and applications. In: Polymerase chain reaction for biomedical applications. InTech, pp 15–38
2. Stevens J, Heid C, Livak KJ et al (1996) Real time quantitative PCR. *Genome Res* 6:986–994
3. Stevanato P, Broccanello C, Biscarini F et al (2013) High-throughput RAD-SNP genotyping for characterization of sugar beet genotypes. *Plant Mol Biol Rep* 32:691–696
4. Stevanato P, Trebbi D, Saccomani M et al (2017) Single nucleotide polymorphism markers linked to root elongation rate in sugar beet. *Biol Plantarum* 61:48–54
5. Broccanello C, Stevanato P, Biscarini F et al (2015) A new polymorphism on chromosome 6 associated with bolting tendency in sugar beet. *BMC Genet* 16(1)
6. Biscarini F, Marini S, Stevanato P et al (2015) Developing a parsimonious predictor for binary traits in sugar beet (*Beta vulgaris*). *Mol Breed* 35(1)
7. Biscarini F, Nazzicari N, Broccanello C et al (2016) “Noisy beets”: impact of phenotyping errors on genomic predictions for binary traits in *Beta vulgaris*. *Plant Methods* 12:36
8. Stevanato P, Broccanello C, Moliterni VMC et al (2018) Innovative approaches to evaluate sugar beet responses to changes in sulfate availability. *Front Plant Sci* 9:14
9. Barone V, Baglieri A, Stevanato P et al (2017) Root morphological and molecular responses induced by microalgae extracts in sugar beet (*Beta vulgaris* L.). *J Appl Phycol* 30:1061–1071
10. Muranty H, Troggio M, Sadok IB et al (2015) Accuracy and responses of genomic selection on key traits in apple breeding. *Hortic Res* 215:60
11. Trebbi D, Ravi S, Broccanello C et al (2019) Identification and validation of SNP markers linked to seed toxicity in *Jatropha curcas* L. *Scientific reports* 9:10220
12. Ross JP, Mohtashami S, Leveille E et al (2018) Association study of essential tremor genetic loci in Parkinson’s disease. *Neurobiol Aging* 66:178–e13
13. Denomme GA, Schanen MJ (2015) Mass-scale donor red cell genotyping using real-time array technology. *Immunohematology* 31:69–74
14. Cook RW, Middlebrook B, Wilkinson J et al (2018) Analytic validity of decision dx-melanoma, a gene expression profile test for determining metastatic risk in melanoma patients. *Diagn Pathol* 13:13
15. Cavalleri T, Angelici L, Favero C et al (2017) Plasmatic extracellular vesicle microRNAs in malignant pleural mesothelioma and asbestos-exposed subjects suggest a 2-miRNA signature as potential biomarker of disease. *PLoS One* 12 (5)
16. Gnani D, Romito I, Artuso S et al (2017) Focal adhesion kinase depletion reduces human hepatocellular carcinoma growth by repressing enhancer of zeste homolog 2. *Cell Death Differ* 24:889
17. Almeida TA, Quispe-Ricalde A, de Oca FM et al (2014) A high-throughput open-array qPCR gene panel to identify housekeeping genes suitable for myometrium and leiomyoma expression analysis. *Gynecol Oncol* 134:138–143
18. Hudson J, Duncavage E, Tamburrino A et al (2013) Overexpression of miR-10a and miR-375 and downregulation of YAP1 in medullary thyroid carcinoma. *Exp Mol Pathol* 95:62–67
19. Farr RJ, Januszewski AS, Joglekar MV et al (2015) A comparative analysis of high-throughput platforms for validation of a circulating microRNA signature in diabetic retinopathy. *Sci Rep* 5:10375
20. Noce A, Pazzola M, Dettori ML et al (2016) Variations at regulatory regions of the milk protein genes are associated with milk traits and coagulation properties in the Sarda sheep. *Anim Genet* 47:717–726

21. Viale E, Zanetti E, Ozdemir D et al (2017) Development and validation of a novel SNP panel for the genetic characterization of Italian chicken breeds by next-generation sequencing discovery and array genotyping. *Poultry Sci* 96:3858–3866
22. Pomeroy R, Duncan G, Sunar-Reeder B et al (2009) A low-cost, high-throughput, automated single nucleotide polymorphism assay for forensic human DNA applications. *Anal Biochem* 395:61–67



Chapter 16

Digital PCR and the QuantStudio™ 3D Digital PCR System

Marion Laig, Christie Fekete, and Nivedita Majumdar

Abstract

The great promise of digital PCR is the potential for unparalleled precision enabling accurate measurements for detection and quantification of genetic material. This chapter walks the reader through the fundamentals of digital PCR technology including digital PCR modeling using Poisson statistics. It describes a highly successful implementation of digital PCR technology using the chip-based nanofluidic Applied Biosystems™ QuantStudio™ 3D digital PCR system. It reviews the large number of applications where digital PCR is poised to make significant impacts. These include applications where detection of rare genetic targets is prioritized such as liquid biopsy, rare mutation detection, confirmation of NGS variant detection, detection of fusion transcripts, detection of chimerism and GMO detection and monitoring. These further include applications where accurate quantification of genetic targets is prioritized such as generation of references and standards, copy number variation, and NGS Library quantification.

Key words Digital PCR, Poisson, Limit of detection, Rare mutation, Copy number, GMO

1 Introduction to Digital PCR

1.1 How Digital PCR Works

Digital PCR (dPCR) is rapidly gaining traction as a method of choice to detect and quantify nucleic acid molecules (Fig. 1). The capability of dPCR to accurately determine the concentration of rare targets (like rare cancer mutations) has been the key driver of adoption, particularly for clinical research applications like liquid biopsy and cancer recurrence monitoring.

Although the term “digital PCR” was first coined in 1999 [1], the method of quantification by performing endpoint PCR on highly diluted samples (limiting dilution PCR) was in use as early as 1990 [2]. This technique was developed in parallel with real-time PCR in the 1990s. Early digital PCR was highly laborious, requiring careful dilutions and gel electrophoretic analysis, so it was quickly overtaken by simple and powerful real-time PCR for the vast majority of applications. However, recent availability of easy-to-use, commercial digital PCR systems has prompted a resurgence in digital PCR, with publication rate increasing from <10 per year

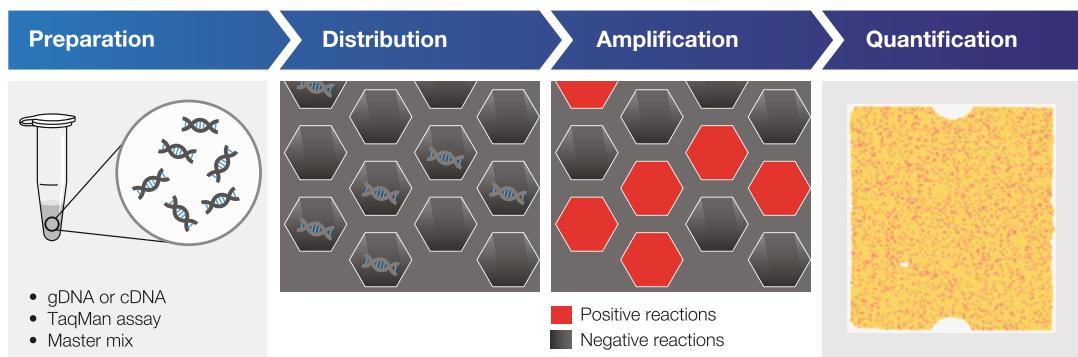


Fig. 1 Digital PCR uses partitioning to achieve low level detection. The PCR reaction mix is distributed across 20,000 fixed partitions and subjected to amplification. Postamplification, the resulting fluorescent signal is captured and the results quantified using AnalysisSuite™ Cloud software

in 2009 to >400 in 2018 (<https://www.ncbi.nlm.nih.gov/pubmed/?term=%22digital+pcr%22>. Accessed 28 Nov 2018).

Digital PCR works by partitioning a sample of nucleic acid into many individual, parallel PCR reactions. This partitioning is typically achieved with droplets or fixed partitions. The fundamental principles are equivalent, so here we will focus on fixed partitions. In an idealized case, the sample is diluted such that each partition contains one target molecule (positive) on average. Some partitions will not contain any target molecule (negative), and some will contain more than one target molecule. During PCR, amplification will occur in each partition containing one or more target molecules. This amplification can be detected using TaqMan chemistry with dye-labeled probes to ensure target specificity. Following PCR, the fluorescent signal from each partition is captured and used to assign a result to each of those partitions; either a positive or negative amplification. Calculations using the ratio of positive to negative results are then used to generate an absolute count of the number of target molecules in the sample, without the need for standards or endogenous controls.

The Applied Biosystems™ QuantStudio™ 3D Digital PCR System uses nanofluidic chip technology featuring 20,000 reaction wells for consistent volume sample partitioning. The chip consumable is sealed, which minimizes contamination from amplicons and other contaminating nucleic acids. This simple, manual system enables even users with limited experience and space to take advantage of all digital PCR has to offer.

1.2 Modeling: The Poisson and Poisson Plus Algorithm

After each partition undergoes PCR amplification and analysis, partitions with at least one molecule will emit a much higher range of fluorescence than the partitions with no molecules, making it possible to directly count the number of partitions with and without amplified product. The QuantStudio 3D AnalysisSuite

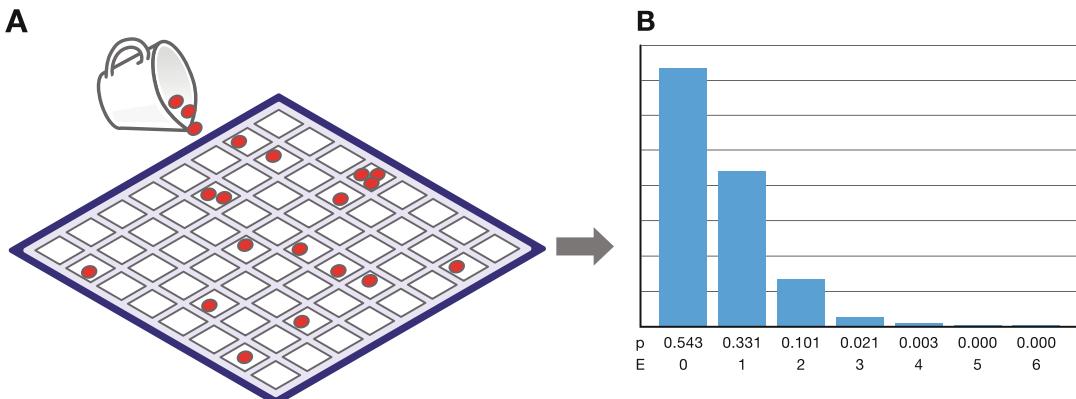


Fig. 2 Due to random distribution, we cannot be assured that each positive reaction received only a single molecule. The probability of a reaction receiving zero, one, two, three etc. copies is described by the Poisson model. Poisson statistics “corrects” for reactions containing multiple molecules

Cloud Software assesses the quality of each PCR reaction (like other digital PCR systems) and assigns a quality score. Only those through-holes that meet a user settable quality threshold are included in downstream analysis. This flexibility ensures there is minimal dead volume (wasted reaction) and minimal chances of missing a rare event.

Now, due to the random assortment, we cannot be assured that each positive reaction receives only a single molecule (Fig. 2). Thus, a straight count of partitions showing amplification may lead to undercounting of the total number of molecules present at the start of the experiment. The QuantStudio 3D AnalysisSuite Cloud Software corrects the count by applying Poisson modeling. The measured number of negatives occurring across the total number of partitions can be used to estimate the average number of molecules per partition using Poisson statistics. This average count of molecules per partition is then converted into concentration by dividing by partition volume (for a detailed description, see [3]).

In this standard approach using Poisson modeling, identical partition sizing is assumed. Violations of this assumption result in underestimation of target quantity, especially at higher concentrations. The QuantStudio 3D AnalysisSuite Cloud Software implements the Poisson-Plus Model to correct for this underestimation [4]. The extent of the volume variation is measured on a per chip array basis, using the ROX fluorescence level as a proxy for effective load volume per through-hole.

1.3 Limit of Detection (LoD), Sensitivity, Precision, and Dynamic Range

The limit of detection (LoD) is defined as the lowest quantity or concentration that can be reliably detected. Consideration of LoD becomes important in cases where you are making measurements near the limits of concentration detectable by your system. For example, when quantifying rare alleles, detecting early disease

stage in cancer research samples or monitoring disease progression, LoD becomes critical. The LoD varies with sample quality and the individual assay used for detection.

Generally, to apply the Poisson model, you must have at least one negative reaction, which determines the upper limit of detection, and in parallel, at least one positive reaction, which determines the lower limit of detection. While this determines the theoretical dynamic range of the digital PCR system, the measurement variability at both extremes of concentration (nearly all negative or nearly all positive) is very high. This variability translates into a reduced level of precision at these concentrations. The limit of detection is also challenged by the presence of false positive and false negative calls at both ends of the concentration spectrum. These are practical noise factors that plague all digital PCR systems. To work around them for the measurement of rare alleles, it is recommended to run wild-type only negative control chips to understand the distribution of false positives occurring for the given sample-assay system. For reliable detection, the expectation would be to detect higher counts of positives in the rare cluster than detected when chips are running wild-type only control samples.

Another practical recommendation for improving the sensitivity or improving the lower limit of detection is to interrogate as high a total volume of reaction as possible. Pooling the data across multiple QuantStudio 3D chips is one way to interrogate a higher total reaction volume. This also improves the measurement precision owed to the higher number of total reactions now participating in the measurement.

While the measurable dynamic range for a simple sample is fundamentally dictated by the number of partitions used to interrogate that sample, it is also possible to increase the supported range by running multiple dilutions of the sample across multiple chips, then combining the resultant data. The QuantStudio 3D AnalysisSuite Cloud Software uses an advanced Poisson model that performs this combination such that low concentration dilution points drive the upper limit of detection, high concentration dilution points drive the lower limit of detection and the data from all dilution points raise the precision by increasing the partition count (for a detailed description, see [3]).

1.4 Digital PCR Compared to qPCR

The digital PCR method is the next generation of qPCR technique and complements the strengths of qPCR analysis. A brief comparison of the techniques highlighting the strengths of each is as follows:

1.4.1 Absolute Quantification

Digital PCR can eliminate the need for use of reference or standards by providing absolute quantification. This saves cost and effort, eliminating the requirement for running standard curves. Digital PCR also provides higher quality answers in terms of measurement

accuracy, sensitivity and precision. The improved measurement precision reduces the burden of running large numbers of replicate qPCR reactions. However, it is important to note that the biological variation is not captured without running biological replicates in both digital and quantitative PCR. A further advantage with digital PCR is the support for linear detection of small fold changes. The main advantage for quantitative PCR over digital comes from the wider dynamic range supported by qPCR.

1.4.2 Digital PCR Enrichment Effect

Digital PCR partitions the reaction into multiple subpartitions, effectively enriching the target of interest. Suppose there is exactly one copy of rare target in a background of 100 wild-type targets in your original reaction. Figure 3 illustrates this principle. In qPCR (a), all molecules are allocated to the same reaction. The ratio of rare to wild-type target is 1:100, so the rare target is difficult to detect in the extensive background of wild-type targets. In digital PCR (b), as you split the sample into many subparts, the rare target is allocated to its own reaction. The ratio of rare to wild-type target is 1:0. There is no competition from the wild-type targets. The wild-type molecules get distributed, effectively increasing the rare to wild-type ratio in the well where the rare target is intercepted. This is the key to the great sensitivity that digital PCR can deliver.

1.4.3 Resistance to Inhibitors

Digital PCR is more resistant to inhibitors in the reaction than qPCR (Fig. 4). In qPCR, sample quality will impact amplification curve shape and hence the Cq results (a). In digital PCR, the read derives a yes or no answer differentiated by the nonamplified vs. amplified end point signal. The course of the amplification curve itself is irrelevant as long as these can be told apart. Thus, digital PCR digitizes across the noise (b) and still yields a high-quality answer.

FFPE samples are notoriously difficult template for PCR [5]. The higher resistance of digital PCR to inhibitors makes it ideal for detecting low level molecules in FFPE samples. We tested an FFPE sample that had previously shown inconsistent results with another method (ARMS-PCR). The presence of melanin that was copurified with the DNA interfered with detection by ARMS. Mutation detection in the same samples by digital PCR using the QuantStudio 3D digital PCR system confirmed the mutation with much higher confidence and consistency during replicate testing (Fig. 5).

1.5 Assays for Digital PCR with QuantStudio 3D Digital PCR System

Two types of assays are used in digital PCR applications with the QuantStudio 3D digital PCR system (Table 1): TaqMan SNP genotyping/Liquid Biopsy dPCR Assays and TaqMan Gene Expression/Copy number assays:

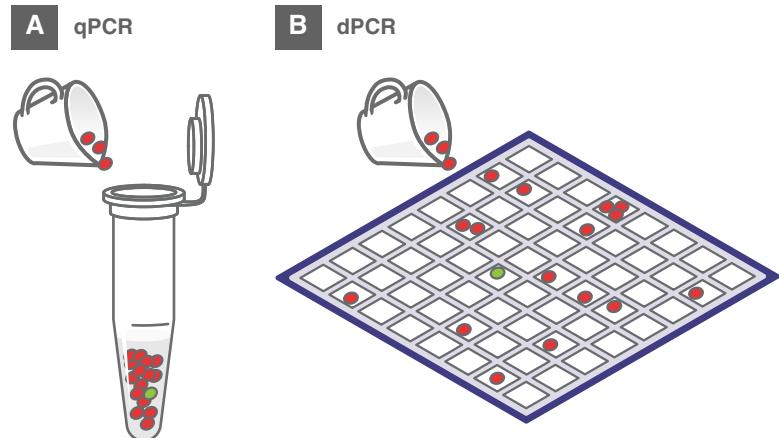


Fig. 3 Enrichment effect of digital PCR. (a) In qPCR, multiple molecules are allocated to the same reaction well. A rare mutation allele (green) is difficult to detect in the large number of wild-type alleles (red). (b) In digital PCR, the sample is diluted and partitioned into multiple reaction wells so that each reaction well receives a single molecule on average. There is an enrichment for the mutant allele away from the wild-type allele, and the mutant allele is detected

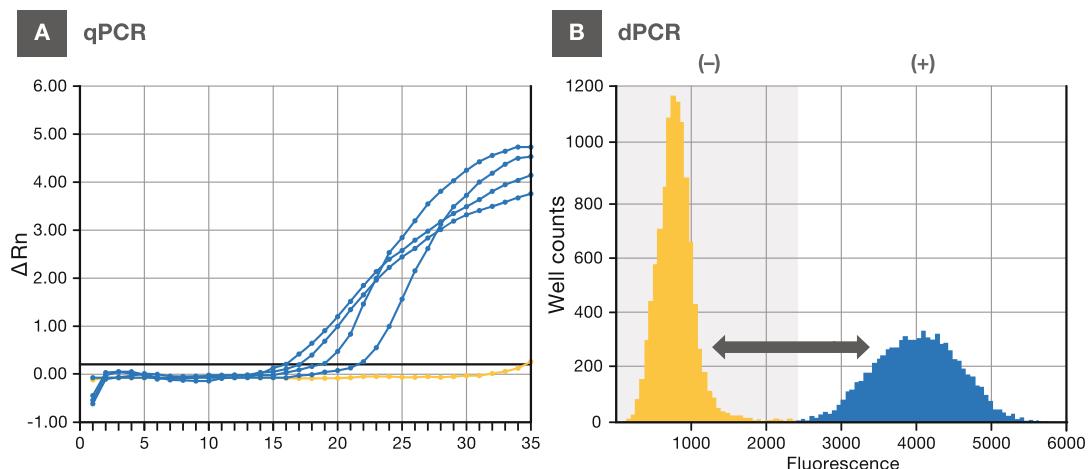


Fig. 4 Digital PCR is more resistant to inhibitors than qPCR

1. TaqMan SNP genotyping assay-based assays featuring two MGB probes, one probe carrying a VIC label and the other probe carrying a FAM label, for simultaneous detection of mutant and wild-type alleles.
2. Gene expression assays featuring one probe, usually with a FAM label.

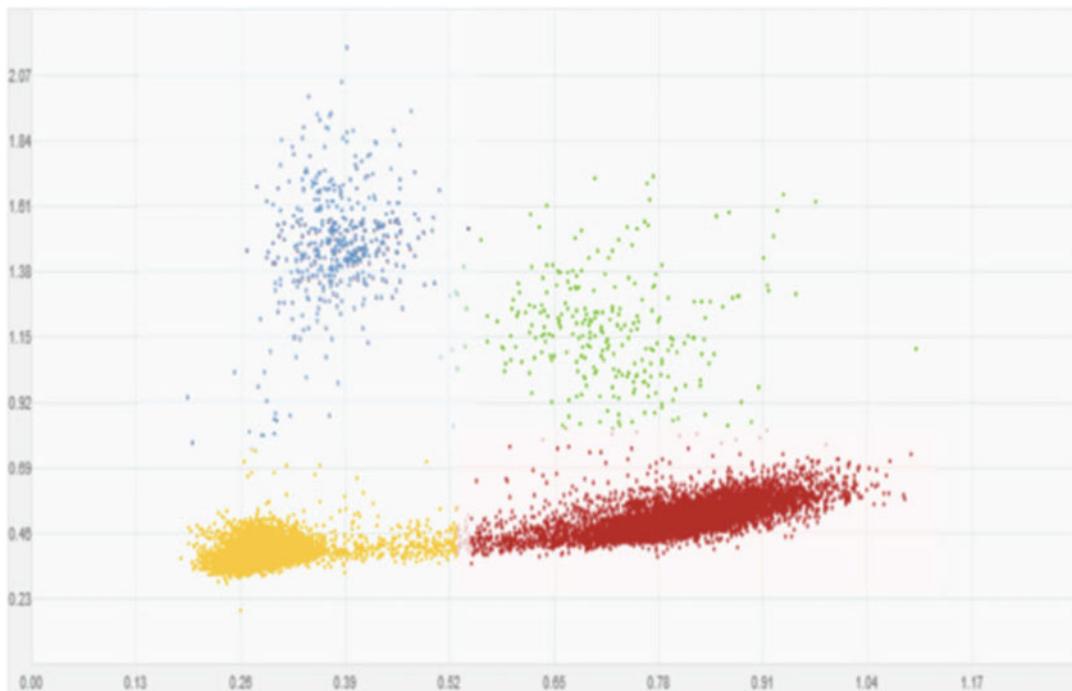


Fig. 5 Example of mutation detection in an FFPE sample. Applied Biosystems™ TaqMan™ Liquid Biopsy dPCR Assays achieve good cluster separation in challenging sample types

Table 1
Assays for digital PCR applications

Type of assay	Number of probes	Dye label	Application
Copy number Gene expression	1	FAM or VIC	Copy number GMO detection Standards quantification NGS library quantification Fusion transcript detection
Liquid biopsy SNP genotyping	2	FAM and VIC	Rare mutation detection

1.5.1 *TaqMan Liquid Biopsy dPCR Assays*

For rare mutation detection in liquid biopsy research, we developed the TaqMan Liquid Biopsy dPCR assay set. All assays have two probes for simultaneous detection of mutant allele (FAM dye label) and wild-type allele (VIC dye label). The assays address clinically actionable somatic mutations, single nucleotide variants (SNVs) and small indels. All assays were optimized for use in digital PCR with the QuantStudio 3D digital PCR system. Cluster separation and signal intensity are enhanced for these assays on the digital PCR platform compared to TaqMan SNP genotyping assays.

Multiple assay designs were wet lab-tested for each somatic variant, and the best performer was selected for the final assay.

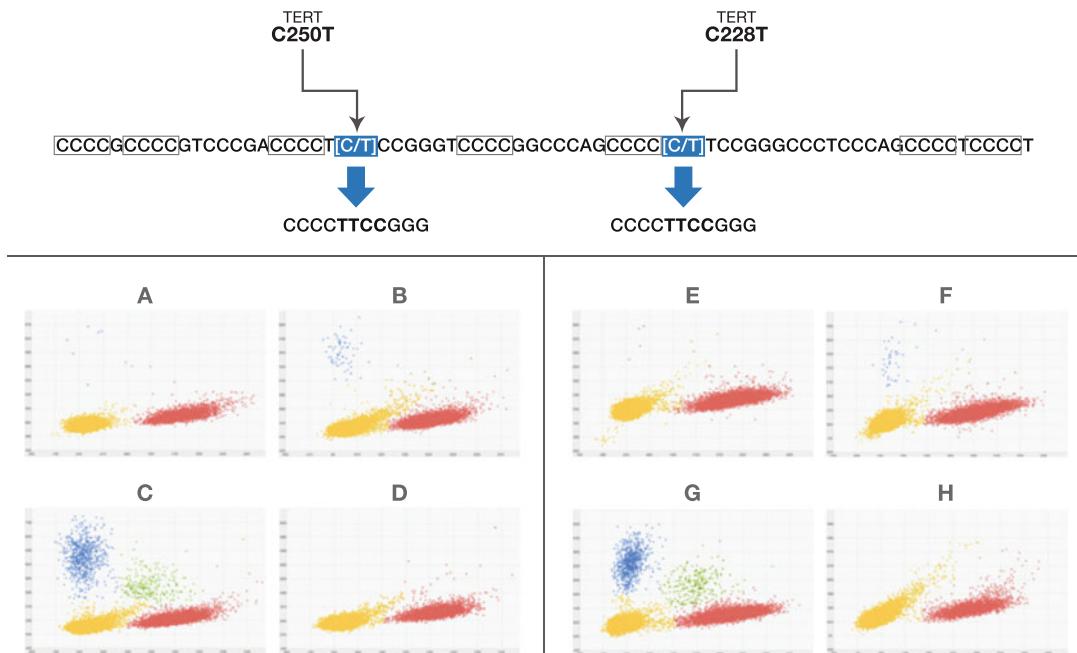


Fig. 6 Context sequence of two mutations in the promoter region of the Telomerase Reverse Transcriptase (TERT) gene showing that both mutations create an identical 11 nucleotide CCCCTTCCGGG motif that serves as binding site for the ETS transcription factor. (a–h) Performance Testing of two TaqMan Liquid Biopsy dPCR Assays detecting mutations in the TERT promoter region. The assays were tested on wild-type genomic DNA with spiked in mutant plasmid. (a–d) Assay Hs000000093_rm detecting the TERT C250T mutation (COSM1716559) located at –146 base pairs upstream of the ATG start codon. (a) 0.1% mutation. (b) 1% mutation. (c) 10% mutation. (d) wild-type only control. (e, f) Assay Hs000000092_rm detecting the TERT C228T (COSM1716558) mutation located at –124 base pairs upstream of the ATG start codon. (e) 0.1% mutation. (f) 1% mutation. (g) 10% mutation. (h) Wild-type only control

Template for performance testing was plasmid containing the mutant allele (GeneArt, Thermo Fisher Scientific) spiked into wild-type genomic DNA (Thermo Fisher Scientific) at concentrations reflecting 10% and 1% mutation rate as well as 0.1% mutation rate for selected assays. Fifteen nanograms of genomic DNA was loaded per digital chip and cycled according to protocol [6]. Data Analysis was done using QuantStudio 3D AnalysisSuite Cloud Software.

Figure 6 shows an example of assay development for the Taq-Man Liquid Biopsy dPCR Assays set for two noncoding mutations in the TERT gene. Both mutations create an identical 11 nucleotide CCCCTTCCGGG motif that serves as binding site for the ETS transcription factor, facilitating gene expression from the TERT promoter [7]. The high frequency of these mutations suggests that they may comprise early genetic events in the genesis of several cancer types.

2 Digital PCR Applications

Applications covered in this section

1. Liquid biopsy applications.
 - (a) Rare mutation detection.
 - (b) Confirmation of NGS variant detections.
 - (c) Detection of fusion transcripts.
 - (d) Detection of chimerism.
2. Generation of references and standards.
3. GMO detection and monitoring.
4. Copy number variation.
5. NGS library quantification.
6. Low-level pathogen detection.
7. Low-fold differential gene expression.

2.1 Liquid Biopsy Applications

Liquid biopsy is a fast-evolving research field that enables analysis of cell-free DNA (cfDNA) and circulating tumor cells (CTCs) isolated from plasma. It captures tumor heterogeneity that may be missed by conventional solid tumor biopsy. Liquid biopsy takes advantage of the fact that tissue cells, including tumor cells, shed DNA into the bloodstream as cell-free DNA (reviewed in [8]). Part of the cfDNA is composed of circulating tumor DNA (ctDNA) that can be analyzed for the presence of cancer-causing mutations. Combined with sensitive and highly specific detection methods, liquid biopsy allows for monitoring mutation burden, detection of minimal residual disease, recurrence, acquired resistance and early disease [9]. Liquid biopsy is noninvasive and does not encounter the spatial temporal restrictions of solid tumor biopsy since ctDNA is shed into the bloodstream. Liquid biopsy detects mutations from remote metastases and at various stages of tumor evolution.

2.1.1 Rare Mutation Detection

Rare mutation analysis is a major application in oncology research, especially in the emerging field of liquid biopsy. Digital PCR offers the sensitivity needed to detect mutant ctDNA in a high background of wild-type cfDNA. TaqMan assays contribute the specificity needed to discriminate mutant and wild-type alleles.

Garcia-Saenz et al. [10] used digital PCR with the QuantStudio 3D digital PCR system and TaqMan Liquid Biopsy dPCR Assays to investigate three hotspot mutations in the PIK3CA oncogene in breast cancer samples. E542K, E545K, and H1047R mutations account for 70% of observed PIK3CA mutations in breast cancer. In this study, cfDNA was extracted from plasma and analyzed by digital PCR. Figure 7 shows the result for two samples. Mutation burden detected in solid tumors and ctDNA were

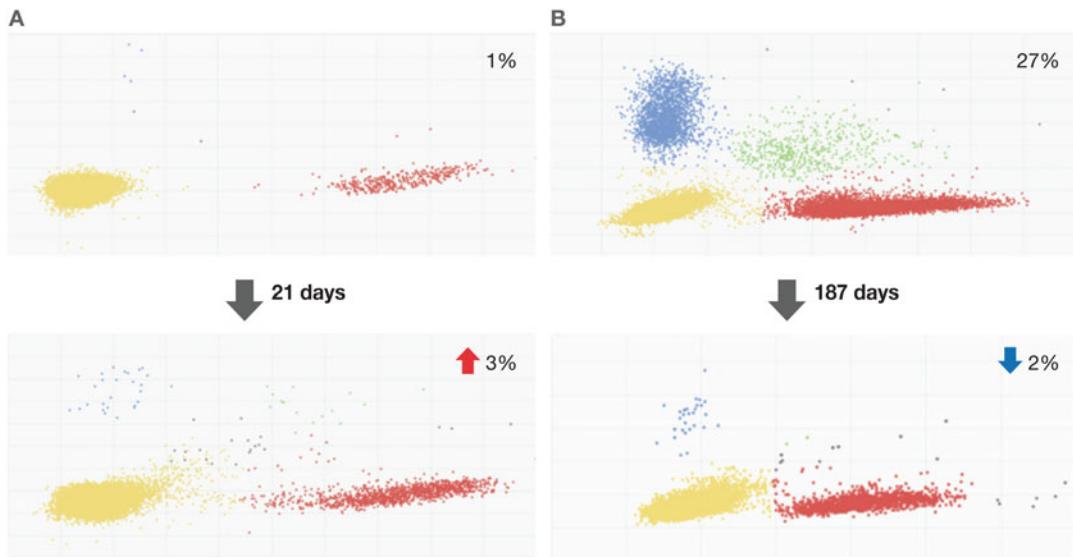


Fig. 7 Monitoring PIK3CA mutation burden to assess drug efficacy using digital PCR in two breast cancer samples. cfDNA was tested before treatment and after treatment. The scatter plots show the mutant allele in the FAM cluster (blue), wild-type allele in the VIC cluster (red). Double positive data points are shown in green. (a) This sample shows low initial mutation burden for the PIK3CA E542K mutation that increases during treatment. This sample did not respond to the treatment. (b) This sample shows high mutation burden for the PIK3CA E545K mutation before treatment that decreases during treatment. This sample responded to the treatment. Data courtesy Dr. Atocha Romero, Medical Oncology Department, Hospital Clinico San Carlos, Madrid, Spain

concordant, and changes in the percent of PIK3CA mutant allele were consistent with changes in lesion size detected on images.

2.1.2 Confirmation of NGS Variant Detection

Tumor genomes evolve constantly, accumulating new somatic mutations, adapting to their environments and undergoing clonal selection [11]. The first step in molecular characterization of a tumor is identification of causative mutations. This is often accomplished by next generation sequencing (NGS).

For a complete solution from discovery to validation, Next Generation Sequencing (NGS) with Oncomine™ cfDNA Assays is used in the discovery step to identify mutations present in a sample (Fig. 8). The Oncomine cfDNA Assays and the Ion S5™ System enable tumor heterogeneity research studies from minimal input DNA. Digital PCR with TaqMan Liquid Biopsy dPCR Assays may be used for orthogonal verification and downstream mutation monitoring.

When testing samples with low mutation frequency, we found good concordance between NGS and digital PCR (Table 2).

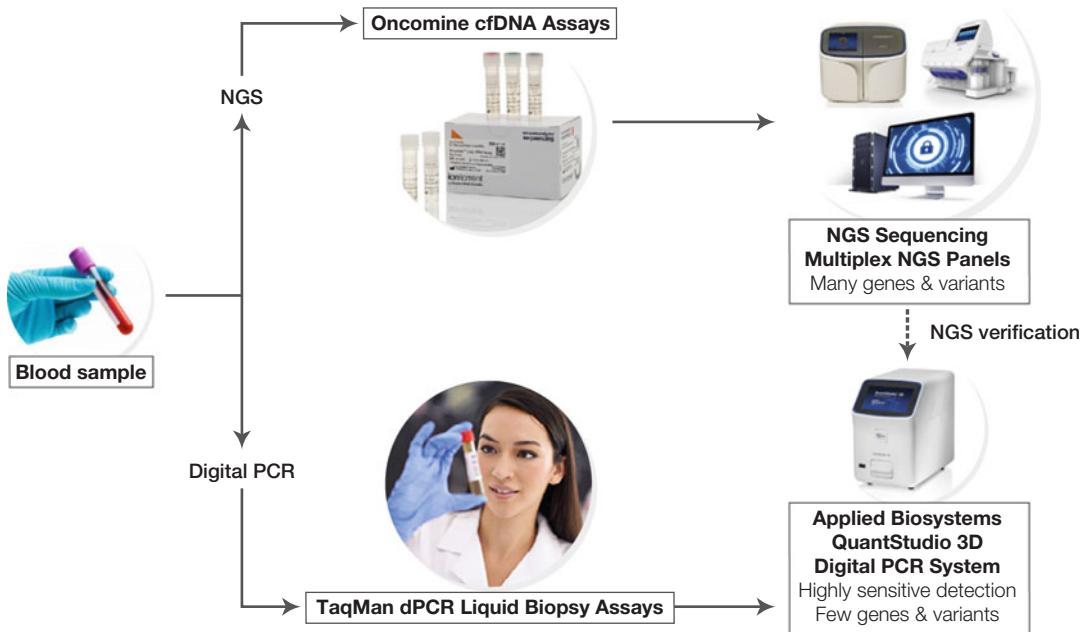


Fig. 8 Complete solution from mutation discovery by Next Generation Sequencing with Oncomine assays to variant confirmation and monitoring by digital PCR with TaqMan Liquid Biopsy dPCR assays

Table 2

The mutation rate of four samples was tested with Oncomine NGS and digital PCR using the QuantStudio 3D digital PCR system

Gene	Mutation	Oncomine cfDNA lung assay	Digital PCR	Sample type
BRAF	V600E	0.20%	0.13%	AcroMetrixOncology hotspot control
EGFR	p.E746_A750delELREA (exon 19 del)	0.12%	0.05%	Horizon cfDNA reference standard
KRAS	G12D	0.19%	0.13%	Horizon cfDNA reference standard
NRAS	G12D	0.07%	0.06%	AcroMetrixOncology hotspot control

We tested mutations in the BRAF, EGFR, KRAS, and NRAS genes. NGS and dPCR data show good agreement in the percentage mutation detected

2.1.3 Detection of Fusion Transcripts

Exchange of DNA segments between genes generates fusion genes, and the resulting proteins may lead to oncogenic changes which are causative for some cancers [11]. Roedel et al. [12] were interested in developing a detection method for one such fusion gene, the TMPRSS2:ERG fusion that occurs in more than 50% of prostate cancer cases [13]. The method was established using cDNA from

two prostate carcinoma cell lines, VCaP (heterozygous mutant for the *TMPRSS2:ERG* gene fusion) and LNCaP (homozygous wild-type). cDNA from the mutant cell line was spiked into wild-type cDNA to reflect varying levels of mutation rate.

A combination of one TaqMan gene expression assay and one TaqMan fusion assay was selected from several options for best coverage of the fusion transcript. The gene expression assay detected the wild-type transcript, the fusion assay detected the mutant transcript. Experimental conditions were optimized to yield the highest fluorescent signal and the best cluster separation. The mutation was detected with a limit of detection (LoD) of 0.05%.

2.1.4 Detection of Chimerism

Chimerisms can be monitored via sensitive PCR-based methods such as digital PCR. The capability of digital PCR to accurately detect small quantities of the recipient allele offers the advantage of earlier detection and potential for earlier intervention.

To illustrate this, Thermo Fisher Scientific collaborated with Dr. Antonio Jimenez-Velasco (Carlos Haya Hospital, Malaga Spain) to analyze leukemia samples post-bone marrow transplant. Assays discriminating between donor and recipient DNA were selected targeting domains where the bone marrow donors possessed a null allele of the locus of interest and the bone marrow recipient's samples possessed an insertion allele in the same locus, facilitating discrimination. Blood samples were collected from the donors and recipients before transplant and at various times after transplant. The presence of the two loci in the samples was quantified by dPCR.

Figure 9 shows typical results. Before the bone marrow transplant (Recipient Pre-SCT) the genotype of the circulating cells is 100% recipient, as expected (Fig. 9a, left panel). In the donor sample (Fig. 9a, right panel), approximately 0.1% recipient allele is detected in this assay. After stem cell transplant, although levels of the recipient allele remain low for 101 days, (Fig. 9b), recurrence of the recipient allele is apparent at 118 days posttransplant. This indicates that the recipient's own cells are repopulating the bone marrow (relapse). Chimerism was not detectable even after 192 days using the conventional PCR diagnostic method (data not shown). As dPCR detected recurrence more readily than the standard PCR method in this study, this highlights the potential of dPCR as a more sensitive alternative to traditional PCR for detecting rare events.

2.2 Generation of References and Standards

Generation of reference standards of known copies per μL is needed for a variety of applications. Accurate quantification of reference standards is essential since all follow-up measurements are based on the standards. This is especially important in the field of metrology.

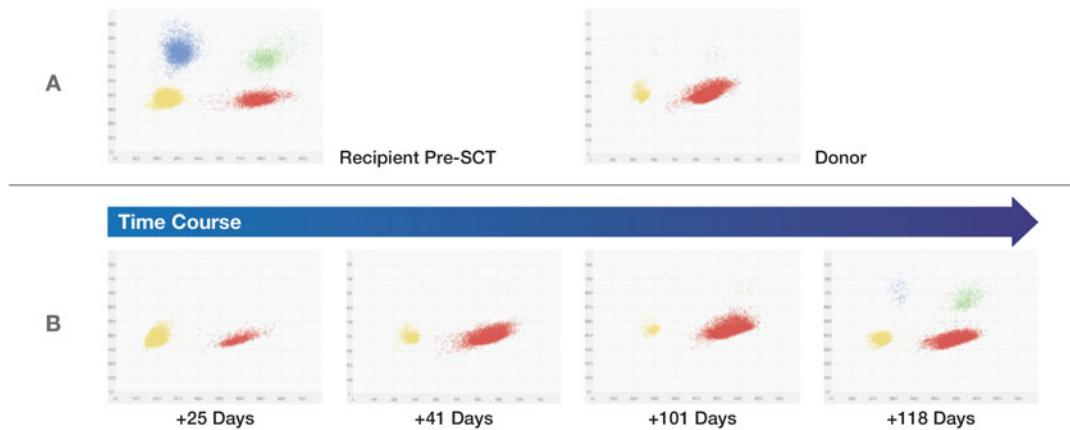


Fig. 9 Detection of chimerism. (a) Before the bone marrow transplant (Recipient Pre-SCT) the genotype of the circulating cells is 100% recipient (left). The donor's null allele is not detected (right). (b) A time course after stem cell transplant shows that the recipient's cells are repopulating the bone marrow at 118 days posttransplant

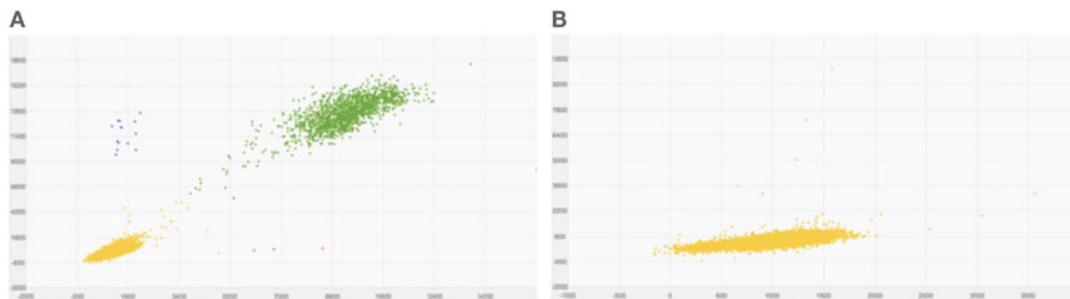


Fig. 10 Reference standard quantification by digital PCR. This standard is used as a control in qPCR experiments with infectious disease assay panels. 1:10 serial dilutions were generated of the plasmid control; three dilutions were tested in triplicates as well as a No template negative control (NTC). Digital PCR was performed using two plasmid-specific gene expression assays for high accuracy: one assay with a FAM label and a second assay with a VIC label. Positive, specific data points are expected in the green double positive cluster. The number of copies per μL was derived from AnalysisSuite Cloud Software and adjusted for dilution on the chip (not shown). (a) Plasmid control, (b) NTC

Digital PCR offers the advantage of absolute quantification. The exact number of copies per μL of a nucleic acid target on the chip can simply be counted. Digital PCR facilitates precise measurements and is ideally suited for quantification without a reference standard of

- (a) Target DNA molecules.
- (b) Reference standards for use in qPCR using a standard curve.

Absolute quantification using digital PCR ensures high quality reference standards (Fig. 10). In infectious disease research,

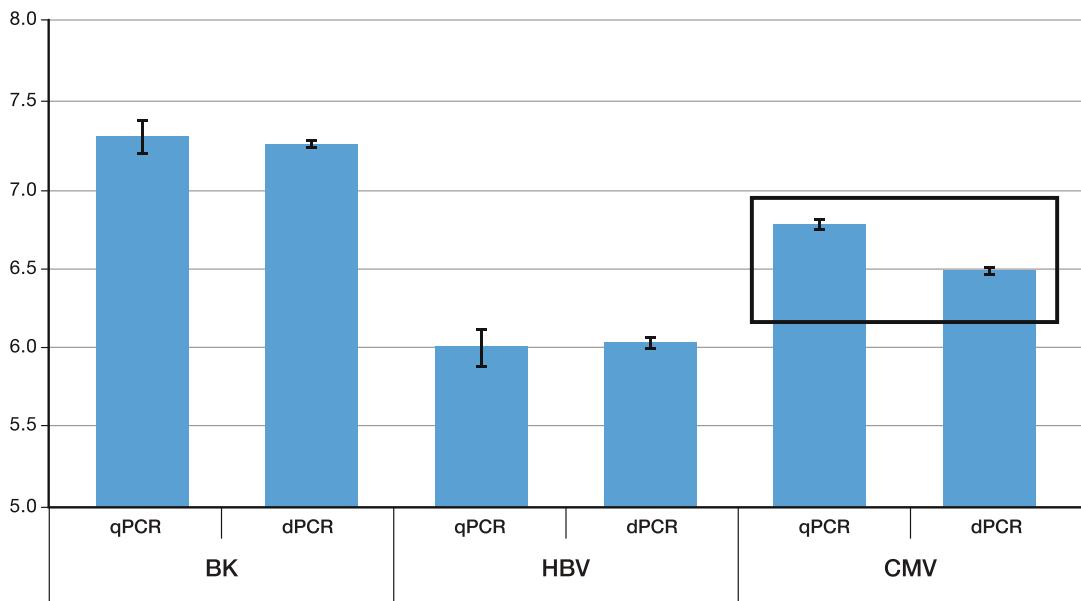


Fig. 11 Determination of viral titer using qPCR and digital PCR. BK is an internal standard. qPCR and digital PCR show good agreement for BK and HBV, with qPCR displaying a higher error rate. The CMV titer was assessed to be higher by qPCR than by digital PCR. The standard later proved to be degraded. It was likely quantified by UV spectrophotometric method. Short degraded DNA would mimic higher concentration of the standard, and the viral titer measurement was inaccurate as a result

calibrated synthetic and natural DNA controls are often used to functionally test assay panels that detect the presence and load of infectious organisms. These standards are used for qualification testing when establishing assays, and as in-line controls to be run alongside unknown samples in routine testing.

A comparison of viral titer assessment using qPCR and digital PCR is shown in Fig. 11. The result for qPCR demonstrates the caveat of depending on a standard curve: the effect of an inaccurately quantified reference standard for CMV. This standard was used in the qPCR experiment leading to a higher than actual viral titer for CMV. The standard later proved to be degraded and was likely quantified by UV spectrophotometry. Short degraded DNA would mimic higher concentration of the standard, and the viral titer measurement was inaccurate as a result.

Proving the point of high precision, we quantified several metrology reference standards. Digital PCR measurements were consistent with orthogonal methods but showed better precision. Results of these experiments are presented in Figs. 12 and 13.

2.3 GMO Detection and Monitoring

The most prominent example of genetically modified organisms (GMO) is genetically modified (GM) crops in agriculture. GM crops can be grown to be disease and herbicide resistant as well as to tolerate harsh growing conditions like dry and salty land and

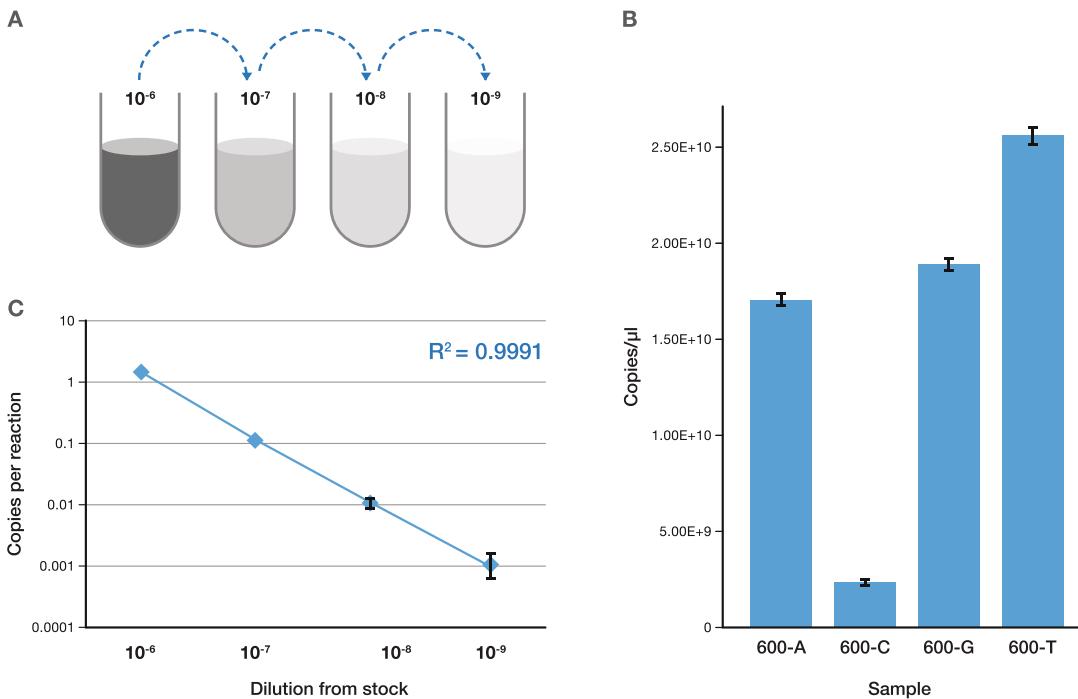


Fig. 12 Absolute quantification by digital PCR. Four certified DNA reference samples (National Metrology Institute of Japan, NMIJ CRM 6203-a) were quantified by digital PCR. Data shown represents two chips per sample **(a)** tenfold serial dilution of sample 660-T. Copies per reaction and dilution document excellent correlation **(b)**. Quantification of four samples shows tight error bars for all samples tested. **(c)** Same as **a**, showing dilution of sample vs. copies per reaction. Correlation is excellent (0.9991)

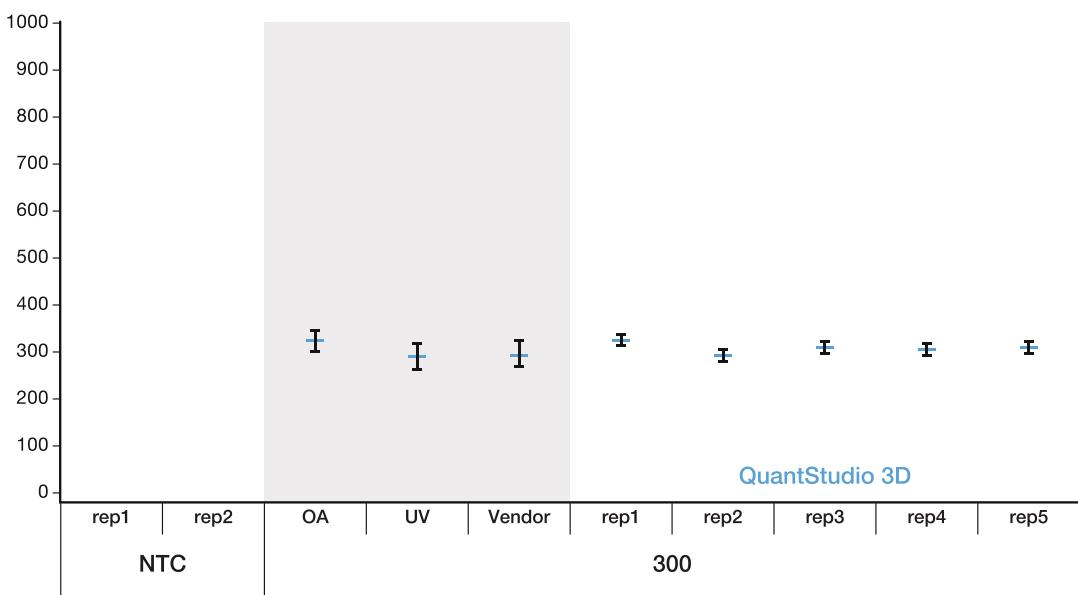


Fig. 13 Orthogonal measurements of GLA metrology standards. Digital PCR measurements are concordant with three orthogonal quantification methods: TaqMan OpenArray, stated vendor concentration, and UV spectrophotometric methods. Digital PCR shows higher precision with tighter error bars. Y-axis is copies per μL

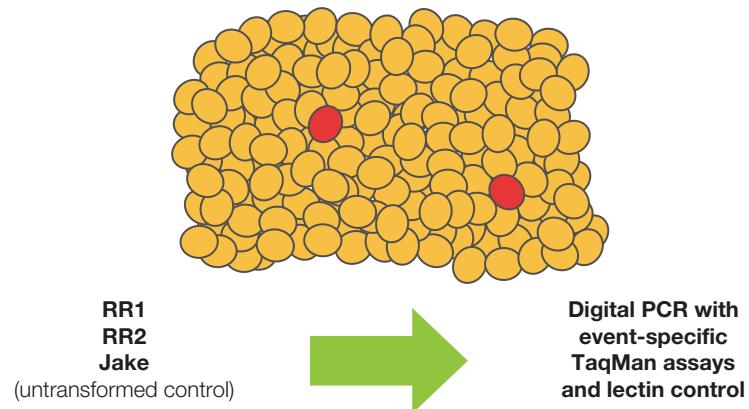


Fig. 14 GMO Detection. RR1 and RR2 transgenic (red) and Jake wild-type soybeans. Transgenic events were detected by digital PCR using custom Taq-Man assays with the QuantStudio 3D digital PCR system

high or low temperatures. They can also be modified to produce higher yield [14]. GM crops could help alleviate the growing food supply needs for an ever-growing human population [15].

At the same time, concerns persist and some countries including the European Union (EU) have implemented strict regulations [16]. Implementing these regulations necessitates development of monitoring systems which can both detect the presence of GM material and quantify the percentage of this material in a sample. Digital PCR using assays targeting known GMO transgenes has been shown to be an effective system for detection and absolute quantification of these targets (Fig. 14). Wan et al. [16] conducted a study analyzing the RR1 and RR2 transgenes used in GM Roundup® Ready soybeans using the QuantStudio 3D digital PCR system.

Custom TaqMan assays were designed targeting RR1 and RR2 transgenic events. Assays targeting the Lectin gene were used as internal control. The conventional nontransgenic soybean (Jake cultivar) was used as wild-type negative control. The amount of the RR transgene measured in copies/ μ L was comparable to the amount of the Lectin internal control. No RR transgene was detected in the Jake wild-type control. The Limit of Detection (LoD) in these experiments was 1%.

2.4 Copy Number Variation

Copy number variation can be determined with high accuracy using digital PCR (Fig. 15). Copy number variation (CNV) is the result of duplications or deletions of a specific locus in the genome. Copy number variation has implications in disease susceptibility and resistance, inherited disease and cancer [17, 18].

When using digital PCR to detect copy number, the first step is to establish an absolute count of copies of both the target gene and

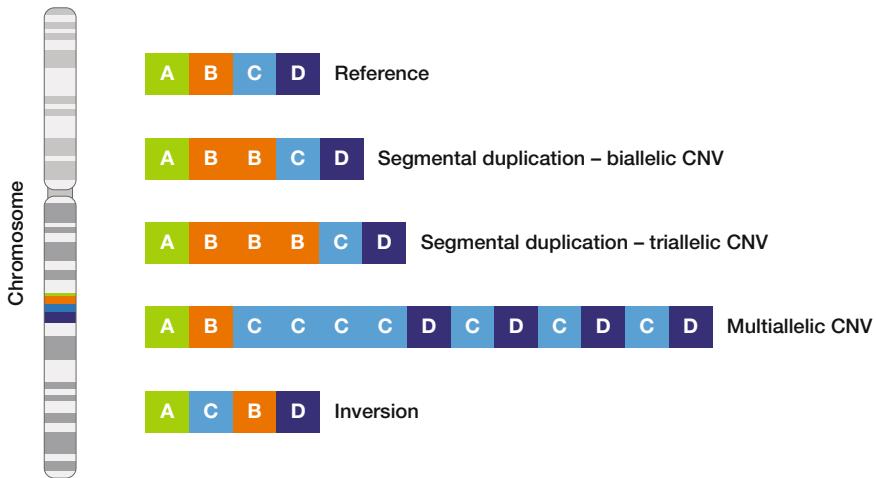


Fig. 15 Structural variation caused by duplication and inversion

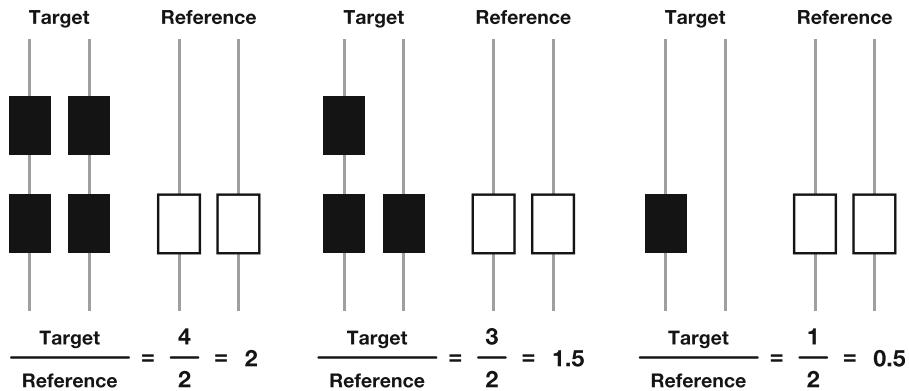


Fig. 16 Abnormal ratios of target to reference locus based on deletion of the target locus

a reference gene with known copy number. The reference gene is typically a diploid locus like RNase P. Two assays are needed, one assay detecting the target locus and one assay detecting the reference locus. The copy number of the target gene is the ratio of the absolute count of copies of the target locus to the absolute count of copies of the reference locus. A ratio above one indicates a duplication, while a ratio of less than 1 indicates a deletion.

Copies target/copies reference = 1	Wild-type
Copies target/copies reference >1	Amplification
Copies target/copies reference <1	Deletion

Some abnormal ratios and their causes are shown in Fig. 16.

Note: When the target is an X-chromosome linked locus, consider the hemizygosity status of male samples. The pseudoautosomal region is an exception

If tandem copies are suspected, the individual repeat units should be separated by restriction digest. If left undigested, the tandem copies are likely to be partitioned into the same well on the digital chip and would be undercounted as only one copy [22].

The ratio of copy numbers of the target locus to the reference locus is not always an integer. In the case of a heterogeneous tumor sample with 90% normal cells carrying 2 copies of the target locus and 10% mutant cells carrying 10 copies of the target locus, the resulting ratio will be 2.8:

90% normal cells	(0.9×2)	=1.8
10% tumor cells	(0.1×10)	=1.0
Total ratio		=2.8

Detection of copy number variation can be achieved by qPCR or by digital PCR. A limitation of qPCR is the reduction in measurement precision the further the measurement is from a simple doubling or loss. Measurements may not be sufficiently precise for determining high absolute copy number or for determining very small copy number differences (Fig. 17). qPCR is not recommended to measure copy numbers above four since the log scale compresses higher copy number differences. By contrast, digital PCR is a technology capable of highly precise measurements to differentiate subtle changes in copy number. It is suitable for detection of fractional copy numbers in heterogeneous sample types like tumor tissue as well as for detection of higher copy numbers above four.

An example of the importance of copy number is the determination of Her-2 status in breast cancer. 15–20% of breast cancers show amplification of Her-2. Two methods have been approved for determining Her-2 status in breast cancer: Immunohistochemistry (IHC) and In Situ Hybridization (ISH). Both methods are time consuming and expensive. ASCO-CAP guidelines and testing algorithms have been developed for analysis for these two methods.

A study performed by the Royal Surrey County Hospital, Guildford, United Kingdom investigated the possibility of substituting digital PCR for the established IHC and ISH methods to achieve time- and cost-saving [19]. No ASCO-CAP guidelines have been established for digital PCR. The study aimed at comparing digital PCR results to the established methods and identifying copy number ranges for Her-2 status in breast cancer as determined by digital PCR. Analyzing samples with all three methods showed excellent concordance of digital PCR results with IHC and ISH results, with digital PCR showing higher precision for fractional

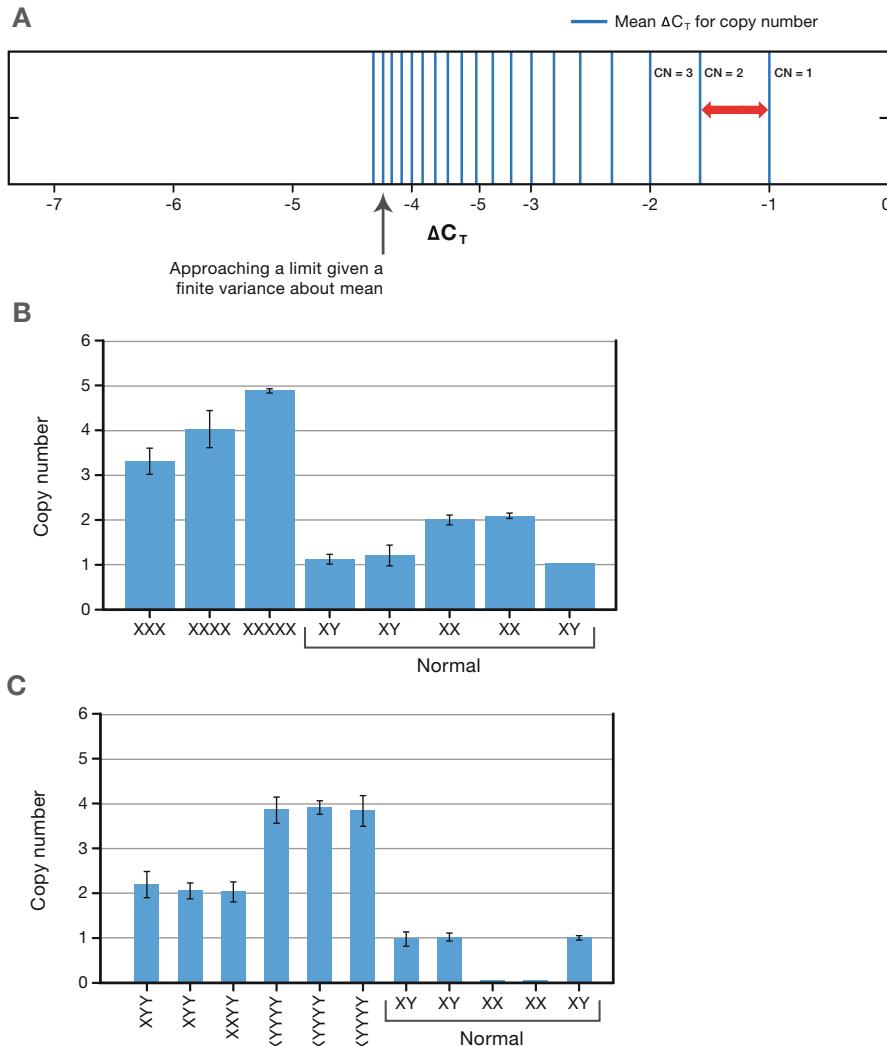


Fig. 17 (a) Compression of higher copy number differences in qPCR due to the log scale. (b, c) Examples of copy number variations in the range of 4–5 copies, which is the maximum distinguishable in qPCR

copy numbers than ICH and ISH. The study suggested employing digital PCR as the sole testing method when Her-2 status falls within either the negative or the positive range (Fig. 18). This enables tremendous cost and time saving. Additional testing with IHC and ISH is still recommended when test results fall in the equivocal (indeterminate) range.

2.5 NGS Library Quantification

Library quantification is an important part of successful template preparation for Next Generation Sequencing (NGS). This section describes library quantification for Ion Torrent and Illumina NGS libraries using digital PCR with the QuantStudio 3D Digital PCR System. NGS libraries are prepared by fragmenting DNA and

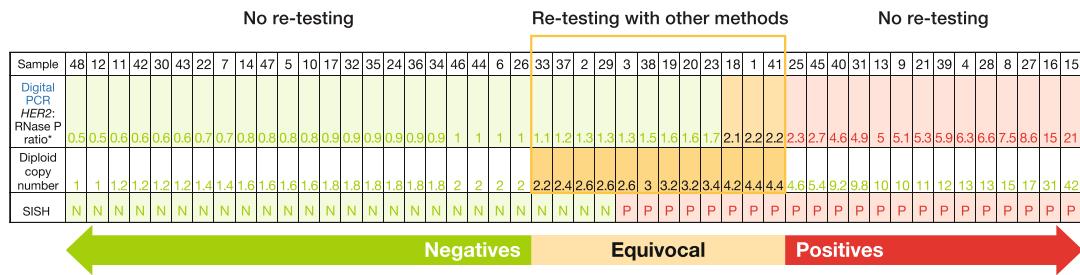


Fig. 18 A study performed by the Royal Surrey County Hospital, Guildford, United Kingdom. Excellent concordance was found for Her-2 status determined by digital PCR and with IHC and ISH methods. Digital PCR may be used as the sole testing method when Her-2 status falls within either the negative or the positive range. Additional testing with IHC and ISH is recommended when test results fall in the equivocal (indeterminate) range

ligating adapters for PCR. Library quantification by digital PCR utilizes TaqMan assays with primers and probes complementary to adapter sequences. Therefore, this approach will quantify only amplifiable (sequenceable) DNA fragments that have both adaptors incorporated. Absolute quantification results in accurately measured library input amounts.

2.5.1 Quantification of Ion Torrent Libraries

Optimal sequencing yield on the Ion PGM™ and Ion Proton™ platforms is highly dependent on the amount of library input. Library concentrations that are too high or too low lead to reduced total reads and reduce the overall throughput of the system. An accurate method for quantifying high-quality libraries is critical to maximizing output from a sequencing run.

Digital PCR using a prevalidated assay for quantification of Ion Torrent libraries offers a precise quantification method and is perfectly suited for this application. Once a high-quality library has been constructed, the QuantStudio 3D Digital PCR System is used to quantify the number of sequenceable molecules (Fig. 19). Based upon this result, the library is diluted to a level consistent with template bead preparation recommendations using the Ion One-Touch 2 or Ion Chef Systems. This is followed by sequencing on the Ion PGM, S5 or Ion Proton System.

TaqMan gene expression assay for Ion Torrent NGS Libraries:

Library	Assay ID
Ion Torrent, all libraries	Ac04347676_a1

Assay Ac04347676_al works for all Ion Torrent library sizes and types: Ion Fragment, Ion AmpliSeq™ Exome, Ion AmpliSeq™ Comprehensive Cancer Panel, Ion AmpliSeq™ Cancer HotSpot Panel v2, Whole Transcriptome, and small RNA. A detailed protocol is provided in Publication CO020253 0216 [20].

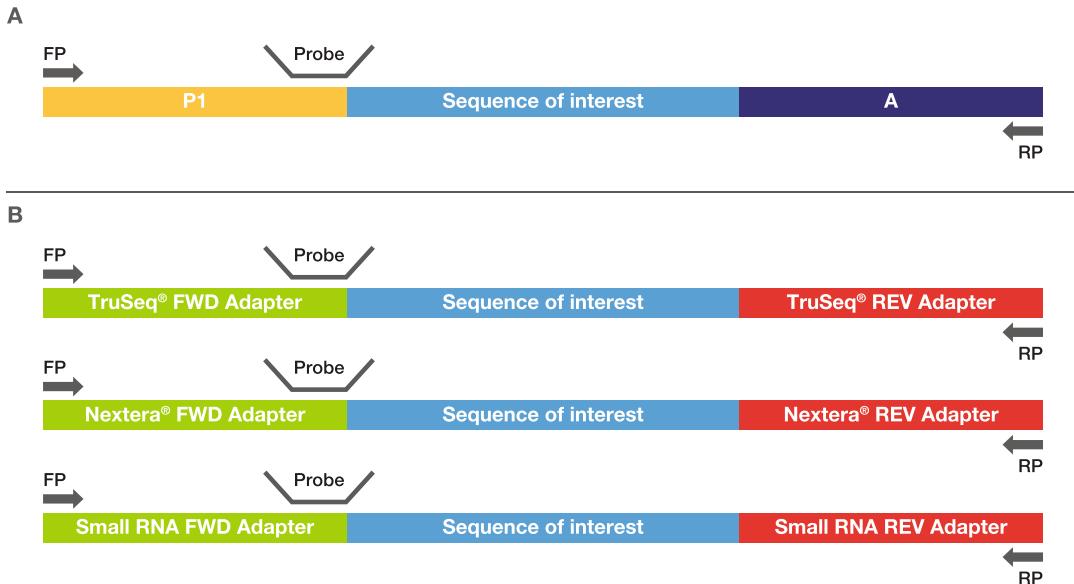


Fig. 19 NGS library quantification. TaqMan assays are designed to span the DNA fragment. Primers and probes are complementary to adapter sequences. Only amplifiable (sequenceable) fragments will be quantified. **(a)** Quantification of Ion Torrent libraries. **(b)** Quantification of three Illumina libraries

2.5.2 Quantification of Illumina Libraries

Illumina sequencing technology uses solid-phase bridge PCR to generate clonal sequence clusters. The cluster generation step is critical to obtaining maximal sequence data quality and yields and is dependent on the library input amount (Fig. 20). A precise method for quantification is needed. Digital PCR with the QuantStudio® 3D Digital PCR System may be used to directly measure the number of molecules that contain both library adapters and can be sequenced.

Based on this quantification, the appropriate library concentration is used to achieve optimal cluster density. Following clonal amplification, sequencing is performed on an Illumina® MiSeq®, HiSeq®, or NextSeq™ 500 platform. TaqMan® assays for Illumina® library quantification are designed to span both the P5 (forward primer, FP) and P7(reverse primer, RP) adapters.

TaqMan gene expression assay for Illumina NGS libraries:

Library	Assay ID
TruSeq® DNA/RNA library quantification	Ac04364396_a1
Nextera® library quantification	Ac04347696_a1
SmallRNA library quantification	Ac04347697_a1

A detailed protocol is provided in publication CO35377 1014 [21].

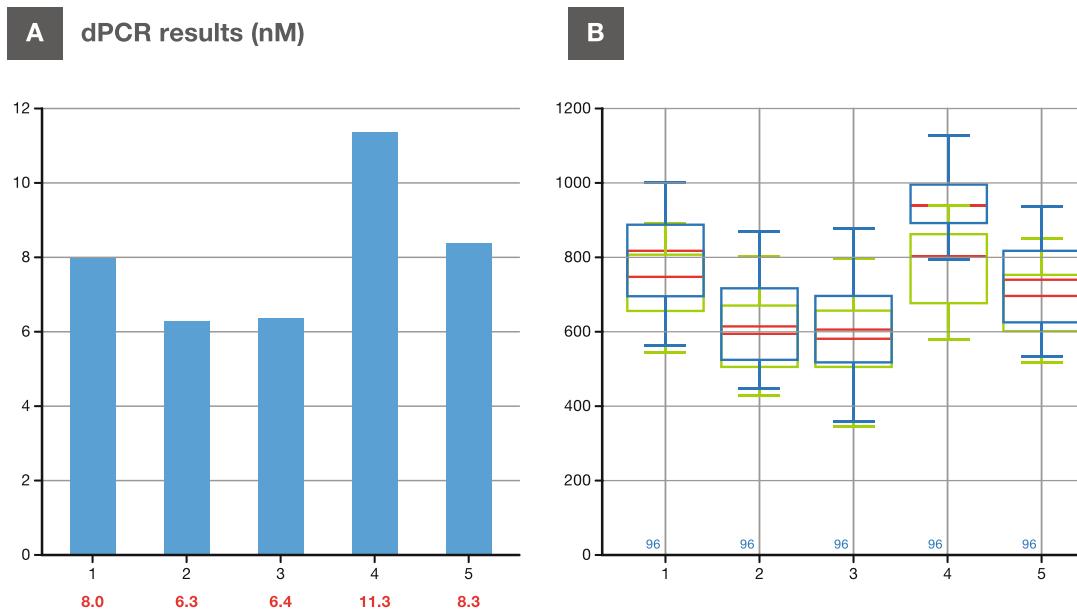


Fig. 20 Correlation of QuantStudio 3D Digital PCR data with Illumina® cluster density data. **(a)** Graph depicting the concentration (nM) for five Illumina TruSeq RNA libraries as determined by the QuantStudio 3D digital PCR system. The expected concentration for all five libraries was 10 nM. Libraries 2 and 3 had lower concentrations at 6.3 nM and 6.4 nM and library 4 had the highest concentration at 11.3 nM. **(b)** Cluster densities for the same libraries, generated by the Illumina HiSeq 2500 system. Libraries 2 and 3 had the lowest cluster densities and library 4 had the highest cluster density, which correlates well with the digital PCR concentrations obtained. Blue boxes: total clusters, green boxes: pass filter clusters

2.6 Low Level Pathogen Detection

Detection of low-level pathogens is essential when bacterial or viral pathogens cause disease through contaminated food or water. In these situations, it is important to detect pathogens as early as possible and in the presence of potential PCR inhibitors in crude sample. While real-time PCR (qPCR) can be a useful tool for pathogen detection in general, it is sensitive to PCR inhibitors often present in crude samples common to food and water testing.

By contrast, digital PCR is less sensitive to inhibitors and offers a more robust testing platform (<https://www.thermofisher.com/us/en/home/life-science/pcr/digital-pcr/pathogen-detection.html>). In addition, digital PCR provides an absolute count of pathogenic sequences, eliminating the need for reference standards.

2.7 Low-Fold Differential Gene Expression

Detection of differential gene expression by qPCR is generally limited to detecting changes of twofold or more. For some studies, detection of expression changes less than twofold may be required. Furthermore, it is often necessary to express differential gene expression with respect to a reference gene, such as a housekeeping gene like beta actin.

With the ability to achieve highly precise measurements of $\pm 10\%$ or better, digital PCR is capable of resolving changes of

twofold or less (<https://www.thermofisher.com/us/en/home/life-science/pcr/digital-pcr/differential-gene-expression.html>). In addition, the ability of digital PCR to determine absolute quantification of a transcript obviates the need for a reference gene.

Acknowledgments

For Research Use Only. Not for use in diagnostic procedures. We thank David Keys and Ross Stolsmark for thoughtful review of the manuscript.

References

- Vogelstein B, Kinzler KW (1999) Digital PCR. Proc Natl Acad Sci U S A 96:9236–9241
- Simmonds P, Balfe 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 64:864–872
- Majumdar N, Wessel T, Marks J (2015) Digital PCR modeling for maximal sensitivity, dynamic range and measurement precision. PLoS One 10:1–17
- Majumdar N, Banerjee S, Pallas M et al (2017) Poisson plus quantification for digital PCR systems. Sci Rep 7:2045–2322
- Dietrich D, Uhl B, Sailer V et al (2013) Improved PCR performance using template DNA from formalin-fixed and paraffin-embedded tissues by overcoming PCR inhibition. PLoS One 8(10):e77771. eCollection 2013
- Publication COL04004 0417 (2017) Rare mutation analysis using the QuantStudio 3D Digital PCR system. Quick Reference Protocol. Thermo Fisher Scientific
- Huang FW, Hodis E, Xu MJ et al (2013) Highly recurrent TERT promoter mutations in human melanoma. Science 339:957–959
- Bardelli A, Pantel K (2017) Liquid biopsies, what we do not know (yet). Cancer Cell 31:172–179
- Crowley E, Di Nicolantonio F, Loupakis F et al (2013) Liquid biopsy: monitoring cancer genetics in the blood. Nat Rev Clin Oncol 10:472–484
- Garcia-Saenz JA, Ayillon P, Laig M et al (2017) Tumor burden monitoring using cell-free tumor DNA could be limited by tumor heterogeneity in advanced breast cancer and should be evaluated together with radiographic imaging. BMC Cancer 17:210
- Yates LR, Campbell PJ (2012) Evolution of the cancer genome. Nat Rev Genet 13:795–806
- Roedel A, Laig M, Schmitz A (2016) Detection of the TMPRSS2:ERG fusion transcript. Application Note, Thermo Fisher Scientific, Publication COL31191 0616
- Tomlins SA, Laxman B, Varambally S et al (2008) Role of the TMPRSS2-ERG gene fusion in prostate cancer. Neoplasia 10:177–188
- Friedman D (2013) Are engineered foods evil? Sci Am 309:80–85
- Gerry C (2015) Feeding the world one genetically modified tomato at a time: a scientific perspective. Blog, special edition on GMOs. <http://sitn.hms.harvard.edu/flash/2015/feeding-the-world/>
- Wan JR, Song L, Wu YL et al (2016) Application of digital PCR in the analysis of transgenic soybean plants. Adv Biosci Biotechnol 7:403–417
- Almal SH, Padh H (2012) Implications of gene copy number variation in health and diseases. J Hum Genet 57:6–13
- Inaki K, Liu ET (2012) Structural mutations in cancer: mechanistic and functional insights. Trends Genet 28:550–559
- Publication CO09771 0514 (2014) Copy number variation in breast cancer translational research. Publication CO09771 0514, Thermofisher.com
- Publication CO020253 0216 (2016). Precise quantification of Ion Torrent libraries on the QuantStudio 3D Digital PCR system. Application Note, Thermofisher.com
- Publication CO35377 1014 (2014). Precise quantification of Illumina® libraries on the QuantStudio 3D Digital PCR System. Application Note, Thermofisher.com
- Publication CO020251 0216 (2016) Copy number variation analysis using the QuantStudio 3D Digital PCR System. Thermofisher.com

INDEX

A

- Absolute quantification 24, 70, 80, 99, 100, 212, 221, 223, 224, 228, 231
Accuracy 3, 76, 97, 105, 120, 195, 212, 224
Acute lymphoblastic leukemia (ALL) 153, 155
Age 177, 191–197
Amplicon yield 7, 21
Annealing temperature (T_a) 6, 11, 12, 14, 182
Apple 205
Ascomycota 80

B

- Barcodes 29, 80, 121
Basidiomycota 80
BCR-ABL1 153–162, 164, 166–170
BestKeeper 52, 129, 187
Beta vulgaris ssp. *maritima* 205
Bioanalyzer 25, 26, 28, 29, 51, 116, 136, 144
Bone marrow transplantation (BMT) 154, 155, 220, 221
Breast cancer 217, 218, 226

C

- Chimerisms 220, 221
Chlorella vulgaris extract 205
Chronic myeloid leukemia (CML) 153–172
Circulating tumor cells (CTCs) 139–150, 193, 217
Coefficient of variation (CV) 17, 31, 117, 182, 187, 188
Comparative Ct method
 ΔC_t method/d C_t 186, 187, 195
 $\Delta\Delta C_t$ method/dd C_t 186
Complementary DNA (cDNA) 2, 14, 16, 24, 29, 31, 32, 35, 36, 40, 42, 45–47, 49, 66, 69, 70, 73, 74, 108, 111–113, 115–117, 124, 125, 127, 128, 132, 133, 135, 142, 146, 156, 158, 160, 163, 164, 166, 168, 170, 176, 179, 182, 186, 188, 201, 202, 219
Confidence interval (CI) 17–20
Copy number variation (CNV) 224, 226, 227
C-Terminal Binding Protein 1 (CTBP1) 182, 184
Cytotoxic T-lymphocytes (CTL) 56

D

- Digital PCR (dPCR) 209–231
Diplodia sapinea 95–101

E

- Efficiencies 2, 6, 7, 15–17, 19, 36, 40, 43, 45, 50, 51, 53, 60, 62, 70, 75, 106, 113, 117, 126–128, 130–133, 135, 164, 167, 169, 175, 176, 185, 186
18S rRNA genes 18, 81, 82, 90
Escherichia coli 83, 85, 86, 130, 133
Ethanol 25, 50, 68, 74, 75, 97–99, 107, 110, 121, 123, 131, 142–144, 147, 148, 185, 202
Eukaryotic Translation Initiation Factor 2B Subunit Alpha (EIF2B1) 182, 184
External quality assessment (EQA) 3, 141

F

- FAM™ dye 40, 42, 57, 135, 200, 214, 215, 218, 221
Fastq 29
Ficoll gradient 58
Fluorescence resonance energy transfer (FRET) 40
Formalin-fixed paraffin-embedded (FFPE) 177–179, 213, 215
Forward primer (FP) 40, 41, 43, 99, 100, 124, 148, 229
Fungal 6, 80–83, 85–86, 89, 90, 95–104
Fusion transcripts 156, 215, 219, 220

G

- GC clamp 14
Genetically modified organisms (GMO) 215, 221, 224
GeNorm 31, 33, 34, 52, 112, 113, 116, 129, 187
Gliomas
 astrocytomas 175
 glioblastomas 175
 GBM 175, 180
 mixed oligoastrocytomas 175
 oligodendrogiomas 175

H

- HIV-1 56
 Human leucocyte antigen (HLA) 55, 59
 HLA-A 55, 56
 HLA-B 55, 56
 HLA-C 55–63

I

- In silico analysis 5
 Internal amplification control (IAC) 67, 71, 73, 75
 Internal transcribed spacer (ITS)
 ITS1 region 80
 ITS2 region 80–82
 Isopropanol 25, 50, 98, 99

J

- Jatropha curcas* 205

K

- Killer immunoglobulin-like receptors (KIRs) 55

L

- Leukocyte receptor complex (LRC) 55
 Limit of detection (LoD) 18, 19, 72,
 76, 211–212, 220, 224
 Limit of quantification (LoQ) 17, 19–21, 72
 Liquid biopsy 209, 213, 215–219
 Locked nucleic acid (LNA) 7, 26, 32,
 35, 43, 44, 148
 Lower limit of quantification (LLOQ) 76
 Low-fold differential gene expression 230, 231
 Low level pathogen detection 230

M

- Mastermix 2, 5, 8, 10–12, 14–18,
 57, 58, 60, 62, 160, 161, 164, 166, 168, 170
 Melanoma cells 140, 141, 143, 146–149
 Melt curves 8–10, 12–17, 158
 Melting temperature (Tm) 2, 13,
 35, 40, 42, 112, 116, 131, 132, 158, 182
 Messenger RNA (mRNA) 11, 119,
 120, 135, 140–143, 146, 155, 163, 170, 179,
 186, 188
 MicroRNA (miRNA) 23–37,
 39–53, 56, 60, 179, 199, 206
 miR 41, 47, 49
 pre-miRNAs 39
 pri-miRNAs 39
 Minimum Information for Publication of Quantitative
 Real-Time PCR Experiments
 (MIQE) guidelines 2, 21, 24,
 27, 43, 53, 105, 120, 171, 172, 182

Minor groove binder (MGB) 40, 42,

167, 171, 214

Mitochondrial Ribosomal Protein L9

(MRPL19) 182, 184

Multiplex 26, 28, 29, 31,
 43, 67, 71, 72, 154–158, 164–166, 171

N

- Natural killer (NK) cells 55
 Next-generation sequencing (NGS) 23,
 29–30, 33, 40, 215, 218, 219, 227–229
 NGS library quantification
 Illumina libraries 227, 229
 Ion Torrent libraries 227–229
 Nonfluorescent quencher (NFQ) 40, 42
 No-reverse transcriptase (noRT) 106,
 111, 113, 116, 117, 125
 NormFinder 31, 33, 34,
 52, 128, 129, 132, 135, 187
 No template controls (NTCs) 7, 8, 11,
 13–15, 32, 71, 72, 76, 99, 101, 113, 116, 125,
 136, 164, 166, 168–170, 221
 NS5 region 67, 69, 71–73

O

- Oligo(dT) 35, 43–45, 179

P

- Passive reference dye
 carboxy-X-rhodamine (CXR) reference dye 182
 ROX 37, 57, 167, 182, 211
 PCR efficiencies 2, 6, 7, 19, 106, 117, 126
 Peptide binding groove 55
 Peripheral blood mononuclear cells (PBMC) 57
 Philadelphia positive (Ph+) 153, 154
 Phosphate buffered saline (PBS) 178
 Phred score
 quality score 29
 PIK3CA oncogene 217, 218
Pinus nigra 96–98, 100, 101
 Poisson error 18
 Poisson model 211, 212
 Precision 58, 76, 127,
 186, 211–213, 222, 223, 226
 Primer design
 annealing times 7

 dimer formation 7

 GC content 7, 112

 in silico analysis 5

 mispriming 7

 no complementarity between primers

 optimization 7

 optimization 2, 5–7, 9, 11, 70, 73, 171

- primer combination 5, 7, 8, 14, 21
 primer concentration 7, 9, 10, 12, 116
 primer dimer 7, 13, 21, 36, 116, 185
 primer specificity 9, 11, 37, 112, 182
P
 Prostate carcinoma
 secondary structures 6, 7, 21, 131, 164
 TMPRSS2:ERG fusion 219
- Q**
 Quantification cycle (C_q) 7, 8, 11,
 12, 14, 15, 17, 18, 21, 32–36, 53, 89, 92, 111,
 113, 126–128, 131, 132, 134, 135, 167, 169,
 172, 213
- QuantStudioTM 12K Flex system
 OpenArray[®] plate 199–207
- R**
 Rare mutation detection 215, 217
 ReffFinder 129, 187
 Regression coefficient of linearity
 r² values 16, 19
 regression line 16, 127, 196
- Reverse primer (RP) 7, 35, 40,
 43, 44, 62, 99–101, 108, 124, 148, 229
- Reverse transcription quantitative real-time polymerase
 chain reaction (RT-qPCR) 24,
 26–28, 31–36, 39–53, 119–136, 140–144,
 146, 149, 188
- RNA integrity number (RIN) 28, 51, 136, 144
 RNA isolation 25, 28, 31,
 50–52, 105, 121, 123
- RNA quantification and quality control
 Agilent 2100 Bioanalyzer 25, 26, 28, 116
 NanoDrop ND-1000 25, 28
- RNA sequencing (RNA-Seq) 23–37
- RNases 27, 46, 47, 49, 51, 66,
 67, 69, 71, 72, 74, 89, 108, 114, 119, 120,
 123, 131, 142, 144, 158, 163, 178, 225
- Robustness 11, 12, 14, 200
- rRNA operons 81
- S**
Scenedesmus quadricauda extract 205
 Signal joint T-cell receptor rearrangement excision circle
 (sjTREC) 192, 193, 195, 196
- Single nucleotide polymorphism (SNP) 21,
 56, 63, 199, 200, 202–206, 213, 215
- Singleplex 67, 71, 72, 162, 167
- SNP genotyping 199, 200, 213, 215
- Sugar beet 203–205
- Sulfate nutrition 205
- SYBR[®] green dye 40, 43, 44, 108, 135
- T**
 TaqMan[®] 40–43, 45–47,
 49–51, 53, 67, 69–71, 73–75, 97–100, 102,
 110, 193–195, 200–202, 206, 210, 213,
 215–220, 223, 224, 228, 229
- T-cell receptors (TCRs) 192
- 3'-end 7, 9, 14, 15, 40, 42, 45, 62, 171, 179
- 3'-Untranslated region (3' UTR) 56, 179
- Threshold cycle (C_t) 2, 60,
 70–72, 92, 100–102, 172, 182, 185–188, 194,
 195
- 260/230 nm ratio 28, 110, 193, 202
- 260/280 nm ratio 28, 85, 131, 193, 202
- Tyrosinase assay 140–143, 146
- Tyrosine kinase inhibitor (TKI) therapy
 bosutinib 154
 dasatinib 154
 imatinib 154
 nilotinib 154
 ponatinib 154
- U**
 Upper limit of quantification (ULOQ) 76
- V**
 Variant confirmation 218, 219
 VIC[™] dye 40, 57, 67,
 71, 74, 102, 167, 193, 200, 214, 215, 218,
 221
- Viral load 66, 72
- W**
 Wastewater treatment plants (WWTPs) 79, 81, 82, 90
- Y**
 Yellow fever virus (YFV) 65–76