

Research and Perspectives in Neurosciences

Rudolf Jaenisch
Feng Zhang
Fred Gage *Editors*

Genome Editing in Neurosciences

FONDATION
IPSEN
POUR LA RECHERCHE
THERAPEUTIQUE

OPEN



Springer

Research and Perspectives in Neurosciences

More information about this series at <http://www.springer.com/series/2357>

Rudolf Jaenisch • Feng Zhang • Fred Gage
Editors

Genome Editing in Neurosciences

 Springer

OPEN

Editors

Rudolf Jaenisch
Whitehead Institute and
Department of Biology
Massachusetts Institute of Technology
Cambridge, MA
USA

Feng Zhang
Department of Brain and Cognitive Science
Broad Institute of MIT and Harvard
Cambridge, MA
USA

Fred Gage
Laboratory of Genetics
Salk Institute for Biological Studies
La Jolla, CA
USA

Fondation IPSEN
Boulogne-Billancourt
France

Acknowledgement: The editors wish to express their gratitude to Mrs. Mary Lynn Gage for her editorial assistance.



ISSN 0945-6082 ISSN 2196-3096 (electronic)
Research and Perspectives in Neurosciences
ISBN 978-3-319-60191-5 ISBN 978-3-319-60192-2 (eBook)
DOI 10.1007/978-3-319-60192-2

Library of Congress Control Number: 2017948865

© The Editor(s) (if applicable) and The Author(s) 2017. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

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

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

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

It was somewhat of a surprise to the scientific community when, in 1944, Oswald Avery definitively proved that DNA encoded the blueprint to life. Many scientists at the time thought that, with just four bases, DNA was chemically too simple to contain so much information. Nearly 75 years later, though, we are still trying to parse all the information contained in a genome. This work has been greatly accelerated in the past decade by two parallel advancements: next-generation DNA sequencing technology and genome editing methods. Current sequencing capacity is leading to the generation of large amounts of genetic data, while our ability to manipulate the genome is rapidly advancing our understanding of that genetic data.

Genome editing based on the microbial CRISPR-Cas adaptive immune system has emerged in recent years as a powerful tool for dissecting genetic circuits. CRISPR-associated enzymes such as Cas9 and Cpf1 are RNA-guided DNA endonucleases that can be precisely targeted to nearly any region of the genome via the guide RNA sequence. These enzymes have been used for both gene disruption and insertion in a wide range of organisms, and they have also been developed as a platform for gene activation, providing another way to modulate gene expression patterns. Finally, RNA-guided nucleases can facilitate both loss- and gain-of-function genome-wide screening applications. This technology has significantly advanced our ability to perform forward genetics in mammalian systems, model human diseases in tractable systems, and interrogate complex genetic processes. Moreover, it has the potential to revolutionize the way we treat human disease.

The Fondation IPSEN *Colloque Médecine et Recherche* in the Neuroscience Series, held in Paris on April 22, 2016, highlighted how genome editing is enabling breakthroughs in how we study the brain and how we may be able to apply this powerful method to understand and treat central nervous system (CNS) disorders. The use of CRISPR-Cas-based technologies was a common thread that ran throughout the meeting: it was used to either develop new cell lines relevant to studying the CNS or it made it possible to use new model organisms to study the CNS; it

powered large-scale interrogation of neuronal genetic circuits; and it was used for proof-of-principle therapeutic restoration of disease-causing mutations.

In contrast to Avery's discovery, nobody has ever doubted the complexity of the human brain. Neuroscientists have struggled for decades with seemingly intractable questions about the nature of the brain, and CNS disorders have proven to be some of the most difficult human diseases to study, in large part because the tools simply were not available. Genome editing, along with other recent technological advances such as next-generation sequencing advances and optogenetics, is unlocking hundreds of new ways to study the brain. The work that is described in this volume exemplifies the lines of research that can now be pursued and offers a tantalizing glimpse of where this work will lead us.

MA, USA
June 2017

Feng Zhang

Contents

In Vitro Modeling of Complex Neurological Diseases	1
Frank Soldner and Rudolf Jaenisch	
Aquatic Model Organisms in Neurosciences: The Genome-Editing Revolution	21
Jean-Stéphane Joly	
Genome-Wide Genetic Screening in the Mammalian CNS	31
Mary H. Wertz and Myriam Heiman	
CRISPR/Cas9-Mediated Knockin and Knockout in Zebrafish	41
Shahad Albadri, Flavia De Santis, Vincenzo Di Donato, and Filippo Del Bene	
Dissecting the Role of Synaptic Proteins with CRISPR	51
Salvatore Incontro, Cedric S. Asensio, and Roger A. Nicoll	
Recurrently Breaking Genes in Neural Progenitors: Potential Roles of DNA Breaks in Neuronal Function, Degeneration and Cancer	63
Frederick W. Alt, Pei-Chi Wei, and Bjoern Schwer	
Neuroscience Research Using Non-human Primate Models and Genome Editing	73
Noriyuki Kishi and Hideyuki Okano	
Multiscale Genome Engineering: Genome-Wide Screens and Targeted Approaches	83
Neville E. Sanjana	

Using Genome Engineering to Understand Huntington's Disease 87

Barbara Bailus, Ningzhe Zhang, and Lisa M. Ellerby

Therapeutic Gene Editing in Muscles and Muscle Stem Cells 103

Mohammadsharif Tabebordbar, Jason Cheng, and Amy J. Wagers

List of Contributors

Shahad Albadri Institut Curie, PSL Research University, INSERM, Paris, France

Frederick W. Alt Howard Hughes Medical Institute, Boston, MA, USA

Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA, USA

Department of Genetics, Harvard Medical School, Boston, MA, USA

Cedric S. Asensio Department of Biological Sciences, University of Denver, Denver, CO, USA

Barbara Bailus Buck Institute for Research on Aging, Novato, CA, USA

Jason Cheng Department of Stem Cell and Regenerative Biology, Harvard University and Harvard Stem Cell Institute, Cambridge, MA, USA

Flavia De Santis Institut Curie, PSL Research University, INSERM, Paris, France

Filippo Del Bene Institut Curie, PSL Research University, INSERM, Paris, France

Vincenzo Di Donato Institut Curie, PSL Research University, INSERM, Paris, France

Lisa M. Ellerby Buck Institute for Research on Aging, Novato, CA, USA

Fred Gage Salk Institute for Biological Studies, La Jolla, CA, USA

Myriam Heiman MIT Department of Brain and Cognitive Sciences, Cambridge, MA, USA

Picower Institute for Learning and Memory, Cambridge, MA, USA

Broad Institute of MIT and Harvard, Cambridge, MA, USA

Salvatore Incontro Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA, USA

Rudolf Jaenisch The Whitehead Institute for Biomedical Research, Cambridge, MA, USA

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA

Jean-Stéphane Joly INRA CASBAH Group, Neuro-Paris Saclay Institute, CNRS, Gif-sur-Yvette, France

Noriyuki Kishi Laboratory for Marmoset Neural Architecture, RIKEN Brain Science Institute, Wako-shi, Saitama, Japan

Department of Physiology, Keio University School of Medicine, Shinjuku-ku, Tokyo, Japan

Roger A. Nicoll Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA, USA

Hideyuki Okano Laboratory for Marmoset Neural Architecture, RIKEN Brain Science Institute, Wako-shi, Saitama, Japan

Department of Physiology, Keio University School of Medicine, Shinjuku-ku, Tokyo, Japan

Neville E. Sanjana New York Genome Center, New York, NY, USA

Department of Biology, New York University, New York, NY, USA

Bjoern Schwer Howard Hughes Medical Institute, Boston, MA, USA

Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA, USA

Department of Genetics, Harvard Medical School, Boston, MA, USA

Department of Neurological Surgery and Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, University of California, San Francisco, CA 94158, USA

Frank Soldner The Whitehead Institute for Biomedical Research, 455 Main Street, Cambridge, MA 02142, USA

Mohammadsharif Tabebordbar Department of Stem Cell and Regenerative Biology, Harvard University and Harvard Stem Cell Institute, Cambridge, MA, USA

Amy J. Wagers Department of Stem Cell and Regenerative Biology, Harvard University and Harvard Stem Cell Institute, Cambridge, MA, USA

Pei-Chi Wei Howard Hughes Medical Institute, Boston, MA, USA

Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA, USA

Department of Genetics, Harvard Medical School, Boston, MA, USA

Mary H. Wertz Picower Institute for Learning and Memory, Cambridge, MA, USA

Broad Institute of MIT and Harvard, Cambridge, MA, USA

Feng Zhang Broad Institute, McGovern Institute, MIT, Cambridge, MA, USA

Ningzhe Zhang Buck Institute for Research on Aging, Novato, CA, USA

In Vitro Modeling of Complex Neurological Diseases

Frank Soldner and Rudolf Jaenisch

Abstract A major reason for the lack of effective therapeutics and a deep biological understanding of complex diseases, which are thought to result from a complex interaction between genetic and environmental risk factors, is the paucity of relevant experimental models. This review describes a novel experimental approach that allows the study of the functional effects of disease-associated risk in complex disease by combining genome wide association studies (GWAS) and genome-scale epigenetic data to prioritize disease-associated risk variants with efficient gene editing technologies in human pluripotent stem cells (hPSCs). As a proof of principle, we recently used such a genetically precisely controlled experimental system to identify a common Parkinson's disease-associated risk variant in a non-coding distal enhancer element that alters the binding of transcription factors and regulates the expression of α -synuclein (*SNCA*), a key gene implicated in the pathogenesis of Parkinson's disease.

Introduction

One of the main challenges to understanding the onset and progression of human disease is to develop effective model systems that combine known genetic elements with disease-associated phenotypic readouts. The identification of genes linked to familial forms of diseases such as cystic fibrosis, sickle cell anemia or monogenetic forms of neurodegenerative disorders has fundamentally changed our understanding of many diseases and provided vital clues into the underlying pathogenesis (Botstein and Risch 2003; Altshuler et al. 2008; McClellan and King 2010).

F. Soldner

The Whitehead Institute for Biomedical Research, 455 Main Street, Cambridge, MA 02142, USA

R. Jaenisch (✉)

The Whitehead Institute for Biomedical Research, 455 Main Street, Cambridge, MA 02142, USA

Department of Biology, Massachusetts Institute of Technology, 31 Ames Street, Cambridge, MA 02139, USA

e-mail: jaenisch@wi.mit.edu

© The Author(s) 2017

R. Jaenisch et al. (eds.), *Genome Editing in Neurosciences*, Research and Perspectives in Neurosciences, DOI 10.1007/978-3-319-60192-2_1

1

Detailed knowledge of disease-causing mutations and genes allows the establishment of reliable and disease-relevant cellular and animal models and facilitates the systematic analysis of molecular and cellular disease mechanisms and the development and validation of novel and effective therapeutic approaches.

In contrast to such predominantly rare and monogenic disorders, the majority of the most common medical conditions, such as obesity, heart disease, diabetes, autoimmune disease or sporadic neurodegenerative disease, have no well-defined genetic etiology and do not follow Mendelian inheritance patterns. Population genetics suggest that such sporadic or polygenic diseases result from a complex interaction between multiple genetic and non-genetic, lifestyle and environmental risk factors (Botstein and Risch 2003; Altshuler et al. 2008). The complexity and our limited knowledge of the underlying genetic component have largely prevented the generation of genetically defined disease models. The paucity of disease-relevant experimental systems represents one of the major reasons for our limited biological understanding of complex diseases and an almost complete lack of disease-modifying effective therapeutics.

In the following, we will summarize recent progress in genetics and developmental and molecular biology, which may provide a solution for generating disease-relevant in vitro models for complex disease. By combining human pluripotent stem cell (hPSC)-technology with genome editing and genome-scale epigenetic and genome-wide association studies (GWAS) data to identify disease-associated risk variants, we will provide a blueprint to create genetically defined experimental model systems that allow the functional analysis of disease-associated risk variants. As a proof of principle, we describe how we applied this approach to sporadic Parkinson's disease and identified a common risk variant in a non-coding distal enhancer element that regulates the expression of *SNCA*, a key gene implicated in the pathogenesis of Parkinson's disease (Soldner et al. 2016).

Induced Pluripotent Stem Cells to Model Complex Diseases

The ability to reprogram somatic cells into human induced pluripotent stem cells (hiPSCs) has opened the intriguing possibility of studying complex human disease in a cell culture dish (Takahashi and Yamanaka 2006; Takahashi et al. 2007; Yu et al. 2007). Following in vitro differentiation, patient-derived hiPSCs provide access to large amounts of human disease-relevant cells that carry all the genetic alterations involved in disease development (Saha and Jaenisch 2009; Soldner and Jaenisch 2012; Takahashi and Yamanaka 2013; Yu et al. 2013). Without precise knowledge of the underlying genetics, such patient-derived cells, therefore, allow the generation of relevant cellular model systems based on disease-associated genetic elements. This approach has already been used to model a range of primarily monogenic diseases, including neurodegenerative diseases such as Alzheimer's disease, Parkinson's disease and amyotrophic lateral sclerosis (ALS; Cooper et al. 2012; Israel et al. 2012; Reinhardt et al. 2013; Alami et al. 2014; Wainger et al. 2014; Young et al. 2015). Despite the unprecedented potential and excitement of this approach, it became

apparent that individual hiPSC lines, independent of disease status or genotype, displayed highly variable biological properties in vitro, such as the propensity to differentiate into functional cell types (Bock et al. 2011; Boulting et al. 2011; Soldner and Jaenisch 2012; Nishizawa et al. 2016). This observation significantly limits their value to identify robust disease-associated phenotypes by simply comparing patient-derived cells with unrelated controls. This system-immanent variability has proven to be particularly challenging in the context of age-related diseases including neurodegenerative diseases such as Alzheimer's and Parkinson's disease, considering that disease-associated phenotypes typically progress slowly over many years in patients, which suggests that expected in vitro phenotypes would be rather mild and subtle. The reasons for the observed cell-to-cell differences include genetic background variations, genetic and epigenetic changes resulting from reprogramming and extended maintenance of hiPSCs and the lack of robust in vitro differentiation protocols (Soldner and Jaenisch 2012; Liang and Zhang 2013).

Some of the above-described limitations have been overcome by improved reprogramming and culture conditions (Warren et al. 2010; Hou et al. 2013), directed differentiation approaches including transcription factor-induced reprogramming (Zhang et al. 2013), insertion of cell type-specific fluorescent marker proteins to monitor differentiation (Di Giorgio et al. 2008; Hockemeyer et al. 2009, 2011; Chambers et al. 2012; Mica et al. 2013) or by consortium-size experiments to significantly increase the number of independent experimental samples (The HD iPSC Consortium 2012). However, variable genetic backgrounds between patient-derived and control cells remain an unresolved major limitation of the current hiPSC approach, due to the well-established influence of uncharacterized genetic modifiers on disease development and progression in patients and, accordingly, on disease-associated phenotypes in vitro.

Gene Editing to Generate Genetically Controlled Disease Models

The recent progress in gene editing technologies by using engineered nucleases such as meganucleases, zinc finger nucleases (ZFNs), transcription activator-like effector-based nucleases (TALEN) and the CRISPR/Cas9 system is thought to provide an elegant solution to control for differences in genetic background (Soldner et al. 2011; Soldner and Jaenisch 2012; Hockemeyer and Jaenisch 2016). In particular, the simplicity and ease of the CRISPR/Cas9 system to efficiently modify the genome in human cells, even at multiple loci simultaneously, allow us to engineer genetically controlled hPSC lines that differ only at known genetic disease-causing variants (Jinek et al. 2012, 2013; Cong et al. 2013; Mali et al. 2013).

As a proof of principle, we recently used ZFNs to either seamlessly correct Parkinson's disease-associated mutations in the *SNCA* gene in patient-derived hiPSCs or to insert similar variants into wild-type human embryonic stem cells (hESCs; Soldner et al. 2011). Such isogenic pairs of hPSC lines provided an

experimental system with a controlled genetic background in which the engineered disease-associated risk variants were the only experimental variables. Analyzing disease-associated phenotypes in this genetically controlled system allowed identification of nitrosative stress, accumulation of endoplasmic reticulum (ER)-associated degradation substrates, and ER stress as early Parkinson's disease-associated pathological phenotypes (Chung et al. 2013). A further study revealed that nitrosative and oxidative stress result in S-nitrosilation of the transcription factor MEF2C and inhibition of the MEF2C-PGC1 α transcriptional network contributing to mitochondrial dysfunction and apoptotic neuronal cell death (Ryan et al. 2013). By combining this monogenic disease model with disease-associated environmental stressors, the experiments further provide new mechanistic insight into gene-environmental (GxE) interaction in the pathogenesis of Parkinson's disease (Ryan et al. 2013). Notably, both studies relying on a genetically controlled in vivo model identified novel therapeutic targets and small molecules that reversed the observed pathological phenotypes in neurons, which are currently perused as novel therapeutics for Parkinson's disease (Chung et al. 2013; Ryan et al. 2013). The above-described approach clearly overcomes many of the limitations of the current hiPSC technology. Due to the simplicity of the CRISPR/Cas9 system to efficiently edit the genome in hiPSCs, the use of isogenic cell lines is becoming the gold standard for analyzing disease-associated phenotypes in vitro (Reinhardt et al. 2013; Kiskinis et al. 2014; Paquet et al. 2016). However, such an approach seems currently limited to monogenetic diseases in which the disease-causing genetic alterations are well established and the expected disease-associated phenotypes display robust and highly penetrant effects.

Functional Role of GWAS-Identified Risk Variants in Complex Disease

Translating the concept of engineering genetically controlled model systems to complex disease seems daunting and will require a detailed understanding of the underlying genetic component. GWAS and genome-scale next generation sequencing (NGS) approaches have significantly advanced our understanding of the genetic basis of complex disease. GWAS in particular have identified numerous common single-nucleotide polymorphisms (SNPs) associated with human traits and diseases, pinpointing the genomic loci and genes thought to play important roles in the pathophysiology of the respective diseases (Botstein and Risch 2003; Altshuler et al. 2008; McClellan and King 2010).

However, the interpretation of this permanently increasing amount of data is limited by the fact that disease-associated SNPs only statistically correlate with the underlying disease and the vast majority of risk variants have no established biological relevance to disease or clinical utility for prognosis or treatment (Altshuler et al. 2008; McClellan and King 2010). Any SNP in linkage

disequilibrium (LD) with a GWAS-identified risk variant is equally likely to be causative for the risk to develop a specific disease. It has therefore been difficult to distinguish variants that are functional and disease-relevant from those that are in LD and thus only mark the underlying haplotype containing the functional variant. Advancing from genetic association to causal biologic processes has been challenging for two additional reasons. First, the majority of disease-associated genetic variants fall into the non-coding part of the genome, which impedes any functional analysis through simple transgenic overexpression or disruption in established cell lines or any analysis in non-human model systems due to the limited conservation of non-coding elements between species. Second, the prevailing hypothesis about the heritability of complex diseases suggests that multiple common or potentially rare SNPs cooperatively contribute to the risk of developing a specific disease; however, each individual risk variant will have only a small or at most medium-size additive or multiplicative effect on disease phenotypes (Gibson 2012). Indeed, disease-associated genetic variants are also prevalent in the healthy population, although with lower frequency, and the majority of carriers of risk SNPs do not develop a disease, implying that individual risk variants are not sufficient to cause disease-associated phenotypes. Consequently, only very few risk variants have been functionally linked to specific diseases, such as a common polymorphism at the 1q13 locus, which alters the expression of the SORT1 gene and is correlated with both plasma low-density lipoprotein cholesterol (LDL-C) and myocardial infarction (Musunuru et al. 2010).

Under the assumption that specific risk haplotypes contribute through dysregulation of the same molecular pathways to disease risk, a current approach suggests that we stratify patient-derived hiPSCs according to specific genetic risk variants rather than according to disease status. This approach may be sufficient in some cases to reduce the genetic heterogeneity based on known disease haplotypes and to reveal previously masked disease-associated phenotypes. Indeed, this approach was successfully used to dissect the function of a common Alzheimer's disease-associated non-coding genetic variant in the 5' region of the SORL1 (sortilin related receptor 1; Young et al. 2015). However, the main limitation of this approach remains the uncontrolled effect of additional genetic modifiers and the inability to identify the specific causative sequence variant that is required for further functional analysis.

Epigenomic Signatures to Prioritize GWAS-Identified Risk Variants

Cis-acting effects of genetic variants on gene expression have been proposed to be a major factor for phenotypic variation of complex traits and disease susceptibility (Schadt et al. 2003; Morley et al. 2004; Cheung et al. 2005, 2010; Lee and Young 2013; GTEx Consortium 2015). The widespread availability of cell- and

tissue-specific transcriptome-wide expression data along with the corresponding genotyping data has greatly facilitated the identification of expression quantitative trait loci (eQTLs; GTEx Consortium 2015). Although able to detect statistical correlation between specific risk variants and gene expression, this approach entails limitations that are comparable to traditional GWAS in identifying the functional risk variants. Recent genome-scale epigenetic studies such as the ENCODE (ENCODE Project Consortium 2012) and Roadmap Epigenomics project (Roadmap Epigenomics Consortium 2015) have allowed us to reliably identify and catalogue regulatory elements in a cell type-, tissue- and in some cases disease-specific manner. These studies specifically have highlighted the enrichment of GWAS-identified risk variants in regulatory DNA elements specific to tissues and cell types (Ernst et al. 2011; Degner et al. 2012; Maurano et al. 2012; Hnisz et al. 2013; Trynka et al. 2013; Farh et al. 2014; Pasquali et al. 2014; Ripke et al. 2014) affected by the respective diseases. These results suggest that disease-associated risk variants may affect gene regulation by modifying the function of tissue-specific regulatory elements. In particular, distal enhancer elements that are bound by key transcription factors (TFs) and known to precisely control spatial and temporal gene expression during embryonic development and tissue homeostasis in a cell type-specific manner (Ward and Kellis 2012; Lee and Young 2013; Farh et al. 2014; Ripke et al. 2014; Wamstad et al. 2014) are found to be enriched for GWAS variants in many complex diseases.

A number of recent studies have correlated changes in TF binding in enhancer regions with sequence-specific, heritable changes in chromatin state and gene regulation (Kasowski et al. 2013; Kilpinen et al. 2013; McVicker et al. 2013), thus providing a molecular mechanism for how individual sequence variants contribute to the development of complex diseases. Recent progress in defining TF binding specificities using high throughput SELEX and chromatin immunoprecipitation sequencing (ChIP-seq) approaches has largely increased our understanding of sequence-specific TF binding in the genome and significantly improved our ability to analyze or predict TF binding on a genome-wide scale (Jolma et al. 2013, 2015). Based on the rapidly increasing availability of epigenetic data, mapping of GWAS-identified variants to TF binding sites within tissue-specific enhancer elements has been proposed as a valuable approach to prioritize and identify functional and disease-relevant risk variants (Ward and Kellis 2012; Rivera and Ren 2013; Claussnitzer et al. 2014; Wamstad et al. 2014). Indeed, such integration of GWAS with epigenetic signatures for heart-specific enhancers allowed for the identification of novel functional risk variants for cardiac phenotypes (Wang et al. 2016). Likewise, a similar approach identified an obesity-associated risk variant in the FTO locus, which alters early adipose differentiation by disrupting a TF binding site at a pre-adipocyte-specific enhancer (Claussnitzer et al. 2015).

The 3-dimensional (3D) organization of the genome is thought to contribute to the regulation of gene expression (Bickmore 2013; de Graaf and van Steensel 2013; de Laat and Duboule 2013). The recent development of chromosome conformation capture techniques ("3C" and genome-wide 3C-based methods;

Dekker et al. 2002, 2013) or cohesin chromatin interaction analysis by paired-end tag sequencing (ChIA-PET; Downen et al. 2014) allow us to determine long-range chromatin interactions such as cell type-specific promoter-enhancer interaction. These analyses suggest that active enhancer elements are bound by transcription factors and loop over long distances to contact target genes to regulate transcription. An emerging model suggests promoter-enhancer interactions typically only occur within megabase-sized topological-associated domains (TAD; Dixon et al. 2012; Nora et al. 2012), as defined by high DNA interaction frequency based on genome-wide chromosome capture data or within such TADs in insulated neighborhoods restricted by cohesin-associated CTCF-CTCF loops (Handoko et al. 2011; DeMare et al. 2013; Downen et al. 2014; Rao et al. 2014; Ji et al. 2016). Notably, there is mounting evidence that changes in 3D structure, potentially through sequence-specific disruption of CTCF interaction, might contribute to disease development (Ji et al. 2016). Integrating datasets of cell type-specific changes in enhancer-promoter interactions and information about the 3D structure of the genome will further help us to assign disease-associated risk variants in enhancer sequences to target genes and provide supporting evidence to identify functional disease-associated risk variants and deregulated target genes.

Functional Analysis of Parkinson's Disease-Associated Risk Variants

As a proof of principle, we describe below how we recently applied the above-elucidated approach to sporadic Parkinson's disease as a prototypical complex disorder, to identify common risk variants in non-coding distal enhancer elements that functionally modulate the risk to develop the disease (Soldner et al. 2016). Parkinson's disease is the second most common chronic progressive neurodegenerative disease, with a prevalence of more than 1% in the population over the age of 60. Although the discovery of genes linked to rare Mendelian forms of PD such as *SNCA*, *LRRK2*, *PARKIN*, *PINK1* and *DJ1* has provided insight into the molecular and cellular pathogenesis of the disease (Gasser et al. 2011; Singleton et al. 2013), the etiology leading to neuronal cell loss is largely unknown. Importantly, over 90% of Parkinson's cases do not show Mendelian inheritance patterns; however, substantial clustering of cases within families suggests that sporadic, late age of onset Parkinson's disease results from a complex interaction between genetic risk alleles and environmental factors. A recent GWAS meta-analysis has identified 26 genomic loci containing risk variants for sporadic Parkinson's disease (Nalls et al. 2014); however, as for the majority of neurodegenerative disorders, little mechanistic insight is available on how specific sequence variations contribute to disease development and progression.

Identification of Parkinson's Disease-Associated Risk Variants in Brain-Specific Enhancer Elements

A recent analysis of Histone H3 acetylated at lysine 27 (H3K27ac)-marked regions in the post-mortem adult brain suggests a significant enrichment of Parkinson's disease-associated risk SNPs within distal enhancer elements (Vermunt et al. 2014). This finding supports the hypothesis that sequence-specific changes in enhancer function and deregulated transcription of linked genes mediate the risk to develop the disease. A number of specific epigenetic modifications, such as p300, mono-methylation of Histone H3 at lysine 4 (H3K4me1), H3K27ac and DNase I hypersensitive sites (DHSs) have been established as surrogate marks to reliably identify candidate enhancer sequences (Visel et al. 2009, 2013; Creighton et al. 2010; Rada-Iglesias et al. 2011; Maurano et al. 2012). Thus, to identify specific candidate risk variants in distal enhancers, we intersected Parkinson's disease-associated risk SNPs (Nalls et al. 2014) with publicly available epigenetic data (Roadmap Epigenomics Consortium 2015). This analysis allowed us to compile a list of risk variants ranked by the overlap of active enhancer elements. Interestingly, many of the top-ranked risk variants were located to the *SNCA* locus. Because changes in TF binding are thought to be the major mediator of SNP-specific changes in gene expression (Kasowski et al. 2013; Kilpinen et al. 2013; McVicker et al. 2013) we incorporated this idea to further prioritize the risk variants in enhancers by analyzing predicted TF binding for known TF binding specificities comparing both alternative genotypes for each Parkinson's disease-associated SNP. This analysis highlighted the Parkinson's disease-associated SNP rs356168 in an enhancer in intron-4 of *SNCA* as the risk variant with the highest number of genotype-dependent differential TF binding in the *SNCA* locus. The functional relevance of this enhancer was further supported by chromosome conformation capture data, which indicate a physical interaction (looping) between the enhancer and the promoter region of *SNCA* that is thought to be necessary for the cis-acting effects on gene expression (Vermunt et al. 2014).

It is well established that *SNCA* plays a central role in the pathogenesis of Parkinson's disease. Point mutations in *SNCA* were the first genetic variants linked to familial forms of Parkinson's disease, and the *SNCA* protein is the major component of Lewy bodies and Lewy neuritis, which are considered the pathological hallmark of familial and sporadic Parkinson's disease (Gasser et al. 2011; Singleton et al. 2013). In addition, the *SNCA* locus represents one of the strongest Parkinson's disease-associated GWAS hits (Nalls et al. 2014). Notably, multiplication of the entire *SNCA* locus was identified as causal for a rare autosomal-dominant form of Parkinson's disease, indicating that a moderate increase of wild-type *SNCA* expression (1.5 times in the case of genomic duplications) is sufficient to cause an autosomal-dominant form of Parkinson's disease (Singleton et al. 2003; Miller et al. 2004; Devine et al. 2011; Kim et al. 2012). This observation is highly suggestive of a molecular mechanism by which risk variants in the *SNCA* locus modify the risk to develop Parkinson's disease by slightly modulating the expression

of *SNCA*. This clear link between *SNCA* expression and the development of Parkinson's disease in the context of genomic amplification therefore provides a good rationale for gene expression as a disease-relevant phenotypic readout to connect genetic variation to disease risk (Devine et al. 2011). Indeed, the first indication that the *SNCA* locus may contain risk alleles that modulate *SNCA* expression came from the identification of *SNCA-Rep1*, a complex polymorphic microsatellite repeat region approximately 10 kb upstream of the transcription start site. Multiple candidate gene association studies suggested that individuals who are homozygous for a shorter, "protective" repeat region (Rep1-257 or Rep1-259) have a significantly lower risk of developing Parkinson's disease compared to individuals carrying the longer forms (Rep1-261 or Rep1-263; Kruger et al. 1999; Maraganore et al. 2006). Several functional studies, including the analysis of transgenic mice carrying different human *SNCA-Rep1* alleles (Chiba-Falek et al. 2005; Cronin et al. 2009), suggested an "enhancer-like" function of the microsatellite repeat element based on the cis-regulatory correlation between the *SNCA-Rep1* repeat length and *SNCA* expression.

Allele-Specific Gene Expression as a Robust Read-Out to Analyze Cis-Regulatory Effects

As explained in detail above, one of the major limitations of using hPSC-derived somatic cells to model disease in vitro is the considerable variability of the biological properties between individual cell lines. As for *SNCA*, a gene known to be variable between neuronal cell types such as astrocytes, oligodendrocytes and neurons and to be regulated during development and terminal differentiation, cellular heterogeneity and incomplete maturation significantly interfere with the detection of subtle differences in gene expression between distinct risk-genotypes or patient compared to control cells, respectively. Indeed, individual in vitro differentiation experiments from genetically identical sub-clones resulted in up to fourfold differences in *SNCA* expression (Soldner et al. 2016). To address this problem, we recently described an experimental approach that is based on determining the effect of individual regulatory elements on the transcription of the cis-regulated gene by analyzing allele-specific gene expression (Soldner et al. 2016). The deletion of just a single copy (heterozygous) of a candidate regulatory element or its exchange with an alternative disease-associated element affects only the gene expression of the cis-regulated gene on the same allele while maintaining the expression of the other, homologous allele, unaltered. Consequently, allele-specific gene expression would be biased towards lower or higher expression of the cis-regulated allele depending on the introduced genetic modification. Because expression is measured as the *ratio* between two individual alleles in every cell, this analysis is expected to be largely independent of cell homogeneity and can be applied to heterogeneous cell populations. In this respect, the non-targeted *SNCA* allele allows for a simple normalization and serves as internal control across isogenic samples.

Functional Analysis of Parkinson's-Associated Risk Variants

To analyze allele-specific expression, we developed a robust, sensitive and highly quantitative reverse transcription polymerase chain reaction (qRT-PCR) assay based on the detection of a heterozygous SNP in the 3'UTR of *SNCA*. Using CRISPR/Cas9 genome editing, we generated an allelic series of isogenic cell lines by either heterozygous deletion of the entire microsatellite repeat region (thought to have the most pronounced effect on *SNCA* expression) or insertion of *SNCA-Rep1* elements with all of the repeat length alleles (*Rep1-257*, *Rep1-259*, *Rep1-261* and *Rep1-263*) that are present in the normal population. Using allele-specific expression as readout, we showed that neither the deletion of the microsatellite repeat *SNCA-Rep1* element nor its exchange for the shorter or longer repeat length risk alleles affected the cis-regulated expression of the linked *SNCA* allele, suggesting that this element has no clear role in *SNCA* regulation. This result conflicts with previous studies that supported an “enhancer-like” cis-regulatory effect of *SNCA-Rep1* on the expression of *SNCA*. It is possible that difficulties in controlling the experimental variables of the transgenic mouse (Cronin et al. 2009) or neuroblastoma cell system (Chiba-Falek et al. 2005) used in the functional analyses, species-specific differences of non-coding regulatory elements or the variability in analyzing human post-mortem brain tissue (Fuchs et al. 2008; Dumitriu et al. 2012) affected the validity of these conclusions. However, because in vitro differentiated cells allow only for the analysis of early events, due to the limited time in culture, we cannot completely exclude an effect of the *SNCA-Rep1* element at later time points or only in combination with additional environmental factors.

In contrast to *SNCA-Rep1*, the CRISPR/Cas9-mediated exchange of Parkinson's disease-associated alleles spanning an enhancer element in the fourth intron that carries two risk SNPs (rs356168 and rs3756054) showed a significant effect on allele-specific expression of *SNCA* (Fig. 1; Soldner et al. 2016). When the protective A-allele at SNP rs356168 was exchanged for the risk-associated G-allele, the expression of the cis-regulated *SNCA* allele was increased by 6–18%. In contrast, the exchange of the adjacent risk SNP rs3756054 showed no effect on allele-specific *SNCA* expression, suggesting that this variant only reaches genome-wide significance in GWAS because this variant is in LD with the functional risk-modifying SNP (Fig. 1). Given that a 1.5-fold increase in *SNCA* expression is sufficient to cause a familial autosomal-dominant form of Parkinson's disease, these data support the notion that a modest life-long increase of *SNCA* expression may represent the molecular cause of increased risk to develop Parkinson's disease of individuals carrying the G-allele at this risk variant. Moreover, an expression quantitative trait loci (eQTL) analysis of *SNCA* expression in post-mortem adult brain samples suggested that a similar sequence-specific modest increase in *SNCA* expression occurs within the human population, further substantiating a functional role of the risk variant rs356168 in Parkinson's disease (Soldner et al. 2016). This subtle effect on the expression of a disease-relevant gene is consistent with the hypothesis that small effect size of common genetic risk variants contributes to the heritability of sporadic diseases.

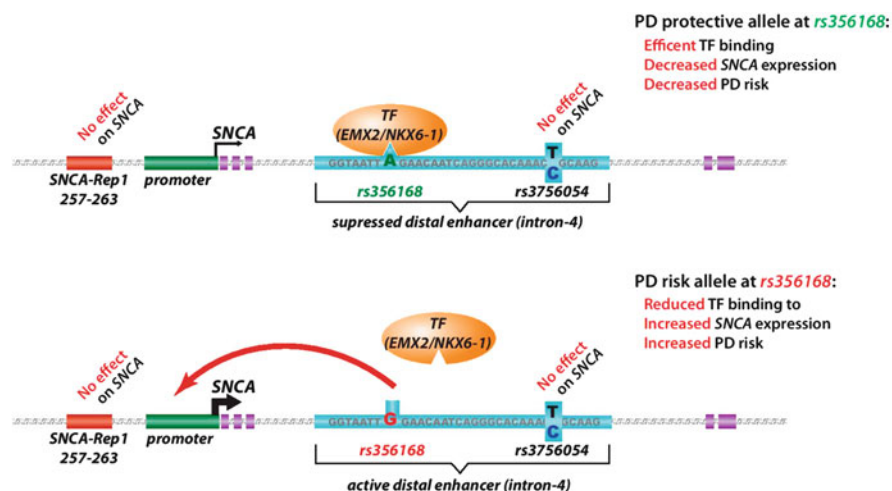


Fig. 1 Proposed model describing the effect of multiple Parkinson's disease (PD)-associated risk variants on *SNCA* expression (modified from Soldner et al. 2016). The schematic illustrates the genomic organization of the *SNCA* locus, including the PD-associated risk variants *SNCA-Rep1* and the risk SNPs rs356168 and rs3756045, both located in a distal enhancer element in the fourth intron of *SNCA*. The analysis described in Soldner et al. (2016) suggests that the brain-specific transcription factors (TF) EMX2 and NKX6-1 show sequence-dependent binding at rs356168 with preference for the A-allele. The efficient TF binding in carriers of the protective A-allele results in a suppressed distal enhancer element and, consequently, in reduced expression of *SNCA* associated with reduced risk to develop PD. In contrast, the reduced TF binding in carriers of the PD risk-associated G-allele at this variant leads to a more active distal enhancer, resulting in increased expression of *SNCA* associated with an increased risk to develop PD. Notably, neither the repeat length of *SNCA-Rep1* nor the PD-risk variant at rs3756054 significantly affects *SNCA* expression, suggesting that these elements are in linkage disequilibrium (LD) with other functional risk-modifying variants

To gain insight into the molecular basis of how risk variants affect target gene expression, we analyzed TF binding data and identified two brain-specific TFs, EMX2 and NKX6-1, that bind to the enhancer element at the risk variant. Further analysis for sequence-specific binding indicated that both TFs, EMX2 and NKX6-1 preferentially bind to the protective, lower *SNCA* expressing A-allele at rs356168 (Fig. 1). These results suggest a model in which the sequence-dependent binding of these TFs at a distal enhancer element represses enhancer activity and thus modulate *SNCA* expression. Indeed, ectopic overexpression of both TFs in neurons reduced *SNCA* expression (Soldner et al. 2016), consistent with previous data in mouse models demonstrating their role as repressors of enhancer function (Ligon 2003; Schisler et al. 2005; Schaffer et al. 2010; Mariani et al. 2012). Thus, our data provide a molecular link between GWAS-identified risk SNP-dependent changes in TF binding at a distal enhancer element, altered expression of *SNCA* and the risk to develop sporadic Parkinson's disease (Fig. 1). EMX2 and NKX6-1 may physically interact and function in a complex to suppress enhancer activity. However, expression analysis indicated that the two TFs are only expressed in a subset of neurons

and are primarily not co-expressed in the same cell, suggesting that they may function at the same enhancer element in different cell types. TF-specific usage of identical regulatory elements in distinct cell populations might be a possible explanation for the selective vulnerability of distinct neuronal populations, as observed in Parkinson's disease.

Mechanistic Study of Sporadic Diseases: Conclusions

As outlined in this review, a major challenge of modeling sporadic diseases in the culture dish is the system-immanent variability in differentiating hESCs or hiPSCs to functional cells. The variability is caused by genetic background differences between patient-derived hiPSCs and cells derived from control individuals as well as the inconsistency of most protocols to generate homogeneous cultures of differentiated cells. These issues complicate, if not exclude, the use of gene expression level as a valid functional readout to define the molecular mechanisms of candidate disease risk variant, which are expected to only subtly alter the transcription of the downstream gene. As our analysis of the *SNCA*-associated risk variants demonstrates, two experimental strategies allow us to overcome these limitations: (1) the use of CRISPR/Cas9-mediated gene editing for generating disease-relevant and control lines that differ exclusively at the risk variant and (2) the development of an allele-specific assay that allows the robust detection of small differences in disease risk-associated gene expression, an assay that is independent of cell heterogeneity and extent of differentiation.

Acknowledgments This work was supported by grants NS088538, MH104610 and HD045022 from the NIH.

References

- Alami NH, Smith RB, Carrasco MA, Williams LA, Winborn CS, Han SSW, Kiskinis E, Winborn B, Freibaum BD, Kanagaraj A, Clare AJ, Badders NM, Bilican B, Chaum E, Chandran S, Shaw CE, Eggan KC, Maniatis T, Taylor JP (2014) Axonal transport of TDP-43 mRNA granules is impaired by ALS-causing mutations. *Neuron* 81:536–543
- Altshuler D, Daly MJ, Lander ES (2008) Genetic mapping in human disease. *Science* 322:881–888
- Bickmore WA (2013) The spatial organization of the human genome. *Annu Rev Genomics Hum Genet* 14:67–84
- Bock C, Kiskinis E, Verstappen G, Gu H, Boulting G, Smith ZD, Ziller M, Croft GF, Amoroso MW, Oakley DH, Gnirke A, Eggan K, Meissner A (2011) Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* 144:439–452
- Botstein D, Risch N (2003) Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet* 33(Suppl):228–237

- Boulting GL, Kiskinis E, Croft GF, Amoroso MW, Oakley DH, Wainger BJ, Williams DJ, Kahler DJ, Yamaki M, Davidow L, Rodolfa CT, Dimos JT, Mikkilineni S, MacDermott AB, Woolf CJ, Henderson CE, Wichterle H, Eggan K (2011) A functionally characterized test set of human induced pluripotent stem cells. *Nat Biotechnol* 29:279–286
- Chambers SM, Qi Y, Mica Y, Lee G, Zhang X-J, Niu L, Bilsland J, Cao L, Stevens E, Whiting P, Shi S-H, Studer L (2012) Combined small-molecule inhibition accelerates developmental timing and converts human pluripotent stem cells into nociceptors. *Nat Biotechnol* 30:715–720
- Cheung VG, Spielman RS, Ewens KG, Weber TM, Morley M, Burdick JT (2005) Mapping determinants of human gene expression by regional and genome-wide association. *Nature* 437:1365–1369
- Cheung VG, Nayak RR, Wang IX, Elwyn S, Cousins SM, Morley M, Spielman RS (2010) Polymorphic cis- and trans-regulation of human gene expression. *PLoS Biol* 8:e1000480
- Chiba-Falek O, Kowalak JA, Smulson ME, Nussbaum RL (2005) Regulation of alpha-synuclein expression by poly (ADP ribose) polymerase-1 (PARP-1) binding to the NACP-Rep1 polymorphic site upstream of the SNCA gene. *Am J Hum Genet* 76:478–492
- Chung CY, Khurana V, Auluck PK, Tardiff DF, Mazzulli JR, Soldner F, Baru V, Lou Y, Freyzo Y, Cho S, Mungenast AE, Muffat J, Mitalipova M, Pluth MD, Jui NT, Schüle B, Lippard SJ, Tsai LH, Krainc D, Buchwald SL, Jaenisch R, Lindquist S (2013) Identification and rescue of α -synuclein toxicity in Parkinson patient-derived neurons. *Science* 342:983–987
- Claussnitzer M, Dankel SN, Klocke B, Grallert H, Glunk V, Berulava T, Lee H, Oskolkov N, Fadista J, Ehlers K, Wahl S, Hoffmann C, Qian K, Rönn T, Riess H, Müller-Nurasyid M, Bretschneider N, Schroeder T, Skurk T, Horsthemke B, DIAGRAM+Consortium, Spieler D, Klingenspor M, Seifert M, Kern MJ, Mejhert N, Dahlman I, Hansson O, Hauck SM, Blüher M, Amer P, Groop L, Illig T, Suhre K, Hsu YH, Mellgren G, Hauner H, Laumen H (2014) Leveraging cross-species transcription factor binding site patterns: from diabetes risk loci to disease mechanisms. *Cell* 156:343–358
- Claussnitzer M, Dankel SN, Kim K-H, Quon G, Meuleman W, Haugen C, Glunk V, Sousa IS, Beaudry JL, Puvion-Vandier V, Abdennur NA, Liu J, Svensson P-A, Hsu Y-H, Drucker DJ, Mellgren G, Hui C-C, Hauner H, Kellis M (2015) FTO obesity variant circuitry and adipocyte browning in humans. *N Engl J Med* 373:895–907
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823
- Cooper O, Seo H, Andrabi S, Guardia-Laguarta C, Graziotto J, Sundberg M, McLean JR, Carrillo-Reid L, Xie Z, Osborn T, Hargus G, Deleidi M, Lawson T, Bogetoft H, Perez-Torres E, Clark L, Moskowitz C, Mazzulli J, Chen L, Volpicelli-Daley L, Romero N, Jiang H, Uitti RJ, Huang Z, Opala G, Scarffe LA, Dawson VL, Klein C, Feng J, Ross OA, Trojanowski JQ, Lee VM, Marder K, Surmeier DJ, Wszolek ZK, Przedborski S, Krainc D, Dawson TM, Isacson O (2012) Pharmacological rescue of mitochondrial deficits in iPSC-derived neural cells from patients with familial Parkinson's disease. *Sci Transl Med* 4:141ra90–141ra90
- Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci USA* 107:21931–21936
- Cronin KD, Ge D, Manninger P, Linnertz C, Rossoshek A, Orrison BM, Bernard DJ, El-Agnaf OMA, Schlossmacher MG, Nussbaum RL, Chiba-Falek O (2009) Expansion of the Parkinson disease-associated SNCA-Rep1 allele upregulates human α -synuclein in transgenic mouse brain. *Hum Mol Genet* 18:3274–3285
- de Graaf CA, van Steensel B (2013) Chromatin organization: form to function. *Curr Opin Genet Dev* 23:185–190
- de Laat W, Duboule D (2013) Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* 502:499–506

- Degner JF, Pai AA, Pique-Regi R, Veyrieras J-B, Gaffney DJ, Pickrell JK, De Leon S, Michelini K, Lewellen N, Crawford GE, Stephens M, Gilad Y, Pritchard JK (2012) DNase I sensitivity QTLs are a major determinant of human expression variation. *Nature* 482:390–394
- Dekker J, Rippe K, Dekker M, Kleckner N (2002) Capturing chromosome conformation. *Science* 295:1306–1311
- Dekker J, Marti-Renom MA, Mirny LA (2013) Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet* 14:390–403
- DeMare LE, Leng J, Cotney J, Reilly SK, Yin J, Sarro R, Noonan JP (2013) The genomic landscape of cohesin-associated chromatin interactions. *Genome Res* 23:1224–1234
- Devine MJ, Gwinn K, Singleton A, Hardy J (2011) Parkinson's disease and α -synuclein expression. *Mov Disord* 26:2160–2168
- Di Giorgio F, Boulting G, Bobrowicz S, Eggan K (2008) Human embryonic stem cell-derived motor neurons are sensitive to the toxic effect of glial cells carrying an ALS-causing mutation. *Cell Stem Cell* 3:637–648
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485:376–380
- Dowen JM, Fan ZP, Hnisz D, Ren G, Abraham BJ, Zhang LN, Weintraub AS, Schuijers J, Lee TI, Zhao K, Young RA (2014) Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* 159:374–387
- Dumitriu A, Moser C, Hadzi TC, Williamson SL, Pacheco CD, Hendricks AE, Latourelle JC, Wilk JB, DeStefano AL, Myers RH (2012) Postmortem interval influences α -synuclein expression in Parkinson disease brain. *Park Dis* 2012:1–8
- ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57–74
- Ernst J, Kheradpour P, Mikkelsen TS, Shores N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, Ku M, Durham T, Kellis M, Bernstein BE (2011) Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 473:43–49
- Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, Shores N, Whitton H, Ryan RJ, Shishkin AA, Hatan M, Carrasco-Alfonso MJ, Mayer D, Luckey CJ, Patsopoulos NA, De Jager PL, Kuchroo VK, Epstein CB, Daly MJ, Hafler DA, Bernstein BE (2014) Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* 518:337–343
- Fuchs J, Tichopad A, Golub Y, Munz M, Schweitzer K, Wolf B, Berg D, Mueller J, Gasser T (2008) Genetic variability in the SNCA gene influences alpha-synuclein levels in the blood and brain. *FASEB J* 22:1327–1334
- Gasser T, Hardy J, Mizuno Y (2011) Milestones in PD genetics. *Mov Disord* 26:1042–1048
- Gibson G (2012) Rare and common variants: twenty arguments. *Nat Rev Genet* 13:135–145
- GTEx Consortium (2015) Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 348:648–660
- Handoko L, Xu H, Li G, Ngan CY, Chew E, Schnapp M, Lee CW, Ye C, Ping JL, Mulawadi F, Wong E, Sheng J, Zhang Y, Poh T, Chan CS, Kunarso G, Shahab A, Bourque G, Cacheux-Rataboul V, Sung WK, Ruan Y, Wei CL (2011) CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat Genet* 43:630–638
- Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-André V, Sigova AA, Hoke HA, Young RA (2013) Super-enhancers in the control of cell identity and disease. *Cell* 155:934–947
- Hockemeyer D, Jaenisch R (2016) Induced pluripotent stem cells meet genome editing. *Cell Stem Cell* 18:573–586
- Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, Dekelver RC, Katibah GE, Amora R, Boydston EA, Zeitler B, Meng X, Miller JC, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R (2009) Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol* 27:851–857

- Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, Zeitler B, Cherone JM, Meng X, Hinkley SJ, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R (2011) Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol* 29:731–734
- Hou P, Li Y, Zhang X, Liu C, Guan J, Li H, Zhao T, Ye J, Yang W, Liu K, Ge J, Xu J, Zhang Q, Zhao Y, Deng H (2013) Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* 341:651–654
- Israel MA, Yuan SH, Bardy C, Reyna SM, Mu Y, Herrera C, Hefferan MP, Van Gorp S, Nazor KL, Boscolo FS, Carson CT, Laurent LC, Marsala M, Gage FH, Remes AM, Koo EH, Goldstein LSB (2012) Probing sporadic and familial Alzheimer's disease using induced pluripotent stem cells. *Nature* 482:216–220
- Ji X, Dadon DB, Powell BE, Fan ZP, Borges-Rivera D, Shachar S, Weintraub AS, Hnisz D, Pegoraro G, Lee TI, Misteli T, Jaenisch R, Young RA (2016) 3D chromosome regulatory landscape of human pluripotent cells. *Cell Stem Cell* 18:262–275
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816–821
- Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J (2013) RNA-programmed genome editing in human cells. *Elife* 2:e00471
- Jolma A, Yan J, Whittington T, Toivonen J, Nitta KR, Rastas P, Morgunova E, Enge M, Taipale M, Wei G, Palin K, Vaquerizas JM, Vincentelli R, Luscombe NM, Hughes TR, Lemaire P, Ukkonen E, Kivioja T, Taipale J (2013) DNA-binding specificities of human transcription factors. *Cell* 152:327–339
- Jolma A, Yin Y, Nitta KR, Dave K, Popov A, Taipale M, Enge M, Kivioja T, Morgunova E, Taipale J (2015) DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature* 527:384–388
- Kasowski M, Kyriazopoulou-Panagiotopoulou S, Grubert F, Zaugg JB, Kundaje A, Liu Y, Boyle AP, Zhang QC, Zakharia F, Spacek DV, Li J, Xie D, Olarerin-George A, Steinmetz LM, Hogenesch JB, Kellis M, Batzoglou S, Snyder M (2013) Extensive variation in chromatin states across humans. *Science* 342:750–752
- Kilpinen H, Waszak SM, Gschwind AR, Raghav SK, Witwicki RM, Orioli A, Migliavacca E, Wiederkehr M, Gutierrez-Arcelus M, Panousis NI, Yurovsky A, Lappalainen T, Romano-Palumbo L, Planchon A, Bielser D, Bryois J, Padioulet I, Udin G, Thurnheer S, Hacker D, Core LJ, Lis JT, Hernandez N, Reymond A, Deplancke B, Dermitzakis ET (2013) Coordinated effects of sequence variation on DNA binding, chromatin structure, and transcription. *Science* 342:744–747
- Kim HJ, Jeon BS, Yoon MY, Park SS (2012) Increased expression of alpha-synuclein by SNCA duplication is associated with resistance to toxic stimuli. *J Mol Neurosci* 47:249–255
- Kiskinis E et al (2014) Pathways disrupted in human ALS motor neurons identified through genetic correction of mutant SOD1. *Cell Stem Cell* 14:781–795
- Kruger R, Vieira-Saecker A, Kuhn W, Berg D, Muller T, Kuhn N, Fuchs G, Storch A, Hungs M, Weitalla D, Przuntek H, Epplen J, Schols L, Riess O (1999) Increased susceptibility to sporadic Parkinson's disease by a certain combined alpha-synuclein/apolipoprotein E genotype. *Ann Neurol* 45:611–617
- Lee TI, Young RA (2013) Transcriptional regulation and its misregulation in disease. *Cell* 152:1237–1251
- Liang G, Zhang Y (2013) Genetic and epigenetic variations in iPSCs: potential causes and implications for application. *Cell Stem Cell* 13:149–159
- Ligon KL (2003) Loss of Emx2 function leads to ectopic expression of Wnt1 in the developing telencephalon and cortical dysplasia. *Development* 130:2275–2287
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013) RNA-guided human genome engineering via Cas9. *Science* 339:823–826

- Maraganore DM, de Andrade M, Elbaz A, Farrer MJ, Ioannidis JP, Krüger R, Rocca WA, Schneider NK, Lesnick TG, Lincoln SJ, Hulihan MM, Aasly JO, Ashizawa T, Chartier-Harlin MC, Checkoway H, Ferrarese C, Hadjigeorgiou G, Hattori N, Kawakami H, Lambert JC, Lynch T, Mellick GD, Papapetropoulos S, Parsian A, Quattrone A, Riess O, Tan EK, Broeckhoven V (2006) Collaborative analysis of α -synuclein gene promoter variability and Parkinson disease. *JAMA* 296:661–670
- Mariani J, Favaro R, Lancini C, Vaccari G, Ferri AL, Bertolini J, Tonoli D, Latorre E, Caccia R, Ronchi A, Ottolenghi S, Miyagi S, Okuda A, Zappavigna V, Nicolis SK (2012) Emx2 is a dose-dependent negative regulator of Sox2 telencephalic enhancers. *Nucl Acids Res* 40:6461–6476
- Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, Shafer A, Neri F, Lee K, Kutayin T, Stehling-Sun S, Johnson AK, Canfield TK, Giste E, Diegel M, Bates D, Hansen RS, Neph S, Sabo PJ, Heimfeld S, Raubitschek A, Ziegler S, Cotsapas C, Sotoodehnia N, Glass I, Sunyaev SR, Kaul R, Stamatoyannopoulos JA (2012) Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337:1190–1195
- McClellan J, King M-C (2010) Genetic heterogeneity in human disease. *Cell* 141:210–217
- McVicker G, van de Geijn B, Degner JF, Cain CE, Banovich NE, Raj A, Lewellen N, Myrthil M, Gilad Y, Pritchard JK (2013) Identification of genetic variants that affect histone modifications in human cells. *Science* 342:747–749
- Mica Y, Lee G, Chambers SM, Tomishima MJ, Studer L (2013) Modeling neural crest induction, melanocyte specification, and disease-related pigmentation defects in hESCs and patient-specific iPSCs. *Cell Rep* 3:1140–1152
- Miller DW, Hague SM, Clarimon J, Baptista M (2004) α -Synuclein in blood and brain from familial Parkinson disease with SNCA locus triplication. *Neurology* 62:1835–1838
- Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, Cheung VG (2004) Genetic analysis of genome-wide variation in human gene expression. *Nature* 430:743–747
- Musunuru K, Strong A, Frank-Kamenetsky M, Lee NE, Ahfeldt T, Sachs KV, Li X, Li H, Kuperwasser N, Ruda VM, Pirruccello JP, Muchmore B, Prokunina-Olsson L, Hall JL, Schadt EE, Morales CR, Lund-Katz S, Phillips MC, Wong J, Cantley W, Racie T, Ejebe KG, Orholm-Melander M, Melander O, Koteliensky V, Fitzgerald K, Krauss RM, Cowan CA, Kathiresan S, Rader DJ (2010) From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature* 466:714–719
- Nalls MA, Pankratz N, Lill CM, Do CB, Hernandez DG, Saad M, AL DS, Kara E, Bras J, Sharma M, Schulte C, Keller MF, Arepalli S, Letson C, Edsall C, Stefansson H, Liu X, Pliner H, Lee JH, Cheng R, International Parkinson's Disease Genomics Consortium (IPDGC), Parkinson's Study Group (PSG) Parkinson's Research: The Organized GENetics Initiative (PROGENI), 23andMe, GenePD, NeuroGenetics Research Consortium (NGRC), Hussman Institute of Human Genomics (HIHG), Ashkenazi Jewish Dataset Investigator, Cohorts for Health and Aging Research in Genetic Epidemiology (CHARGE), North American Brain Expression Consortium (NABEC), United Kingdom Brain Expression Consortium (UKBEC), Greek Parkinson's Disease Consortium, Alzheimer Genetic Analysis Group, Ikram MA, Ioannidis JP, Hadjigeorgiou GM, Bis JC, Martinez M, Perlmutter JS, Goate A, Marder K, Fiske B, Sutherland M, Xiromerisiou G, Myers RH, Clark LN, Stefansson K, Hardy JA, Heutink P, Chen H, Wood NW, Houlden H, Payami H, Brice A, Scott WK, Gasser T, Bertram L, Eriksson N, Foroud T, Singleton AB (2014) Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet* 46:989–993
- Nishizawa M, Chonabayashi K, Nomura M, Tanaka A, Nakamura M, Inagaki A, Nishikawa M, Takei I, Oishi A, Tanabe K, Ohnuki M, Yokota H, Koyanagi-Aoi M, Okita K, Watanabe A, Takaori-Kondo A, Yamanaka S, Yoshida Y (2016) Epigenetic variation between human induced pluripotent stem cell lines is an indicator of differentiation capacity. *Cell Stem Cell* 19:341–354

- Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, Gribnau J, Barillot E, Blüthgen N, Dekker J, Heard E (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485:381–385
- Paquet D, Kwart D, Chen A, Sproul A, Jacob S, Teo S, Olsen KM, Gregg A, Noggle S, Tessier-Lavigne M (2016) Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 533:125–129
- Pasquali L, Gaulton KJ, Rodríguez-Seguí SA, Mularoni L, Miguel-Escalada I, Akerman I, Tena JJ, Morán I, Gómez-Marín C, van de Bunt M, Ponsa-Cobas J, Castro N, Nammo T, Cebola I, García-Hurtado J, Maestro MA, Pattou F, Piemonti L, Berney T, Gloy AL, Ravassard P, Gómez-Skarmeta JL, Müller F, McCarthy MI, Ferrer J (2014) Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nat Genet* 46:136–143
- Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J (2011) A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 470:279–283
- Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159:1665–1680
- Reinhardt P, Schmid B, Burbulla LF, Schöndorf DC, Wagner L, Glatz M, Höing S, Hargus G, Heck SA, Dhingra A, Wu G, Müller S, Brockmann K, Kluba T, Maisel M, Krüger R, Berg D, Tsytsyura Y, Thiel CS, Psathaki OE, Klingauf J, Kuhlmann T, Klewin M, Müller H, Gasser T, Schöler HR, Sterneckert J (2013) Genetic correction of a LRRK2 mutation in human iPSCs links parkinsonian neurodegeneration to ERK-dependent changes in gene expression. *Cell Stem Cell* 12:354–367
- Ripke S, Neale BM, Corvin A, Walters JTR, Farh K-H, Holmans PA, Lee P, Bulik-Sullivan B, Collier DA, Huang H et al (2014) Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511:421–427
- Rivera CM, Ren B (2013) Mapping human epigenomes. *Cell* 155:39–55
- Roadmap Epigenomics Consortium (2015) Integrative analysis of 111 reference human epigenomes. *Nature* 518:317–330
- Ryan SD, Dolatabadi N, Chan SF, Zhang X, Akhtar MW, Parker J, Soldner F, Sunico CR, Nagar S, Talantova M, Lee B, Lopez K, Nutter A, Shan B, Molokanova E, Zhang Y, Han X, Nakamura T, Masliah E, Yates JR 3rd, Nakanishi N, Andreyev AY, Okamoto S, Jaenisch R, Ambasudhan R, Lipton SA (2013) Isogenic human iPSC Parkinson's model shows nitrosative stress-induced dysfunction in MEF2-PGC1 α transcription. *Cell* 155:1351–1364
- Saha K, Jaenisch R (2009) Technical challenges in using human induced pluripotent stem cells to model disease. *Cell Stem Cell* 5:584–595
- Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N, Colinayo V, Ruff TG, Milligan SB, Lamb JR, Cavet G, Linsley PS, Mao M, Stoughton RB, Friend SH (2003) Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422:297–302
- Schaffer AE, Freude KK, Nelson SB, Sander M (2010) Nkx6 transcription factors and Ptf1a function as antagonistic lineage determinants in multipotent pancreatic progenitors. *Dev Cell* 18:1022–1029
- Schisler JC, Jensen PB, Taylor DG, Becker TC, Knop FK, Takekawa S, German M, Weir GC, Lu D, Mirmira RG, Newgard CB (2005) The Nkx6.1 homeodomain transcription factor suppresses glucagon expression and regulates glucose-stimulated insulin secretion in islet beta cells. *Proc Natl Acad Sci USA* 102:7297–7302
- Singleton AB, Farrer M, Johnson J, Singleton A, Hague S, Kachergus J, Hulihan M, Peuralinna T, Dutra A, Nussbaum R, Lincoln S, Crawley A, Hanson M, Maraganore D, Adler C, Cookson MR, Muentner M, Baptista M, Miller D, Blacato J, Hardy J, Gwinn-Hardy K (2003) alpha-Synuclein locus triplication causes Parkinson's disease. *Science* 302:841
- Singleton AB, Farrer MJ, Bonifati V (2013) The genetics of Parkinson's disease: progress and therapeutic implications. *Mov Disord* 28:14–23
- Soldner F, Jaenisch R (2012) Medicine. iPSC disease modeling. *Science* 338:1155–1156

- Soldner F, Laganière J, Cheng AW, Hockemeyer D, Gao Q, Alagappan R, Khurana V, Golbe LI, Myers RH, Lindquist S, Zhang L, Guschin D, Fong LK, BJ V, Meng X, Urnov FD, Rebar EJ, Gregory PD, Zhang HS, Jaenisch R (2011) Generation of isogenic pluripotent stem cells differing exclusively at two early onset parkinson point mutations. *Cell* 146:318–331
- Soldner F, Stelzer Y, Shivalila CS, Abraham BJ, Latourelle JC, Barrasa MI, Goldmann J, Myers RH, Young RA, Jaenisch R (2016) Parkinson-associated risk variant in distal enhancer of α -synuclein modulates target gene expression. *Nature* 533:95–99
- Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126:663–676
- Takahashi K, Yamanaka S (2013) Induced pluripotent stem cells in medicine and biology. *Development* 140:2457–2461
- Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131:861–872
- The HD iPSC Consortium (2012) Induced pluripotent stem cells from patients with Huntington's disease show CAG-repeat-expansion-associated phenotypes. *Cell Stem Cell* 11:264–278
- Trynka G, Sandor C, Han B, Xu H, Stranger BE, Liu XS, Raychaudhuri S (2013) Chromatin marks identify critical cell types for fine mapping complex trait variants. *Nat Genet* 45:124–130
- Vermunt MW, Reinink P, Korving J, de Bruijn E, Creyghton PM, Basak O, Geeven G, Toonen PW, Lansu N, Meunier C, van Heesch S, Clevers H, de Laat W, Cuppen E, Creyghton MP, Bank NB (2014) Large-scale identification of coregulated enhancer networks in the adult human brain. *Cell Rep* 9:767–779
- Visel A, Taher L, Girgis H, May D, Golonzhka O, Hoch RV, McKinsey GL, Pattabiraman K, Silberberg SN, Blow MJ, Hansen DV, Nord AS, Akiyama JA, Holt A, Hosseini R, Phouanavong S, Plajzer-Frick I, Shoukry M, Afzal V, Kaplan T, Kriegstein AR, Rubin EM, Ovcharenko I, Pennacchio LA, Rubenstein JL (2013) A high-resolution enhancer atlas of the developing telencephalon. *Cell* 152:895–908
- Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Ren B, Rubin EM, Pennacchio LA (2009) ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457:854–858
- Wainger BJ, Kiskinis E, Mellin C, Wiskow O, Han SSW, Sandoe J, Perez NP, Williams LA, Lee S, Boulting G, Berry JD, Brown RH Jr, Cudkowicz ME, Bean BP, Eggan K, Woolf CJ (2014) Intrinsic membrane hyperexcitability of amyotrophic lateral sclerosis patient-derived motor neurons. *Cell Rep* 7:1–11
- Wamstad JA, Wang X, Demuren OO, Boyer LA (2014) Distal enhancers: new insights into heart development and disease. *Trends Cell Biol* 24:294–302
- Wang X, Tucker NR, Rizki G, Mills R, Krijger PH, de Wit E, Subramanian V, Bartell E, Nguyen X-X, Ye J, Leyton-Mange J, Dolmatova EV, van der Harst P, de Laat W, Ellinor PT, Newton-Cheh C, Milan DJ, Kellis M, Boyer LA (2016) Discovery and validation of sub-threshold genome-wide association study loci using epigenomic signatures. *Elife* 5:e10557
- Ward LD, Kellis M (2012) Interpreting noncoding genetic variation in complex traits and human disease. *Nat Biotechnol* 30:1095–1106
- Warren L, Manos PD, Ahfeldt T, Loh Y-H, Li H, Lau F, Ebina W, Mandal PK, Smith ZD, Meissner A, Daley GQ, Brack AS, Collins JJ, Cowan C, Schlaeger TM, Rossi DJ (2010) Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* 7:618–630
- Young JE, Boulanger-Weill J, Williams DA, Woodruff G, Buen F, Revilla AC, Herrera C, Israel MA, Yuan SH, Edland SD, Goldstein LSB (2015) Elucidating molecular phenotypes caused by the SORL1 Alzheimer's disease genetic risk factor using human induced pluripotent stem cells. *Cell Stem Cell* 16:373–385
- Yu DX, Marchetto MC, Gage FH (2013) Perspective. *Cell Stem Cell* 12:678–688

- Yu J, Vodyanik M, Smuga-Otto K, Antosiewicz-Bourget J, Frane J, Tian S, Nie J, Jonsdottir G, Ruotti V, Stewart R, Slukvin I, Thomson J (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318:1917–1920
- Zhang Y, Pak C, Han Y, Ahlenius H, Zhang Z, Chanda S, Marro S, Patzke C, Acuna C, Covy J, Xu W, Yang N, Danko T, Chen L, Wernig M, Südhof TC (2013) NeuroResource. *Neuron* 78:785–798

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Aquatic Model Organisms in Neurosciences: The Genome-Editing Revolution

Jean-Stéphane Joly

Abstract The use of aquatic model organisms has been greatly diversified in laboratories. Zebrafish is the most advanced aquatic species for the use of Crispr-Cas9 in laboratories. Because of the simplicity and broad applicability of this later system, knock-out is now efficiently performed at medium scale. Forward genetics in zebrafish can now be performed by CRISPR-based F0 screening using high speed and high content phenotyping for example by confocal imaging. As zebrafish, marine model organisms have the prominent advantage to be transparent, all the more at young stages (embryos and larvae) or when fixed samples are cleared by novel methods. The Cripsr-Cas9 system is routinely used in the ascidian *Ciona intestinalis*. It also starts to be used in many other marine models, such as the medusa *Clythia hemispherica*. We provide at the end of this review a list of aquatic model species and some examples of questions on the origin of our nervous system that can be coped with these models, where the possibility to perform genome editing would constitute a major advance.

Introduction

With the expansion of biochemistry and molecular biology during the twentieth century, researchers focused more and more on a few model organisms such as nematodes, fruit flies, or mice. These models were amenable to many experimental approaches in molecular biology and biochemistry. Recently, a novel species, the zebrafish, has emerged as a major laboratory model. Initially selected because its transparent embryo is an excellent system in which to study development, it has now become the second most used animal in laboratories worldwide. Thus, the current applications of zebrafish studies are now highly diversified in neurobiology, immunology, adult physiology, oncology, and regenerative medicine, exploiting its advantages for in vivo approaches and imaging.

In parallel, due to the explosion of sequencing methods, there has been a clear trend towards the diversification of model organisms, especially those used in

J.-S. Joly

INRA CASBAH Group, Neuro-Paris Saclay Institute, CNRS, 1 Avenue de la Terrasse, 91 198 Gif-sur-Yvette, France

e-mail: joly@inaf.cnrs-gif.fr

neurosciences. In this paper, we will focus on applications of such new models in evolution of development or biomedical research. There are several reasons to use an increasing number of model organisms: first, the classical models are not representative of most branches of the tree of life, and second, many questions now need to be studied *in vivo* and established models are not always well adapted to many of those biological or medical questions.

We will here elaborate why the use of genome editing in these models offers revolutionary perspectives. Future genome-editing experiments should indeed allow us to unveil the function of critical genes in almost any species; this approach was previously restricted to a few model species, or even restricted to mouse for most precise modifications by homologous recombination. With genome editing, it will become possible with functional data to study the evolutionary origin of highly diversified cell types such as neurons and, in addition, to interrogate how extremely complex cellular organizations such as those found in the brain were built in the course of evolution.

Zebrafish: With the CRiSPR-Cas9 System, Forward Genetic Screens Are Back Again

Zebrafish is a vertebrate, so it has a body plan fundamentally similar to ours (Onai et al. 2014): like humans, zebrafish have a notochord that is a central pile of turgescient cells that confers rigidity to embryo and larvae and serves as the support axis for the development of the spine. Muscles are located on both sides of the notochord. The nervous system is found dorsally and intestine ventrally. Zebrafish embryos, during the so-called “phylotypic stage” (Slack et al. 1993), use the same developmental pattern as human embryos, involving the colinear activation with time and space of Hox gene expressions to build axial structures.

Many organs, including brain, also have similar general organisations in zebrafish and humans. Hence, various aspects of neurobiology can be studied in this species. Zebrafish has a tripartite brain. Brains, as other organs in fish such as fins or hearts, regenerate after a mechanical injury. Some complex behaviours like fear (Amo et al. 2014), social behaviour (Chou et al. 2016) and memory can be studied in adults. Additionally, more basic behaviours can be studied in the 5-day larvae, at a stage most amenable for imaging and for which no authorization for animal experimentation is required (Naumann et al. 2016). This makes the model easy to use for applications such as neurotoxicology performed by academic labs or cosmetology companies.

At the end of the twentieth century, the zebrafish was suggested to be a promising model for genetic screens relevant to human diseases (Mullins et al. 1994). However, during the evolution of vertebrates, an additional genome duplication occurred in the teleost fish lineage, leading to the presence of many duplicated genes in fish genomes that greatly complicate the analysis of screen data. Another pitfall was that the zebrafish has maternal factors carried by the egg. So, when a gene is mutated, the effect of the mutation is in most cases not visible at early stages because maternal

factors are present to insure correct development. Therefore, a disappointingly low number of mutations were identified following large-scale screens by random mutagenesis in embryos. Because many zebrafish screens were performed during early larval stages, most identified mutants failed to exhibit phenotypes similar to human rare diseases, which often appear much later in human life.

After 2000, zebrafish became a very useful model for reverse genetics when phenotypical analyses of mutants were performed with spectacular time-lapse imaging in live fish, for example, during early development (Olivier et al. 2010), neurogenesis (Barbosa et al. 2015), hematopoiesis (Renaud et al. 2011) or immune response (Levraud et al. 2014).

Nowadays, the zebrafish is the most advanced aquatic species for the use of Cripsr-Cas9 in laboratories (Shah and Moens 2016). In this species, however, a challenge remains: efficient insertions of point mutations (for example, to mimic human missense mutations) are still generated at low rates (Renaud et al. 2016). For these applications, the very fast early development of the zebrafish is a drawback, as it probably makes the repair events following DNA cutting by the CRISPR protein highly mosaic and hard to detect in the progeny. Improvements to target the repair construct to the nucleus of the one-cell stage embryo will have to be developed. Modified oligonucleotides could improve KI rates. Alternatively, plasmidic constructs with long homology arms have been used in a recently published method (Hoshijima et al. 2016) to perform KI at large scale in zebrafish; unfortunately, this method remains labor intensive.

Knock-out is now efficiently performed in zebrafish at medium scale (Shah et al. 2015). Hence so-called forward genetics in zebrafish again seems to have a bright future. To study the molecular basis of a given phenotype in a particular cell type, large gene families can be targeted for mutations, and, importantly, mutations of duplicated genes and of their close paralogs can be performed jointly, due to the possibility of injecting arrays of CRISPR guide RNAs.

For large-scale forward screens, methods of large-scale phenotyping, at first hand by 3D imaging, still need to be optimized. Thus, a current priority for zebrafish researchers is to improve rapid imaging technologies at large scale and at later stages, to make zebrafish a better model. Such a model would provide a perfect context for analyzing large collections of mutants. Tissue-clearing methods (Seo et al. 2016) and high-speed imaging methods, such as those using highly sensitive video cameras, have recently emerged and will certainly be crucial for these approaches.

Optimizing the Cripsr-Cas9 System in Transparent Marine Animals

Most marine model organisms have the obvious advantages, as zebrafish, to have transparent embryos and larvae, a feature selected in water throughout evolution to escape predators. Transparency is crucial for the microscopic analysis of development

in these models. Eggs can generally be obtained in huge numbers. Although they are sometimes quite big because of the presence of vitelline reserves, embryos are composed of only a few cells and the lineage analysis is thus easy in these simple and compact embryos. In ascidian embryos, for example, the notochord only has 64 cells. Significant progress in understanding human cardiac developmental gene network was made in ascidian models. This unique insight provided direction for the reprogramming of cell lineages in human cell cultures: following the observation that *Ci-esl/2* and *Ci-mesp* generated cardiac progenitors in ascidians, researchers transdifferentiated human dermal fibroblasts into cardiac progenitors (Islas et al. 2012).

The Cripsr-Cas9 system was used in a study using the ascidian *Ciona intestinalis* (Stolfi et al. 2014). This study, from Lionel Christiaen's group, reported the success of tissue-specific genome editing in this species. Optimization of plasmid constructs was performed, in which specific ubiquitous U6 promoters were used to drive guide RNA expression, and tissue-specific promoters were designed to drive the expression of the Cas9 protein. Introducing the CRISPR–Cas9 components in ascidians was quite easy because a large number of eggs could be electroporated with plasmid DNA, producing both the CRISPR protein and the guide RNAs. Nevertheless, while breeding of this species in inland laboratories has been performed (Joly et al. 2007), it remains difficult. Improvements are still required to reliably obtain the culture of stable lines of transgenic animals.

The Cripsr-Cas9 system is beginning to be used in many other aquatic model organisms (for a fascinating example in lampreys, Square et al. 2015). For example, fascinating experiments (unpublished) have been performed in the medusa *Clytia hemisphaerica* (Tsuyoshi Momose, CNRS Villefranche-sur-Mer, personal communication). This model recently emerged as a remarkable cnidarian species useful for evo-devo studies (Houliston et al. 2010). Experiments were in part supported by the French network named “Etude Fonctionnelle sur les Organismes Modèles (EFOR, www.efor.net)”, promoting research—including genome editing and imaging—in metazoan model organisms.

Success in *Clytia* is due to obtaining full life cycles in laboratory aquaria, generating quasi-immortal, vegetatively growing colonies. Moreover, adult medusae spawn daily, generating transparent and easy to inject eggs. Rates of Cripsr-Cas9 knock-outs are strikingly high: over 700 embryos, all with potential knock-outs as seen by the loss of fluorescent protein activity, can be generated in a single injection experiment.

Injection of the NLS Cas9 protein/sgRNA into unfertilized eggs can be performed as soon as 1 hour after ovulation and before subsequent fertilization. In this condition, the Cas9 protein probably has time to be targeted to its cutting site before the first division of the embryo occurs. One first exciting application of this method was the deletion of green fluorescent proteins, making newly generated transgenic lines suitable for imaging applications. Indeed, in these species, endogenous fluorescent activity hinders potential observation of GFP in newly generated transgenic animals. The availability of mutants in such species will offer novel routes for fundamental research in evolution and development (Galliot et al. 2009). In such non-marine species, a challenge remains to keep animals alive in captivity long term. Significant investments will also be needed to obtain colonies of

inbred lines, which should become reference laboratory lines for the corresponding communities of worldwide researchers.

More and More Aquatic Model Organisms for Diversified Uses

Examples of emerging aquatic models of increasing evolutionary distances from vertebrates are described below. These models are located at several key nodes of the metazoan branches of the tree of life. Alternative fish models to zebrafish (Schartl 2014) allow the study of particular evolutionary processes, such as adaptation to cave life in the *Astyanax mexicanus*. To study the emergence of synapomorphic (specific and shared) vertebrate features in evolution, lampreys (including the sea lamprey, *Petromyzon marinus*, an agnathan), lancelets (the cephalochordate amphioxus, *Branchiostoma lanceolatum*) or tunicates (such as the urochordate *Ciona intestinalis*) are very relevant models. Other more distantly related bilaterians such as the polychaete annelid *Platynereis dumerilii* provide insight into which features were already present in the common ancestor of all bilaterian organisms, the so-called “urbilateria.” Even more distant metazoans, with no bilateral symmetry but rather radial symmetry, are also used in laboratories. Thus, Ctenophores (sea gooseberries) and Cnidarians (corals, jellyfish, sea anemones) can be used to study ancient features of nerve cell types. Also, sponges and placozoans constitute fascinating basal Metazoans, with no nervous system.

In Biomedical Research, Why and How Should We Use Aquatic Models to Study Diseases of the Nervous System?

A first obvious use of model organisms is to generate so-called “models of diseases.” In most cases, these models are mutants or transgenic animals that reproduce at best pathological conditions observed in humans, such as neural degeneration. With the advent of precise genome-editing methods, the capacity to generate point mutations in any model by introducing a repair construct bearing the mutations will constitute a true revolution. Indeed, mutations at orthologous positions to variations found in human diseases can be generated in these model animals if genomes can be aligned in the region surrounding the mutation. Then the phenotypical effects of the abnormal protein function can be described, for example, using live imaging to observe abnormal cell behaviors, such as proliferation or migration. Therein resides the great advantage of these aquatic models, with a diversity of developmental and genomic contexts and transparency allowing easy imaging.

Moreover, in an applied perspective in regenerative medicine, understanding how cell types emerged during evolution helps us identify crucial pathways that are, for example, active in stem cells in normal and pathological conditions. The Crispr-Cas9

system applied to aquatic model organisms will offer an unprecedented opportunity to characterize the key genes and pathways that are active following injury or degeneration. They promote regeneration responses in animals with regenerative capacities and could be useful in regenerative medicine, if (re)activated in humans (Karra and Poss 2017).

A Short Natural History of the Nervous System: Several Questions on Its Origin

This chapter provides examples of what marine model organisms bring us in the domain of neuroscience. Many essential questions can indeed be examined with these models, providing us more knowledge about how the human brain was shaped through evolution, and this should help us better understand pathologies and their pleiotropic effects.

Evolution indeed sifts through the noise and allows us to focus on key genes and pathways that have remained crucial for specific cell types throughout evolution. Looking at extant species, and describing common features that are likely to be ancestral and shared, is also a way to “reconstruct” the nervous system of the putative last common ancestor between the two compared species. In this domain, the absence of fossils of nervous system and brains has impeded researchers.

According to Detlev Arendt (Arendt 2008), homology hypotheses are based on the comparison of genes, cytological features and ontological location in the body of the embryo. Functional experiments with genome editing in model organisms will add extremely important cues to this domain.

The Ctenophores are colorful planctonic animals that have a sophisticated nervous net allowing them to swim and to emit beautiful waves of fluorescent flashes. A long-standing debate is whether these animals have neuron-like cells, which would have appeared independently of the neurons of our nervous system during evolution. In this respect, examining the phylogenetic position of these animals is primordial: are they closely related to bilaterians or rather do they form an out-group of metazoans? If they are more distantly related to us than sponges, this would indeed mean that the nervous system was invented twice in evolution, because it is very unlikely that sponges, which would be closer relatives to bilaterians, secondarily lost neural cells.

Until this controversy is resolved, it will not be possible to know whether there were two independent origins of the nervous system in animals, which would, of course, be a very exciting possibility. However, as argued in a recent review (Jager and Manuel 2016), many lines of evidence now suggest that ctenophores are closely related to bilaterians and that neurons appeared only once in evolution. In favour of a single nervous system type is the presence of the well-known neurogenesis SoxB gene and the presence of acetylcholine and numerous common GPCRs. In any case,

ctenophores should provide key insights into deeply conserved features of animal neural cells.

Marine organisms also allow us to examine how the nervous system became condensed and centralized (Arendt et al. 2016) while becoming much more complex in the course of evolution to terrestrial life (Nomaksteinsky et al. 2009). Starting from a diffuse nerve net in a swimming larva, brains became huge, formed from the so-called embryonic neural plate in vertebrates later undergoing neurulation to form the neural tube.

Another question about the origin of the nervous system is how the vertebrate tripartite nervous system arose in the course of evolution. The vertebrate central nervous system is composed of anterior brain, posterior brain and spinal cord. Such organisation occurred with the emergence of borders between brain domains, such as the midbrain/hindbrain boundary. Recently, following studies in an annelid worm, Arendt and colleagues proposed that the nervous system of the common bilaterian ancestor was probably composed of two independent domains (corresponding to the vertebrate forebrain/midbrain and hindbrain) that later fused during evolution (Tosches and Arendt 2013).

Very recently, and in line with Arendt's hypothesis, Chris Lowe's group proposed that the adult body plan of an indirect developing hemichordate develops by adding a Hox pattern trunk to an anterior larval territory, confirming the hypothesis that marine larvae are "swimming heads" (Gonzalez et al. 2017).

Concerning the origin of vertebrate synapomorphic characters, ascidians provide evidence of the precraniate or prevertebrate origins of the neural crest. Neural crest cells are stem cells with migratory behaviors and the capacity to differentiate into incredibly diversified tissues and locations. The evolutionary origin of neural crest is obscure. They were first thought to be a vertebrate synapomorphy. Bill Jeffery identified migratory pigment cells in ascidians (Jeffery et al. 2004; Jeffery 2006) and, more recently, Lionel Christiaen's group identified migrating neuron precursors in the ascidian *Ciona intestinalis*. Interestingly, these precursors were shown to arise from the border of the neural plate, a hallmark of the neural crest in vertebrates (Stolfi et al. 2015).

Conclusion

Crispr-Cas9 system and other genome-editing techniques can now be used in various aquatic model organisms for extremely diversified applications, which will lead to new models of brain diseases for biomedical research. Also, these models will provide strategies to characterize human genetic variations linked to diseases at large scale and suggest new avenues for regenerative medicine, because of the exceptional abilities of most aquatic models to regenerate. Basic biological knowledge will, of course, benefit from this revolution, and one might expect many more revolutionary discoveries from the exploration of the genomes of multiple animal aquatic species using Crispr-Cas9 approaches.

Acknowledgments Marine Joly is warmly acknowledged for help and discussion on the talk and manuscript. Pierre Boudinot contributed a lot to the improvement of earlier versions of this manuscript. I wish to thank T. Momose for sharing unpublished data and M. Manuel for discussion.

References

- Amo R, Fredes F, Kinoshita M, Aoki R, Aizawa H, Agetsuma M, Aoki T, Shiraki T, Kakinuma H, Matsuda M, Yamazaki M, Takahoko M, Tsuboi T, Higashijima S, Miyasaka N, Koide T, Yabuki Y, Yoshihara Y, Fukai T, Okamoto H (2014) The habenulo-raphé serotonergic circuit encodes an aversive expectation value essential for adaptive active avoidance of danger. *Neuron* 84:1034–1048
- Arendt D (2008) The evolution of cell types in animals: emerging principles from molecular studies. *Nat Rev Genet* 9:868–882
- Arendt D, Tosches MA, Marlow H (2016) From nerve net to nerve ring, nerve cord and brain evolution of the nervous system. *Nat Rev Neurosci* 17:61–72
- Barbosa JS, Sanchez-Gonzalez R, Di Giaimo R, Baumgart EV, Theis FJ, Götz M, Neurodevelopment NJ (2015) Live imaging of adult neural stem cell behavior in the intact and injured zebrafish brain. *Science* 348:789–793
- Chou MY, Amo R, Kinoshita M, Cherng BW, Shimazaki H, Agetsuma M, Shiraki T, Aoki T, Takahoko M, Yamazaki M, Higashijima S, Okamoto H (2016) Social conflict resolution regulated by two dorsal habenular subregions in zebrafish. *Science* 352:87–90
- Galliot B, Quiquand M, Ghila L, de Rosa R, Miljkovic-Licina M, Chera S (2009) Origins of neurogenesis, a cnidarian view. *Dev Biol* 332:2–24
- Gonzalez P, Uhlinger KR, Lowe CJ (2017) The adult body plan of indirect developing hemichordates develops by adding a Hox-patterned trunk to an anterior larval territory. *Curr Biol* 27:87–95
- Hoshijima K, Juryneć MJ, Grunwald DJ (2016) Precise genome editing by homologous recombination. *Meth Cell Biol* 135:121–147
- Houlston E, Momose T, Manuel M (2010) *Clytia hemisphaerica*: a jellyfish cousin joins the laboratory. *Trends Genet* 26:159–167
- Islas JF, Liu Y, Weng KC, Robertson MJ, Zhang S, Prejusa A, Harger J, Tikhomirova D, Chopra M, Iyer D, Mercola M, Oshima RG, Willerson JT, Potaman VN, Schwartz RJ (2012) Transcription factors ETS2 and MESP1 transdifferentiate human dermal fibroblasts into cardiac progenitors. *Proc Natl Acad Sci USA* 109:13016–13021
- Jager M, Manuel M (2016) Ctenophores: an evolutionary-developmental perspective. *Curr Opin Genet Dev* 39:85–92
- Jeffery WR (2006) Ascidian neural crest-like cells: phylogenetic distribution, relationship to larval complexity, and pigment cell fate. *J Exp Zool B Mol Dev Evol* 306:470–480
- Jeffery WR, Strickler AG, Yamamoto Y (2004) Migratory neural crest-like cells form body pigmentation in a urochordate embryo. *Nature* 431:696–699
- Joly JS, Kano S, Matsuoka T, Auger H, Hirayama K, Satoh N, Awazu S, Legendre L, Sasakura Y (2007) Culture of *Ciona intestinalis* in closed systems. *Dev Dyn* 236:1832–1840
- Kara R, Poss KD (2017) Redirecting cardiac growth mechanisms for therapeutic regeneration. *J Clin Invest* 127:427–436
- Levrault JP, Palha N, Langevin C, Boudinot P (2014) Through the looking glass: witnessing host-virus interplay in zebrafish. *Trends Microbiol* 22:490–497
- Mullins MC, Hammerschmidt M, Haffter P, Nüsslein-Volhard C (1994) Large-scale mutagenesis in the zebrafish: in search of genes controlling development in a vertebrate. *Curr Biol* 4:189–202

- Naumann EA, Fitzgerald JE, Dunn TW, Rihel J, Sompolinsky H, Engert F (2016) From whole-brain data to functional circuit models: the zebrafish optomotor response. *Cell* 167:947–960
- Nomaksteinsky M, Röttinger E, Dufour HD, Chettouh Z, Lowe CJ, Martindale MQ, Brunet JF (2009) Centralization of the deuterostome nervous system predates chordates. *Curr Biol* 19:1264–1269
- Olivier N, Luengo-Oroz MA, Duloquin L, Faure E, Savy T, Veilleux I, Solinas X, Débarre D, Bourguin P, Santos A, Peyri  ras N, Beaurepaire E (2010) Cell lineage reconstruction of early zebrafish embryos using label-free nonlinear microscopy. *Science* 329:967–971
- Onai T, Irie N, Kuratani S (2014) The evolutionary origin of the vertebrate body plan: the problem of head segmentation. *Annu Rev Genomics Hum Genet* 15:443–459
- Renaud O, Herbolme P, Kissa K (2011) Studying cell behavior in whole zebrafish embryos by confocal live imaging: application to hematopoietic stem cells. *Nat Protoc* 6:1897–1904
- Renaud JB, Boix C, Charpentier M, De Cian A, Cochenne J, Duvernois-Berthet E, Perrouault L, Tesson L, Edouard J, Thinard R, Cherifi Y, Menoret S, Fontani  re S, de Croz   N, Fraichard A, Sohm F, Anegon I, Concordet JP, Giovannangeli C (2016) Improved genome editing efficiency and flexibility using modified oligonucleotides with TALEN and CRISPR-Cas9 nucleases. *Cell Rep* 14:2263–2272
- Schartl M (2014) Beyond the zebrafish: diverse fish species for modeling human disease. *Dis Model Mech* 7:181–192
- Seo J, Choe M, Kim SY (2016) Clearing and labeling techniques for large-scale biological tissues. *Mol Cells* 39:439–446
- Shah AN, Moens CB (2016) Approaching perfection: new developments in zebrafish genome engineering. *Dev Cell* 36:595–596
- Shah AN, Davey CF, Whitebitch AC, Miller AC, Moens CB (2015) Rapid reverse genetic screening using CRISPR in zebrafish. *Nat Methods* 12:535–540
- Slack JMW, Holland PWH, Graham CF (1993) The zootype and the phylotypic stage. *Nature* 361:490–492
- Stolfi A, Ryan K, Meinertzhagen IA, Christiaen L (2015) Migratory neuronal progenitors arise from the neural plate borders in tunicates. *Nature* 527:371–374
- Square T, Rom   ek M, Jandzik D, Cattell MV, Klymkowsky M, Medeiros DM (2015) CRISPR/Cas9-mediated mutagenesis in the sea lamprey *Petromyzon marinus*: a powerful tool for understanding ancestral gene functions in vertebrates. *Development* 142(23):4180–4187
- Stolfi A, Gandhi S, Salek F, Christiaen L (2014) Tissue-specific genome editing in *Ciona* embryos by CRISPR/Cas9. *Development* 141(21):4115–4120
- Tosches MA, Arendt D (2013) The bilaterian forebrain: an evolutionary chimaera. *Curr Opin Neurobiol* 23:1080–1089

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Genome-Wide Genetic Screening in the Mammalian CNS

Mary H. Wertz and Myriam Heiman

Abstract Genes linked to major neurodegenerative diseases, including Alzheimer's, Parkinson's, and Huntington's diseases, were first identified over 15 years ago, but neither a full molecular explanation for the cell loss seen in human patients nor a curative therapy has yet been achieved for any of these diseases. In most model organisms, when new hypotheses are needed to explain a cellular process, genetic screens are the tool of choice. For example, 'synthetic lethal' screens can lead to the identification of genes that enhance the toxicity of a particular mutation, revealing pathways critical for surviving the mutation's effects. To date, however, genome-wide unbiased screens are not feasible in mammalian central nervous system neurons except in vitro, which fails to capture the relevant disease pathologies, and no genome-wide screens have yet been conducted in the mammalian central nervous system. We outline in this short monograph the steps needed to implement a methodology that allows for genome-wide genetic screening in the central nervous system of mice to study both normal and degenerative disease gene function.

Introduction

Genome-wide genetic screens have been used for decades in *S. cerevisiae*, *C. elegans*, and *D. melanogaster* to elucidate many important aspects of cell biology. Such traditional mutagenesis-based genome-wide genetic screens have been impossible to routinely perform in mice due to the prohibitively large number of mice that would be needed. However, the ability to perform such screens in the

M.H. Wertz

Picower Institute for Learning and Memory, Cambridge, MA 02139, USA

Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

M. Heiman (✉)

MIT Department of Brain and Cognitive Sciences, Cambridge, MA 02139, USA

Picower Institute for Learning and Memory, Cambridge, MA 02139, USA

Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

e-mail: mheiman@mit.edu

nervous system would enable the generation of new hypotheses regarding the molecular mechanisms of disease. For example, unbiased genome-wide genetic screens could reveal genes that are involved in the toxicity of disease-associated mutations, such as mutations in the huntingtin gene that are found in human Huntington's disease patients. Such neurodegenerative disease-focused genetic screens have been attempted in *S. cerevisiae*, *C. elegans*, and *D. melanogaster*, but these screens by definition fail to capture the full complexity of mammalian neurons—an important point, given the widely varying susceptibility seen amongst cell types in neurodegenerative diseases. Alternatively, genome-wide genetic screens that utilize mammalian neuron-like cells have been conducted in vitro, but these screens are also unable to recapitulate the many aspects of in vivo neurons in the mammalian central nervous system (CNS). The in vivo context may be essential to many aspects of CNS biology, given for example the diversity of CNS cell types, the likely importance of both cell autonomous and non-cell autonomous factors in neurodegenerative diseases, and the known age dependency of most neurodegenerative diseases. Ideally, these screens would be done in mammalian neurons in their native cellular environment.

To bypass the difficulties associated with classical mutagenesis screening as well as the diploid nature of mammalian genomes, genome-wide short hairpin (shRNA) and clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) screening approaches have been applied to mammalian cells in vitro (e.g., among many others, Moffat et al. 2006; Root et al. 2006; Shalem et al. 2014; Wang et al. 2014; Zhou et al. 2014). Despite the power of these methodologies, there are many challenges to their application in vivo, especially in the CNS. Indeed, mammalian genome-wide shRNA or CRISPR genetic screens have been conducted mainly either in vitro, in transformed cell lines, or else in primary cells manipulated ex vivo and then returned in vivo (Chen et al. 2015; Graham and Root 2015). Based on the insights that have come from such studies, genome-wide genetic screening could be a powerful tool for the study of normal cellular function and degenerative disease processes in the mammalian CNS, provided that such screens are performed in the context of models that recapitulate the relevant biology. For this reason, we recently developed a genetic screening workflow that allows rapid, high-sensitivity screening in the mouse CNS for aging and neurodegenerative disease processes (Shema et al. 2015). This workflow combines the use of (1) pooled lentiviral shRNA libraries; (2) stereotaxic injection of these pools into mouse models of neurodegenerative disease and wild-type littermates; (3) incubation of injected libraries, such that shRNAs that enhance neurodegenerative disease gene toxicity lead to cell death; and (4) sequencing and analysis of the remaining shRNAs elements in all surviving cells in order to determine which constructs have enhanced cell death and thus 'drop out' of library representation (Fig. 1).

For our genetic screening workflow, we first used shRNA viral libraries, since genome-wide shRNA libraries for the mouse genome are available and have been successfully utilized in many studies. In our pilot screen, we chose to target genes that enhanced the lethality of a fragment of the mutant huntingtin gene.

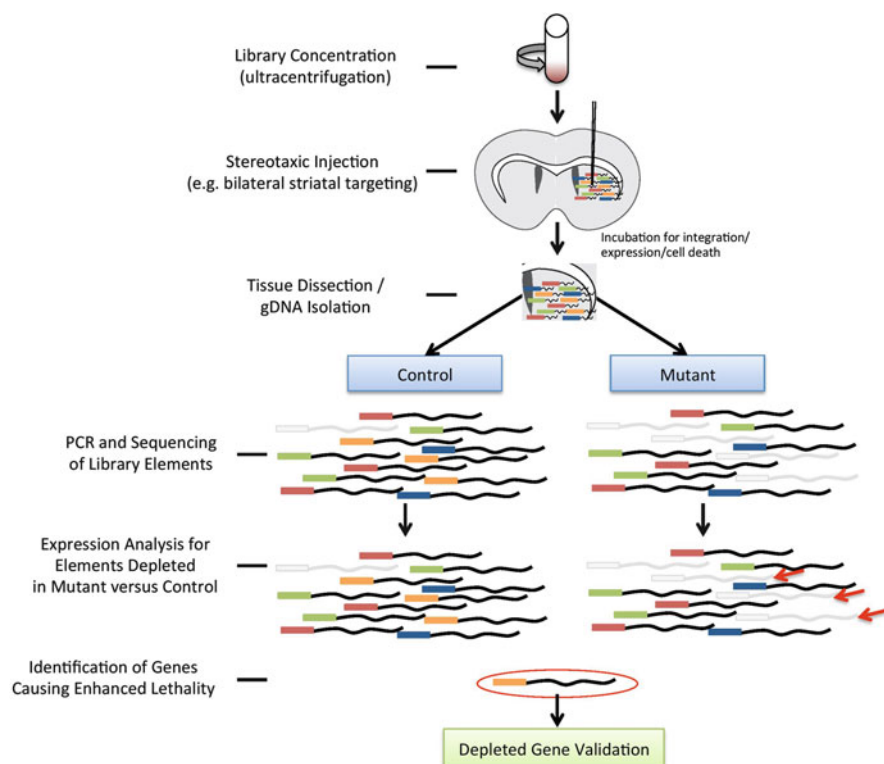


Fig. 1 Genome-wide genetic screening in the mammalian CNS. Pooled viral libraries containing shRNAs, gRNAs, or cDNAs are first concentrated via ultracentrifugation to a high titer suitable for bilateral injection into the striatum (or other CNS target area) for in vivo transduction. After injection, viral payloads are allowed to integrate into the host cell genome and express for several weeks. During this time, genetic perturbations that enhance toxicity in a disease model context may enhance cell death. The targeted tissue is then carefully dissected and the genomic DNA is extracted. After PCR and sequencing of library elements, deconvolution and analysis reveals the library representation. Those genes that enhance cell death in vivo will be depleted or lost from the library (*red arrows*) in the mutant as compared to control animals and thus can be identified as potential modifiers of neuronal toxicity (*orange barcode*). These genes can then be confirmed in follow-up validation experiments

Huntington's disease is the most common inherited neurodegenerative disorder, but the molecular pathways that are essential for mutant Huntingtin protein's toxicity in vivo are not fully understood. Huntington's disease is particularly amenable to genetic screening, as it is a monogenic disease for which several mouse models exist (Huntington's Disease Collaborative Research Group 1993; Mangiarini et al. 1996), and the most greatly affected brain region (caudate-putamen/striatum) is a well-delineated sub-cortical structure. Since Huntington's disease displays an aging component (Mattson and Magnus 2006), we first chose to target a set of genes that showed altered expression both in the context of normal aging and in mutant

Huntingtin expression in CNS neurons. Of these genes, we identified one, *Gpx6*, that enhances the toxicity of mutant Huntingtin protein when its expression is reduced and that partially reverses Huntington's disease-like symptomatology when overexpressed in mouse striatum (Shema et al. 2015). With this proof-of-principle study complete, we outline below parameters that will be essential to extend this methodology to perform genome-wide screening in the mammalian CNS.

Genome-Wide Viral Library Preparation and Delivery

Stable and long-term transduction of post-mitotic neurons by lentivirus has been in use for over 20 years (Naldini et al. 1996a, b). The available genome-wide shRNA or CRISPR guide RNA (gRNA) viral libraries described to date are typically packaged with a vesicular stomatitis virus-G (VSV-G) envelope due to resulting high stability and wide host cell range of the virus (Moffat et al. 2006; Root et al. 2006; Shalem et al. 2014; Wang et al. 2014; Zhou et al. 2014). VSV-G pseudotyping additionally enhances the neuronal tropism of lentivirus (Burns et al. 1993; Yee et al. 1994). Concentration of the initially obtained viral supernatants by ultracentrifugation yields high titers of intact VSV-G pseudotyped virus (Burns et al. 1993; Yee et al. 1994) that are essential for in vivo stereotaxic injections into the brain. As lentivirus is a relatively large virus (~100 nm), its diffusion is limited in the dense neuropil of the mammalian CNS. Given this consideration, injection parameters must be carefully optimized for each target tissue region (Cetin et al. 2006). Adeno-assisted virus (AAV) represents another potential delivery vehicle for pooled screens. As AAV is a small (~20 nm) non-enveloped virus that can be concentrated to very high titers, it is ideal for in vivo CNS delivery and, for this reason, AAV vectors have been widely used in human gene therapy clinical trials (Hocquemiller et al. 2016). Drawbacks to using AAV include its limited payload size (~4.5 kb), which limits the ability to perform cDNA overexpression screens, and the fact that the AAV serotype to be used may need to be optimized for the CNS cell type of interest.

The choice of viral library payload will depend on the experimental goals of the screening project but, in principle, cDNA, shRNA, or CRISPR gRNA libraries could all be used to interrogate CNS gene function. A recent study that compared the results of both shRNA and CRISPR/Cas9 gRNA screens to identify essential genes in a leukemia cell line found modest correlation between screen results (Morgens et al. 2016), and in some biological contexts it may be that both knockdown (shRNA or CRISPRi; Qi et al. 2013) and knockout (CRISPR) strategies should be employed to examine disease-relevant mechanisms (Deans et al. 2016).

Once a viral library is chosen and prepared, the number of cells needed for genome-wide screening should be estimated to determine the feasibility of conducting screening in the desired CNS cell population. Based on past shRNA and CRISPR gRNA screens, approximately 1000 cells should be targeted per

library element, depending on the details of the screen. This number is necessary to average out noise in the assay itself, and also heterogeneity in the genetic perturbation induced in each cell, as well as inherent variability in the response of the screened cells to the perturbation. (Graham and Root 2015). Thus, for a CRISPR gRNA library that contains approximately four gRNAs per protein-coding gene, the 80,000 library elements should each be targeted to approximately 1000 cells (thus 80 million cells in total across all replicates). Reducing either biological or technical variability, for example by employing a more homogeneous cell population, can reduce the number of cells needed in each screen. The time between injection of the library and harvesting of the cells for analysis will be determined by experimental goals and could range from several days to months, depending on the rate of progression of the CNS phenotype being screened.

Interpretation of Results

As in other pooled RNAi/CRISPR screens, in CNS genome-wide screens genomic DNA is extracted from the target tissue and subjected to PCR for constant regions in the shRNA/gRNA sequences. The samples are then barcoded, pooled, sequenced, and run through deconvolution analysis to determine the representation of each individual library element. A few key factors that determine the quality and the interpretation of the results are the number of elements targeting each individual gene, if it is shRNA, gRNA or cDNA, and the depth of sequencing. A number of different methods and tools have been designed to analyze pooled screening data, and these differ based on library complexity and the type of element used to induce the perturbation. There are also a number of analytical tools that have been developed for analysis of RNAi and CRISPR genome-wide screens to assign enrichment/depletion scores, including, for example, Model-based Analysis of Genome-wide CRISPR/Cas9 Knockout (MAGeCK), RNAi Enrichment Gene Ranking (RIGER), and STARS, which rank shRNA or gRNA performance based on magnitude and consistency of elements for each gene that is depleted or enriched (Luo et al. 2008; Li et al. 2014; Doench et al. 2016). Another tool, Cas9 high-Throughput maximum Likelihood Estimator (casTLE), can be used to combine data of shRNA and gRNA screens to increase sensitivity (Morgens et al. 2016).

A primary genome-wide *in vivo* screen may yield hundreds of hits, and independent validation of these targets is necessary to confirm the assay results and the gene specificity of the observed effects and to understand the role of the genes in modifying disease phenotypes (Fig. 2). Two strategies for validation of genome-wide *in vivo* screening can be utilized to assess performance of the primary screen and confirm hits. Creation of sub-pool libraries allows efficient validation of several hundred potential hits. This strategy has been used to validate findings *in vitro* and in cells reintroduced *in vivo* (Chen et al. 2015). Sub-pool elements could include shRNAs or gRNAs that target genes that were unchanged in the primary screen, an additional 4–5 shRNAs or gRNAs for the primary screen hit genes, and carefully

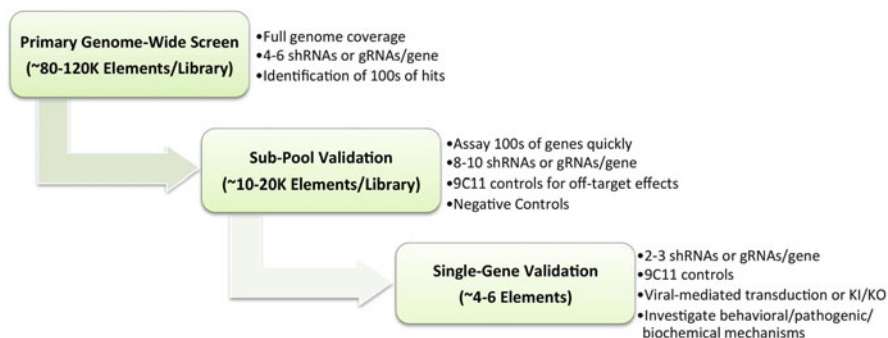


Fig. 2 Validation of in vivo screening hits. A primary genome-wide in vivo screen is completed with at least 4–6 elements targeting a single gene, leading to libraries composed of ~80,000–120,000 elements. Validation of genes identified in the primary screen can be completed with smaller sub-pool libraries of only ~10,000–20,000 elements, which must be carefully designed to include an increased number of unique elements (~10) targeting the positive hits identified in the genome-wide screen as well as appropriate controls. These controls come in the form of elements targeting non-genomic sequences, genes unchanged in the primary screen, and C911 controls that can reveal seed-related off-target effects of hits. Sub-pool validation using a combination of multiple modalities (i.e., cDNA, gRNA and shRNA) may also be used to increase confidence in hits. Additional validation at the single-gene level can then be performed via viral transduction of two to three targeting elements and appropriate controls or else traditional knockdown/knockout/overexpression studies. Such single-gene validation is particularly important for investigation of behavioral and pathogenic readouts of disease processes as well as biochemical mechanisms underlying modification of toxicity

designed 9C11 controls that reveal shRNA seed-related off-target effects (Buehler et al. 2012). A second approach to validation is by interrogation of individual hits via traditional single-gene knockout/knockdown/overexpression studies. To do this, in addition to classical germline genetic perturbations, CNS viral delivery of top screen-hit validated shRNAs/gRNAs/cDNAs by stereotaxic injection can be used to rapidly introduce a single genetic perturbation, as is routinely performed in many CNS studies with AAV or retroviral vectors. This type of more traditional validation approach has the advantage that it can be used to assay various behavioral and pathological readouts of disease progression and to tease out specific biochemical pathways.

In addition to validation of targets from a single primary screen utilizing a particular genetic perturbation, comparison of data from two different modalities, i.e., both shRNA knockdown and gRNA knockout, or cDNA overexpression and gRNA knockout, may be beneficial. This cross-platform approach has shown to produce varying degrees of overlap in identified targets (Deans et al. 2016; Evers et al. 2016; Morgens et al. 2016), highlighting the possible utility of applying several types of perturbations in a multi-armed screen to enhance the specificity of hits or else to expand the type of hits that can be obtained (e.g., certain phenotypes may only be revealed upon gene knockdown, not knockout). While primary genome-wide cDNA screens may be challenging due to the efficiency of

packaging genome-wide cDNAs into viral vectors, the potential for use in sub-pool screening of a smaller number of genes is much higher. Therefore, a combination of these techniques (cDNA overexpression, shRNA knockdown, gRNA CRISPR or CRISPRi) may yield increased sensitivity to uncover biological pathways relevant to neuronal function and dysfunction.

Future Directions

Looking forward, the ability to perform cell type-specific genome-wide genetic screens will be helpful to fully understand CNS disease mechanisms, as most neurological diseases display cell type-specific patterns of vulnerability, including the two most prevalent neurodegenerative diseases, Alzheimer's disease and Parkinson's disease (Mattson and Magnus 2006). The use of a conditional Cas9-expressing mouse line crossed to one that expresses *Cre* recombinase in the cell type of interest should allow such cell type-specific CRISPR knockout or CRISPRi gRNA screens. Conditional or inducible systems for use with mammalian retroviral vectors (Beier et al. 2011) could be useful for lentiviral-based shRNA or cDNA overexpression screens. Genome-wide genetic screening in the mammalian CNS may make it possible to interrogate molecular mechanisms linked to all the major neurodegenerative diseases and eventually to identify common vulnerability factors that may exist among these diseases, for example, aging-related and proteostasis pathways. Finally, the ability to perform genetic screening in the CNS around a non-death phenotype (e.g., biomarker expression using flow-sorting to isolate the hit cells) would greatly expand the power of genome-wide approaches.

Acknowledgments We wish to thank Dr. David E. Root and Dr. John G. Doench for useful discussions and advice. This work was supported by grants to M.H. by the JPB Foundation and NIH/NINDS.

References

- Beier KT, Samson ME, Matsuda T, Cepko CL (2011) Conditional expression of the TVA receptor allows clonal analysis of descendants from Cre-expressing progenitor cells. *Dev Biol* 353:309–320
- Buehler E, Chen YC, Martin S (2012) C911: a bench-level control for sequence specific siRNA off-target effects. *PLoS One* 7:e51942
- Burns JC, Friedmann T, Driever W, Burrascano M, Yee JK (1993) Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc Natl Acad Sci USA* 90:8033–8037
- Cetin A, Komai S, Eliava M, Seeburg PH, Osten P (2006) Stereotaxic gene delivery in the rodent brain. *Nat Prot* 1:3166–3173

- Chen S, Sanjana NE, Zheng K, Shalem O, Lee K, Shi X, Scott DA, Song J, Pan JQ, Weissleder R, Lee H, Zhang F, Sharp PA (2015) Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell* 160:1246–1260
- Deans RM, Morgens DW, Okesli A, Pillay S, Horlbeck MA, Kampmann M, Gilbert LA, Li A, Mateo R, Smith M, Glenn JS, Carette JE, Khosla C, Bassik MC (2016) Parallel shRNA and CRISPR-Cas9 screens enable antiviral drug target identification. *Nat Chem Biol* 12:361–366
- Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, Smith I, Tothova Z, Wilen C, Orchard R, Virgin HW, Listgarten J, Root DE (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* 34:184–191
- Evers B, Jastrzebski K, Heijmans JP, Grenrum W, Beijersbergen RL, Bernards R (2016) CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat Biotechnol* 34:631–633
- Graham DB, Root DE (2015) Resources for the design of CRISPR gene editing experiments. *Genome Biol* 16:260
- Hocquemiller M, Giersch L, Audrain M, Parker S, Cartier N (2016) Adeno-associated virus-based gene therapy for CNS diseases. *Hum Gene Ther* 27:478–496
- Huntington's Disease Collaborative Research Group (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 72:971–983
- Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, Irizarry RA, Liu JS, Brown M, Liu XS (2014) MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* 15:554
- Luo B, Cheung HW, Subramanian A, Sharifnia T, Okamoto M, Yang X, Hinkle G, Boehm JS, Beroukhi R, Weir BA, Mermel C, Barbie DA, Awad T, Zhou X, Nguyen T, Qi Q, Li C, Golub TR, Meyerson M, Hacohen N, Hahn W, Lander ES, Sabatini DM, Root DE (2008) Highly parallel identification of essential genes in cancer cells. *Proc Natl Acad Sci USA* 105:20380–20385
- Mangiarini L, Sathasivam K, Seller M, Cozens B, Harper A, Hetherington C, Lawton M, Trotter Y, Leach H, Davies SW, Bates GP (1996) Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell* 87:493–506
- Mattson MP, Magnus T (2006) Ageing and neuronal vulnerability. *Nat Rev Neurosci* 7:278–294
- Moffat J, Gruenberg DA, Yang X, Kim SY, Kloepper AM, Hinkle G, Qi Q, Eisenhaure TM, Luo B, Grenier JK, Carpenter AE, Foo SY, Stewart SA, Stockwell BR, Hacohen N, Hahn WC, Lander ES, Sabatini DM, Root DE (2006) A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* 124:1283–1298
- Morgens DW, Deans RM, Li A, Bassik MC (2016) Systematic comparison of CRISPR/Cas9 and RNAi screens for essential genes. *Nat Biotechnol* 34:634–636
- Naldini L, Blomer U, Gage FH, Trono D, Verma IM (1996a) Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *Proc Natl Acad Sci USA* 93:11382–11388
- Naldini L, Blomer U, Gallay P, Ory D, Mulligan R, Gage FH, Verma IM, Trono D (1996b) In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* 272:263–267
- Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152:1173–1183
- Root DE, Hacohen N, Hahn WC, Lander ES, Sabatini DM (2006) Genome-scale loss-of-function screening with a lentiviral RNAi library. *Nat Methods* 3:715–719
- Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, Heckl D, Ebert BL, Root DE, Doench JG, Zhang F (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343:84–87

- Shema R, Kulicke R, Cowley GS, Stein R, Root DE, Heiman M (2015) Synthetic lethal screening in the mammalian central nervous system identifies Gpx6 as a modulator of Huntington's disease. *Proc Natl Acad Sci USA* 112:268–272
- Wang T, Wei JJ, Sabatini DM, Lander ES (2014) Genetic screens in human cells using the CRISPR-Cas9 system. *Science* 343:80–84
- Yee JK, Miyanohara A, LaPorte P, Bouic K, Burns JC, Friedmann T (1994) A general method for the generation of high-titer, pantropic retroviral vectors: highly efficient infection of primary hepatocytes. *Proc Natl Acad Sci USA* 91:9564–9568
- Zhou Y, Zhu S, Cai C, Yuan P, Li C, Huang Y, Wei W (2014) High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature* 509:487–491

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



CRISPR/Cas9-Mediated Knockin and Knockout in Zebrafish

Shahad Albadri, Flavia De Santis, Vincenzo Di Donato,
and Filippo Del Bene

Abstract The zebrafish (*Danio rerio*) has emerged in recent years as a powerful vertebrate model to study neuronal circuit development and function, thanks to its relatively small size, rapid external development and translucency. These features allow the easy application of in vivo microscopy analysis and optical perturbation of neuronal function. So far, genetic manipulation in zebrafish has been limited to the generation of constitutive loss-of-function alleles and transgenic models. CRISPR/Cas9 offers unprecedented possibilities for genomic manipulation that can be exploited to study neuronal function. In the past few years, we have successfully used CRISPR/Cas9-based technology in zebrafish to achieve two goals crucial for neuronal circuit analysis by developing two CRISPR/Cas9-based approaches that overcome previous major limitations to the study of gene and neuron functions in zebrafish. The study of gene function via tissue- or cell-specific mutagenesis remains challenging in zebrafish when the study of the function of certain loci might require tight spatiotemporal control of gene inactivation, which is particularly true in studying the function of a particular gene in post mitotic neurons, when the same gene may have had an earlier developmental function. To circumvent this limitation, we developed a simple and versatile protocol to achieve tissue-specific and temporally controlled gene disruption based on Cas9 expression under the control of the Gal4/UAS binary system (Di Donato et al. 2016). This strategy allows us to induce somatic mutations in genetically labeled cell clones or single cells and to follow them in vivo via reporter gene expression. We have also been able to target endogenous genomic loci to specifically label the great variety of neuronal cell types with reporter genes such as the transcriptional activator Gal4 (Auer et al. 2014). As a result, we can specifically target the expression of fluorescent proteins, a genetically encoded calcium indicator or optogenetic actuators in defined neuronal subpopulations.

We will present ways that these two methods can be applied to the study of the development of the nervous system in larval zebrafish.

Shahad Albadri, Flavia De Santis and Vincenzo Di Donato made equal contribution.

S. Albadri • F. De Santis • V. Di Donato • F. Del Bene (✉)
Institut Curie, PSL Research University, INSERM, U 934, CNRS UMR3215, 75005 Paris,
France
e-mail: Filippo.Del-Bene@curie.fr

CRISPR/Cas9 and Gal4/UAS Combination for Cell-Specific Gene Inactivation

Over the last decades, the analysis of gene function has relied on mutagenesis approaches leading to the generation of loss-of-function alleles. The CRISPR/Cas9 system represents a major step forward towards achieving precise and targeted gene disruption. Being readily applicable for the creation of knockout loci in a great variety of animal models used in neuroscience studies, this technology has led to significant advances in the fields of developmental and functional neurobiology (Heidenreich and Zhang 2016). Nonetheless, constitutive gene disruption is often associated with side effects, such as compensation mechanisms and embryonic lethality, representing an important limitation on the analysis of phenotypes specific to the nervous system, since neural circuits are fully established at late stages of development. Recently, studies in worms (Shen et al. 2014), fruit flies (Port et al. 2014), mice (Platt et al. 2014) and zebrafish (Ablain et al. 2015) have pioneered the use of the CRISPR/Cas9 methodology to generate conditional gene knockouts via tissue-specific expression of *cas9*. This strategy takes advantage of cell type-specific promoters to control the spatiotemporal expression of the Cas9 enzyme. Importantly, one of the most common methodologies ensuring cell-specific expression of transgenes in zebrafish is the Gal4-UAS binary system (derived from yeast), in which the transcription of genes placed 3' of an upstream activating sequence (UAS) relies on the DNA binding of the Gal4 transcriptional activator (Asakawa and Kawakami 2008). Gene- and enhancer-trap methods have been applied to establish a significant number of Gal4 transgenic lines (Davison et al. 2007; Asakawa et al. 2008; Scott and Baier 2009; Kawakami et al. 2010; Balciuniene et al. 2013), several of which are neural-specific (Scott et al. 2007; Asakawa et al. 2008). Notably, in these lines the Gal4 open reading frame (ORF) is randomly integrated in the fish genome through Tol2-based transposition, and the insertion site is not mapped; therefore, the sequence of the promoter elements driving Gal4 expression is unknown. In our work, we have developed a flexible conditional knockout strategy based on the CRISPR/Cas9 technology that combines Gal4/UAS-mediated expression of the Cas9 enzyme with a constitutive expression of sgRNAs driven by PolIII U6 promoter sequences. Our strategy does not require previous knowledge of promoter sequences to induce *cas9* expression since this is provided by cell type-specific Gal4 transcription. Additionally, to enable the analysis of the phenotypes arising from Cas9-induced gene disruption, we marked the population of the *cas9*-expressing cells by using the viral T2A self-cleaving peptide (Provost et al. 2007), ensuring the stoichiometric synthesis of the Cas9 enzyme and the fluorescent reporter GFP from the same mRNA. To test our conditional knockout strategy, we used our vector system to target the *tyrosinase* (*tyr*) locus, coding for a key enzyme involved in melanin production (Camp and Lardelli 2001). We were able to induce eye-specific loss of pigmentation by expressing our transgene exclusively in the progenitors of the neural retina and the retinal-pigmented

epithelium (RPE). For this purpose we used a transgenic line, *Tg(rx2:gal4)*, in which the Gal4 trans-activator is specifically driven in the optic primordium by the promoter of the zebrafish *retinal homeobox* gene 2 (*rx2*; Heermann et al. 2015). This result confirmed the ability of our strategy to induce Gal4- and Cas9-mediated tissue-specific gene inactivation. Remarkably, in this first approach, GFP expression was strictly dependent on the temporal activity of the promoter driving Gal4 expression, thus restricting direct detection of potential mutant cells to a limited time window. This caveat reduces the possibility of analyzing loss-of-function phenotypes after Gal4 transactivation activity has terminated. To circumvent this issue, we proposed to use the activity of the Cre enzyme, a topoisomerase that catalyzes the site-specific recombination of DNA between *loxP* sites (Branda and Dymecki 2004; Pan et al. 2005), to constitutively label the population of *Cas9*-expressing cells. We therefore developed a construct where we substituted the GFP with a Cre reporter, enabling the analysis of gene disruption after Cas9 activity has terminated. The visualization of *cre*-expressing cells is commonly achieved with the use of transgenic lines carrying a cassette where a constitutive promoter drives the expression of a fluorescent reporter upon the Cre-mediated excision of a floxed stop codon. Thus, in cells carrying floxed alleles, the concomitant expression of Cas9 and Cre enzymes by a tissue-specific Gal4 promoter would ensure, respectively, double-strand breaks (DSBs) at the targeted locus as well as the recombination of the floxed locus. Notably, if the Cre-dependent expression of a reporter is constitutive after recombination, all the cells deriving from a *cas9*-expressing progenitor will be fluorescent, allowing long-term visualization of potentially mutated clones of cells. By using our system in retinal stem cells, we successfully disrupted the *atoh7* gene, which is involved in the specification of retinal ganglion cells (RGC) in the developing retina. In this case, we could modify cell fate determination of retinal progenitor cells and generate labeled loss-of-function clones lacking the population of RGC.

Additionally, we employed our method to create genetic chimeras in which single mutant cells could be differentially tagged in a wild-type tissue. To obtain this labeling, we combined the 2C-Cas9 system with the Brainbow technology. The *Tg(UAS:brainbow)* line (Robles et al. 2013) carries a transgene in which the CDSs of the fluorescent proteins tdTomato, Cerulean and YFP are separated by Cre recombinase sites. In double transgenic embryos *Tg(UAS:brainbow) × Tg(Tissue-specific promoter:gal4)*, tdTomato will be expressed in the Gal4 transactivation domain in the absence of Cre-mediated recombination. In contrast, *cerulean* or *YFP* will be transcribed if Cre recombinase is active. The expression of our transgenesis vector in these embryos provides simultaneous activity of the Cas9 and Cre enzymes. As a result, all the Gal4-positive cells that received the plasmid are potentially mutant and marked by cerulean or YFP fluorescence, whereas the population of Gal4-positive cells that do not express the construct is wild-type and labeled with the reporter tdTomato. This multicolor labeling strategy can be easily applied to neurobiology studies to induce targeted mutations in single neurons and directly compare loss-of-function and wild-type phenotypes in the same animal. To test this potential application, we targeted the genomic locus coding for the motor

protein Kinesin family member 5A, a (*kif5aa*) (Campbell and Marlow 2013; Auer et al. 2015), whose inactivation triggers the reduction of RGC axon arbor complexity via a cell-autonomous mechanism (Auer et al. 2015). To target the *kif5aa* gene with the 2C-Cas9 system in single RGC, we used the Tg(*isl2b:gal4*) line. As expected, after injection of our construct into one-cell stage embryos derived from a cross of Tg(*isl2b:gal4*) and Tg(*UAS:brainbow*) fish, we could observe a strong decrease in total branch length in YFP- or Cerulean-expressing RGC (potentially *kif5aa* mutant) compared to tdTomato-fluorescent RGC (wild-type).

In conclusion, the 2C-Cas9 system represents a versatile tool to induce biallelic conditional gene inactivation. The use of the Gal4/UAS system allows the targeting of a gene of choice in any cell population. The combination of this bipartite system with simultaneous activation of Cas9 And Cre enzymes in progenitor or differentiated cells enables first, the genetic lineage tracing of mutant cells and second, the detection of cell-autonomous gene inactivation at single cell resolution. Additionally, permanent labeling of knockout cells offers the possibility of investigating gene function in adult animals, expanding the applicability of the 2C-Cas9 from neurodevelopment to maintenance and function of neural networks. Finally, because the 2C-Cas9 system is based on genetic tools available in several model organisms, this approach allows the same level of investigation in a broad range of animal models.

In addition to the use of the Crispr/Cas9 application for the generation of loss-of-function alleles, RNA guide nucleases can be used for more sophisticated genome modifications such as homologous recombination (HR) or non-homologous end joining (NHEJ)-mediated knockin. We herein provide a conceptual outline of the steps involved in the generation of knockin lines based on the Crispr/Cas9 strategy and the latest advances made in the zebrafish genome-editing field.

Crispr/Cas9-Mediated Knockin Approaches in Zebrafish

With its advantage of transparency, the zebrafish model organism rapidly emerged as a powerful experimental system for studies in genetics, developmental biology and neurobiology. The possible integration of exogenous genes into any given loci and the analysis of their function in the living animal have dramatically improved over the past few years with the development of genome editing technologies. Prior to this recent explosion in the field of knockin generation, conventional transgenic zebrafish lines were generated by Tol2-mediated transgenesis, which has successfully allowed the making of hundreds of new reporter lines essential to the study of particular gene functions in vivo (Davison et al. 2007; Asakawa et al. 2008; Scott and Baier 2009; Kawakami et al. 2010; Balciuniene et al. 2013). Bacterial artificial chromosome-based transgenesis has been and still is one of the go-to methods for making reporter lines. However, this technique comes with one major limitation: the integration of extra coding copies of hundreds of kbs. In addition, it is not known how the integration of such a large construct affects the neighboring site of

insertion. More recently, the transcription activator-like effectors (TALEs) technology, a milestone in the development of zebrafish mutant and transgenic lines, has lifted the limit of loci-specific targeting. With very low off-targeting effects, TALEs were therefore the first successful genome editing method that permitted homologous-directed recombination (HDR) and NHEJ-mediated knockin in zebrafish (Bedell et al. 2012; Zu et al. 2013). Two reports (Chang et al. 2013; Hwang et al. 2013b) showed that double stranded breaks (DSB), which are simpler in design and have higher mutagenesis efficiency, could also be generated using the Crispr/Cas9 technology based on the same approach used by Bedell et al. (2012). Following these studies, Hruscha et al. (2013) achieved the integration of HA-tags into the sequence of single strand oligonucleotides flanked by two short homology arms of the targeted gene. Similarly to previously observed integration events, insertion of the sequences of interest was detected in most targeted alleles with, however, a majority of imprecise and error-prone repair mechanisms. In 2013, Zu et al. reported the first HR gene-targeting event using TALENs and a double stranded vector containing an eGFP cassette flanked by long homology arms and a germ line transmission rate of 1.5%. More recently many other laboratories have developed various methods to generate knockin alleles by HR followed by CRISPR/Cas9-induced DSB, using as donor single stranded DNA, circular or linear plasmids with short (~40 bp) or long (800–1000 bp) homology arms (Hruscha et al. 2013; Hwang et al. 2013a; Irion et al. 2014; Shin et al. 2014; He et al. 2015; Hisano et al. 2015). Although these methods were proven possible, their efficiency remains variable. To circumvent these problems, in 2014 our laboratory employed a strategy taking advantage of homologous independent repair events shown to be tenfold more active than HR events in the one-cell stage embryo (Auer and Del Bene 2014; Auer et al. 2014). The plasmid donor vector was engineered with an eGFP bait cassette and a Gal4 transcriptional transactivator cassette. Co-injected with a locus-specific sgRNA, an eGFP targeting sgRNA and *cas9* nuclease mRNA, cleavage of the donor vector was generated along with the endogenous chromosomal integration site. For better readout, the injection was performed into an outcross of two transgenic lines, the first being an eGFP reporter line and the second a Tg(*UAS:RFP*) line. Injected embryos with a successful in-frame integration event (most probably through homologous independent repair mechanisms) therefore displayed RFP signal in cells where GFP signal was normally detected. In this system, the offspring transmission was evaluated at about 30% and increased to 40% when a selection for the RFP signal was performed after injection. The generation of such a donor vector allowed the direct assessment of the efficiency of the strategy by targeting an endogenous locus of the zebrafish genome. Targeting the transcriptional starting site of the *kif5aa* gene, integration of the donor vector was successfully induced and shown to be independent from the orientation of the sgRNA targeting *kif5aa*. In addition, no homologous sequences between the vector and the endogenous targeted site were required for the integration, allowing the re-use of the vector in combination with any given site-specific sgRNA. Using the same approach, Kimura et al. (2014) improved the strategy by adding a heat shock cassette (*Hsp70*) upstream of the transcription trans-activator Gal4 cassette

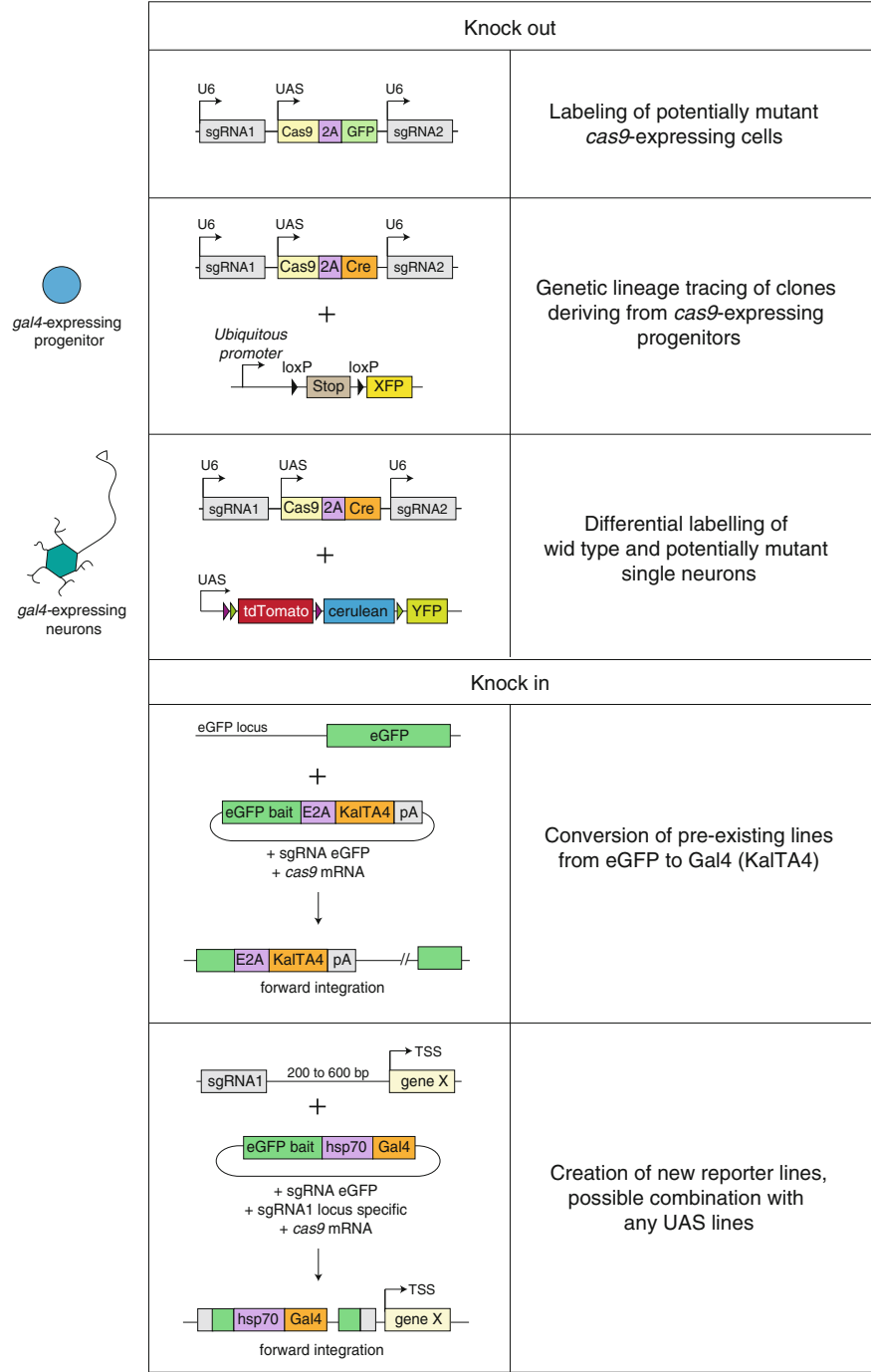


Fig. 1 Knockout and knockin strategies based on the Crispr/Cas9 technology in zebrafish. Schematic representation of the different methods and applications of Crispr/Cas9-mediated genome modifications. From top to bottom: (1) labeling with GFP of *cas9*-expressing cells

into the donor vector, allowing its expression independently from in-frame insertion events within the transcriptional starting site of the gene of interest. To date, several new reporter lines have been generated using this strategy, providing a powerful alternative for homology-independent repair over HR-mediated integration. Key points for its success are (1) the identification of efficient sgRNAs targeting the chromosomal site of choice, for which new prescreening methods have been developed (Carrington et al. 2015; Prykhodzij et al. 2016); (2) the injection of the sgRNA mix with Cas9 nuclease mRNA over purified Cas9 protein that seems to prevent the donor plasmid insertion; and (3) further screening for the identification of founders due to the error-prone nature of junction sites between the endogenous locus and the donor vector. Hisano et al. (2015) addressed this last point by introducing 10–40 bp homology arms into the donor vector to trigger integration events mediated by HR repair mechanisms. In parallel, Li et al. (2015) developed another approach by targeting intronic regions of the gene of interest, therefore non-HR dependent. While this strategy allows keeping the integrity of the targeted coding sequence, the enriched presence of repeat sequences within the introns makes it difficult to achieve a specific targeting. Finally, the latest advance in knockin approaches is the development of traceable genome editing events that allow the easy recovery of edited alleles (Hoshijima et al. 2016) (Fig. 1).

References

- Ablain J, Durand EM, Yang S, Zhou Y, Zon LI (2015) A CRISPR/Cas9 vector system for tissue-specific gene disruption in zebrafish. *Dev Cell* 32:756–764
- Asakawa K, Kawakami K (2008) Targeted gene expression by the Gal4-UAS system in zebrafish. *Dev Growth Differ* 50:391–399
- Asakawa K, Suster ML, Mizusawa K, Nagayoshi S, Kotani T, Urasaki A, Kishimoto Y, Hibi M, Kawakami K (2008) Genetic dissection of neural circuits by Tol2 transposon-mediated Gal4 gene and enhancer trapping in zebrafish. *Proc Natl Acad Sci USA* 105:1255–1260
- Auer TO, Del Bene F (2014) CRISPR/Cas9 and TALEN-mediated knock-in approaches in zebrafish. *Methods* 69:142–150
- Auer TO, Durore K, De Cian A, Concordet JP, Del Bene F (2014) Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res* 24:142–153
- Auer TO, Xiao T, Bercier V, Gebhardt C, Durore K, Concordet J-P, Wyart C, Suster M, Kawakami K, Wittbrodt J, Baier H, Del Bene F (2015) Deletion of a kinesin I motor unmasks a mechanism of homeostatic branching control by neurotrophin-3. *Elife* 4:418

←

Fig. 1 (continued) potentially mutated in locus targeted by the sgRNA1 and sgRNA2 expressed with the PolIII U6 promoters. (2) Genetic labeling with Cre recombinase of *cas9*-expressing cells. Cre activity was revealed by the conditional expression of a fluorescent reporter protein (XFP) after removal of a stop cassette. (3) A similar strategy combined with a brainbow reporter cassette allows the visualization of *cas9*-expressing cells in multiple colors. (4) Genetic knockin of a Gal4 reporter transcription factor into GFP locus of preexisting transgenic lines or (5) into an endogenous genomic location (geneX). UAS upstream activating sequence

- Balciuniene J, Nagelberg D, Walsh KT, Camerota D, Georlette D, Biemar F, Bellipanni G, Balciunas D (2013) Efficient disruption of zebrafish genes using a Gal4-containing gene trap. *BMC Genomics* 14:619
- Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug RG II, Tan W, Penheiter SG, Ma AC, Leung AYH, Fahrenkrug SC, Carlson DF, Voytas DF, Clark KJ, Essner JJ, Ekker SC (2012) In vivo genome editing using a high-efficiency TALEN system. *Nature* 491:114–118
- Branda CS, Dymecki SM (2004) Talking about a revolution: the impact of site-specific recombinases on genetic analyses in mice. *Dev Cell* 6:7–28
- Camp E, Lardelli M (2001) Tyrosinase gene expression in zebrafish embryos. *Dev Genes Evol* 211:150–153
- Campbell PD, Marlow FL (2013) Temporal and tissue specific gene expression patterns of the zebrafish kinesin-1 heavy chain family, kif5s, during development. *Gene Expr Patterns* 13:271–279
- Carrington B, Varshney GK, Burgess SM, Sood R (2015) CRISPR-STAT: an easy and reliable PCR-based method to evaluate target-specific sgRNA activity. *Nucl Acids Res* 43:e157–e157
- Chang N, Sun C, Gao L, Zhu D, Xu X, Zhu X, Xiong J-W, Xi JJ (2013) Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res* 23:465–472
- Davison JM, Akitake CM, Goll MG, Rhee JM, Gosse N, Baier H, Halpern ME, Leach SD, Parsons MJ (2007) Transactivation from Gal4-VP16 transgenic insertions for tissue-specific cell labeling and ablation in zebrafish. *Dev Biol* 304:811–824
- Di Donato V, De Santis F, Auer TO, Testa N, Sánchez-Iranzo H, Mercader N, Concordet J-P, Del Bene F (2016) 2C-Cas9: a versatile tool for clonal analysis of gene function. *Genome Res* 26:681–692
- He MD, Zhang FH, Wang HL, Wang HP, Zhu ZY, Sun YH (2015) Efficient ligase 3-dependent microhomology-mediated end joining repair of DNA double-strand breaks in zebrafish embryos. *Mutat Res* 780:86–96
- Heermann S, Schütz L, Lemke S, Krieglstein K, Wittbrodt J (2015) Eye morphogenesis driven by epithelial flow into the optic cup facilitated by modulation of bone morphogenetic protein. *Elife* 4:373
- Heidenreich M, Zhang F (2016) Applications of CRISPR-Cas systems in neuroscience. *Nat Rev Neurosci* 17:36–44
- Hisano Y, Sakuma T, Nakade S, Ohga R, Ota S, Okamoto H, Yamamoto T, Kawahara A (2015) Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish. *Sci Rep* 5:8841
- Hoshijima K, Jurynek MJ, Grunwald DJ (2016) Precise editing of the zebrafish genome made simple and efficient. *Dev Cell* 36:654–667
- Hruscha A, Krawitz P, Rechenberg A, Heinrich V, Hecht J, Haass C, Schmid B (2013) Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development* 140:4982–4987
- Hwang WY, Fu Y, Reyon D, Maeder ML, Kaini P, Sander JD, Joung JK, Peterson RT, Yeh J-RJ (2013a) Heritable and precise zebrafish genome editing using a CRISPR-Cas system. *PLoS One* 8:e68708
- Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh J-RJ, Joung JK (2013b) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* 31:227–229
- Irion U, Krauss J, Nüsslein-Volhard C (2014) Precise and efficient genome editing in zebrafish using the CRISPR/Cas9 system. *Development* 141:4827–4830
- Kawakami K, Abe G, Asada T, Asakawa K, Fukuda R, Ito A, Lal P, Mouri N, Muto A, Suster ML, Takakubo H, Urasaki A, Wada H, Yoshida M (2010) zTrap: zebrafish gene trap and enhancer trap database. *BMC Dev Biol* 10:105
- Kimura Y, Hisano Y, Kawahara A, Higashijima S-I (2014) Efficient generation of knock-in transgenic zebrafish carrying reporter/driver genes by CRISPR/Cas9-mediated genome engineering. *Sci Rep* 4:6545

- Li J, Zhang BB, Ren YG, SY G, Xiang YH, JL D (2015) Intron targeting-mediated and endogenous gene integrity-maintaining knockin in zebrafish using the CRISPR/Cas9 system. *Cell Res* 25:634–637
- Pan X, Wan H, Chia W, Tong Y, Gong Z (2005) Demonstration of site-directed recombination in transgenic zebrafish using the Cre/loxP system. *Transgenic Res* 14:217–223
- Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, Dahlman JE, Parnas O, Eisenhaure TM, Jovanovic M, Graham DB, Jhunjhunwala S, Heidenreich M, Xavier RJ, Langer R, Anderson DG, Hacohen N, Regev A, Feng G, Sharp PA, Zhang F (2014) CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* 159:440–455
- Port F, Chen H-M, Lee T, Bullock SL (2014) Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in *Drosophila*. *Proc Natl Acad Sci U S A* 111:E2967–E2976
- Provost E, Rhee J, Leach SD (2007) Viral 2A peptides allow expression of multiple proteins from a single ORF in transgenic zebrafish embryos. *Genesis* 45:625–629
- Prykhodzhiy SV, Rajan V, Berman JN (2016) A guide to computational tools and design strategies for genome editing experiments in zebrafish using CRISPR/Cas9. *Zebrafish* 13:70–73
- Robles E, Filosa A, Baier H (2013) Precise lamination of retinal axons generates multiple parallel input pathways in the tectum. *J Neurosci* 33:5027–5039
- Scott EK, Baier H (2009) The cellular architecture of the larval zebrafish tectum, as revealed by gal4 enhancer trap lines. *Front Neural Circ* 3:13
- Scott EK, Mason L, Arrenberg AB, Ziv L, Gosse NJ, Xiao T, Chi NC, Asakawa K, Kawakami K, Baier H (2007) Targeting neural circuitry in zebrafish using GAL4 enhancer trapping. *Nat Publ Group* 4:323–326
- Shen Z, Zhang X, Chai Y, Zhu Z, Yi P, Feng G, Li W, Ou G (2014) Conditional knockouts generated by engineered CRISPR-Cas9 endonuclease reveal the roles of coronin in *C. elegans* neural development. *Dev Cell* 30:625–636
- Shin J, Chen J, Solnica-Krezel L (2014) Efficient homologous recombination-mediated genome engineering in zebrafish using TALE nucleases. *Development* 141:3807–3818
- Zu Y, Tong X, Wang Z, Liu D, Pan R, Li Z, Hu Y, Luo Z, Huang P, Wu Q, Zhu Z, Zhang B, Lin S (2013) TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nat Meth* 10:329–331

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Dissecting the Role of Synaptic Proteins with CRISPR

Salvatore Incontro, Cedric S. Asensio, and Roger A. Nicoll

Abstract A significant step forward in the study of synaptic physiology is the application of single cell genetic modifications. In this landscape, the dissection of the role of single proteins or, more significantly, their subunits and sub-domains has increased enormously the basic knowledge of synaptic function. CRISPR/Cas9 is a recently developed genome-editing tool that can be used to inactivate or modify genes of interest. Its ease of implementation and affordable cost, combined with its high efficiency, make it a very valuable tool to study various biological processes. The application of this technique in addition to previous genetic approaches vastly simplifies and accelerates the study of specific synaptic proteins. Here we illustrate different ways that CRISPR/Cas9 can be used in the study of synaptic properties.

Introduction

Over the last two decades, the combination of pharmacology and genetics has been instrumental in our current understanding of the molecular mechanisms controlling diverse neuronal processes. The development of gene targeting through homologous recombination enabled the generation of knockout (KO) transgenic animals, and this ability to completely inactivate genes of interest for synaptic transmission has provided invaluable information about their function. Although germline gene deletion has dramatically advanced our knowledge, the approach suffers from two main limitations: the deletion can be embryonically lethal if the gene is essential or it can lead to physiological compensation during development, masking the real importance of the studied protein. In addition, the development of transgenic animals represents a significant investment of both cost and time.

The more recent development of RNAi has provided an easier and faster way to inactivate proteins, but use of this technique is limited by the efficiency of

S. Incontro (✉) • R.A. Nicoll

Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA

e-mail: Salvatore.Incontro@ucsf.edu

C.S. Asensio

Department of Biological Sciences, University of Denver, Denver, CO 80120, USA

knockdown. Indeed, in the case of incomplete knockdown, residual protein can lead to serious misinterpretation. In addition, off-target effects present an important concern. Indeed, it has been observed that RNAi manipulation can affect the morphology of single spines (Alvarez et al. 2006), suggesting some general non-specific effects of RNAi in neurons.

More recently, the development of conditional knockout (cKO) technology has offered an interesting alternative to the limitations associated with both germline KO and RNAi approaches. Indeed, the cKO approach relies on the generation of transgenic mice with LoxP sites flanking a gene of interest. The subsequent sparse transfection of Cre recombinase in brain slices derived from these LoxP animals results in the removal of the gene of interest from a few neurons and offers a more controllable way to compare genetically manipulated neurons to controls by dual cell patch clamp (Adesnik et al. 2008; Pluck 1996; Hayashi et al. 2000; Schnell et al. 2002; Sauer and Henderson 1988; Tsien et al. 1996). This method is particularly powerful for studying proteins that are essential for the maintenance of synaptic equilibrium. For example, this genetic inactivation approach has been used successfully to determine the function of single subunits of the excitatory post-synaptic AMPA and NMDA receptors (Lu et al. 2009; Gray et al. 2011) and the role of the different isoforms of the SNARE protein complex machinery at the pre-synapse (Hun et al. 2014; Han et al. 2011; Maximov et al. 2007). Nevertheless, the same time and cost considerations associated with the development of germline KO animals apply to the Cre-LoxP system.

Genome Editing Using CRISPR/Cas9

Genome editing generally relies on the guided activity of endonucleases to generate double-strand breaks at a specific location in the genomic DNA in order to modify it. In eukaryotic cells, there are two main types of DNA repair mechanism following double-strand DNA breaks: non-homologous end joining (Barnes 2001; Lieber 2010) and homologous recombinational repair. Non-homologous end joining is generally accompanied by the loss/gain of nucleotides such as deletions, insertions or nucleotide substitutions in the repaired region, thus often leading to inactivation of the targeted gene. On the other hand, homologous recombination uses the complementary DNA as a template to repair the double-strand DNA breaks. The outcome of this type of repair is generally more precise and controllable, so it can be used to either introduce point mutations or knock-in entire proteins through the use of a repair template.

CRISPR/Cas9 is a recently developed genome-editing technique arising from a bacterial adaptive defense system against invading plasmids or phages. The term CRISPR stands for Clusters of Regularly Interspaced Short Palindromic Repeats. These CRISPR loci are found in bacteria and are composed of partially palindromic non-coding repeats that are separated by non-repetitive spacers of similar length. These repeats and spacers are transcribed into one long RNA transcript that is further processed into smaller CRISPR RNAs by endonucleases encoded by

CRISPR-associated (Cas) genes flanking the CRISPR loci (Ishino et al. 1987; Nakata et al. 1989; Pourcel et al. 2005; Jansen et al. 2002). Each individual CRISPR RNA corresponds to one repetitive unit of the original CRISPR array and will guide Cas nucleases to their target by recognizing the homologous DNA region. To work as a defense mechanism, new spacers deriving from invading plasmids or phages are added to the CRISPR locus (Bolotin et al. 2005; Pourcel et al. 2005). Once transcribed and processed into CRISPR RNAs, these new spacers then serve as memory signatures of past invasions, enabling the bacteria to recognize and cleave foreign DNAs (Makarova et al. 2006).

As a genome-editing tool, the technique relies on the nuclease activity of one of these Cas genes (SpCas9) derived from *Streptococcus pyogenes*. The activity of SpCas9 depends on two of these processed RNAs: a CRISPR RNA and a trans-activating CRISPR RNA, which combine to form an RNA complex. The critical features of this complex are the presence of a double-stranded RNA structure at the 3' end that physically interacts with SpCas9 and a 20-nucleotide sequence at the 5' end, which guides the binding of SpCas9 to the target DNA by homology (Jinek et al. 2012). In addition, the proper targeting of SpCas9 requires the presence of a short sequence of the complementary sequence on the target DNA. This sequence is called the protospacer adjacent motif (PAM) and, in the case of SpCas9, consists of a nucleotide triplet (NGG). Importantly, in the absence of the PAM, Cas9 cannot recognize target sequences even when they are fully complementary to the guide RNA (Sternberg et al. 2014). By engineering chimeric single RNAs consisting of a fusion between the trans-activating CRISPR RNAs and the CRISPR RNAs, it becomes possible to mimic the natural RNA complex and to control the targeting of Cas9 to a specific region of the genome by simply changing the 5' complementary sequence of the RNA complex (Jinek et al. 2012; Jiang et al. 2013). This so-called guide RNA consists of 20 nucleotides complementary to the region of interest, whose only requirement for its design is the presence of a PAM at the 3' end (on the target DNA). As this motif is very frequent in eukaryotic genomes (Wu et al. 2014), it becomes possible to target virtually any gene of interest, making CRISPR/Cas9 a very powerful and promising tool for basic research as well as for potential therapeutic use. Unlike other genome-editing tools requiring the design and generation of specific nucleases for each target site, CRISPR/Cas9 relies on a simple two-component system: Cas9 and a target-specific guide RNA.

Practical Considerations for the Use of CRISPR/Cas9

The careful design of guide RNAs represents one of the key steps in successful use of CRISPR/Cas9. The first step consists of choosing the best region to target within the gene of interest and subsequently scanning this sequence for the presence of PAM motifs. When selecting guide RNAs, it is important to consider the possibility that the non-homologous end joining repair mechanism might lead to in-frame deletions resulting from Cas9 cleavage in position -3 from the PAM (see Figs. 1 and 2). If

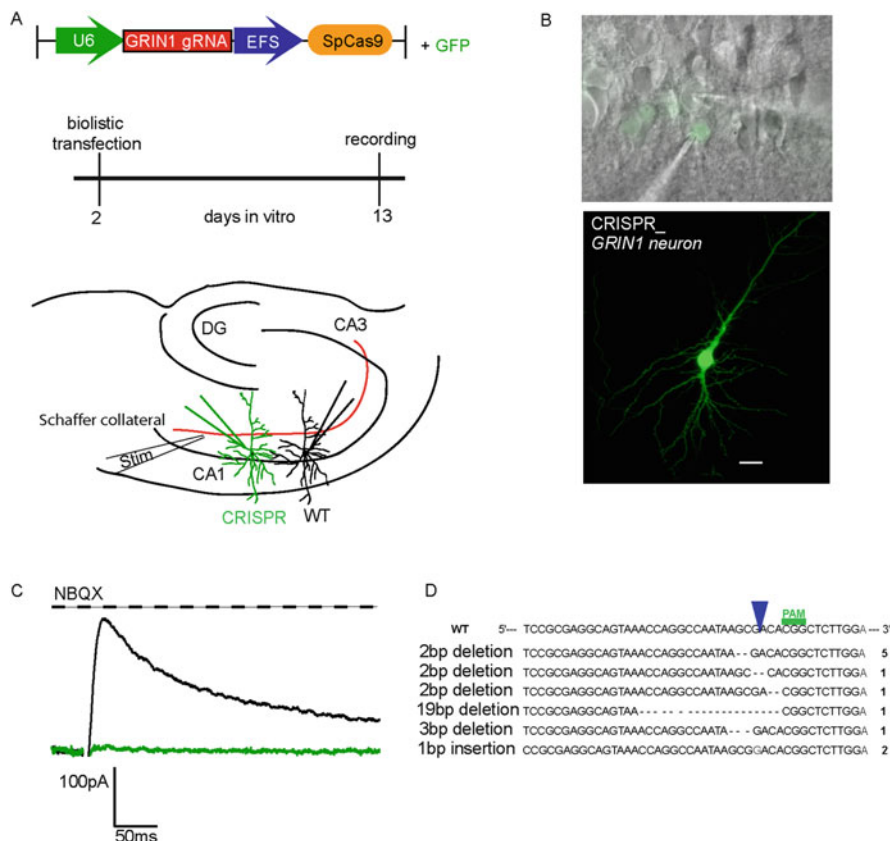


Fig. 1 (a) Timeline of the CRISPR-GRIN1 GluN1 deletion experimentation and scheme of dual whole-cell voltage-clamp recording in organotypic hippocampal slices of a biolistically transfected pX330 CRISPR-GRIN1 neuron and a neighboring wild-type neuron. (b) Representative phase contrast + epifluorescence image of the CA1 region of a hippocampal slice and confocal image of a CRISPR-GRIN1 neuron co-transfected with a FUGW-EGFP plasmid. Scale bar: 20 μ M. (c) Sample traces of NMDAR-evoked EPSCs, from a transfected CRISPR-GRIN1 neuron and a neighboring control in the presence of NBQX (10 μ M). (d) Targeted GRIN1 region and types of insertions or deletions in the DNA after infecting dissociated hippocampal neurons with lentiCRISPR GRIN1 (adapted from Fig. 1 of Incontro et al. 2014)

structural information about the protein is available, it can be used to select a region that is essential for its stability. Unfortunately, for most proteins this information does not exist, and the best strategy to efficiently inactivate the gene of interest is usually to target one of the first exons in order to minimize the chance of generating a truncated, functional protein. When potential 20 bp sequences have been selected, several on-line tools enable users to find sequences with the lowest probabilities for off-target effects based on their lack of similarities to other parts of the genome.

The rescue experiments also provide a powerful tool to assess the role of specific protein domains. By transfecting cDNAs with point mutations or domain deletions, it

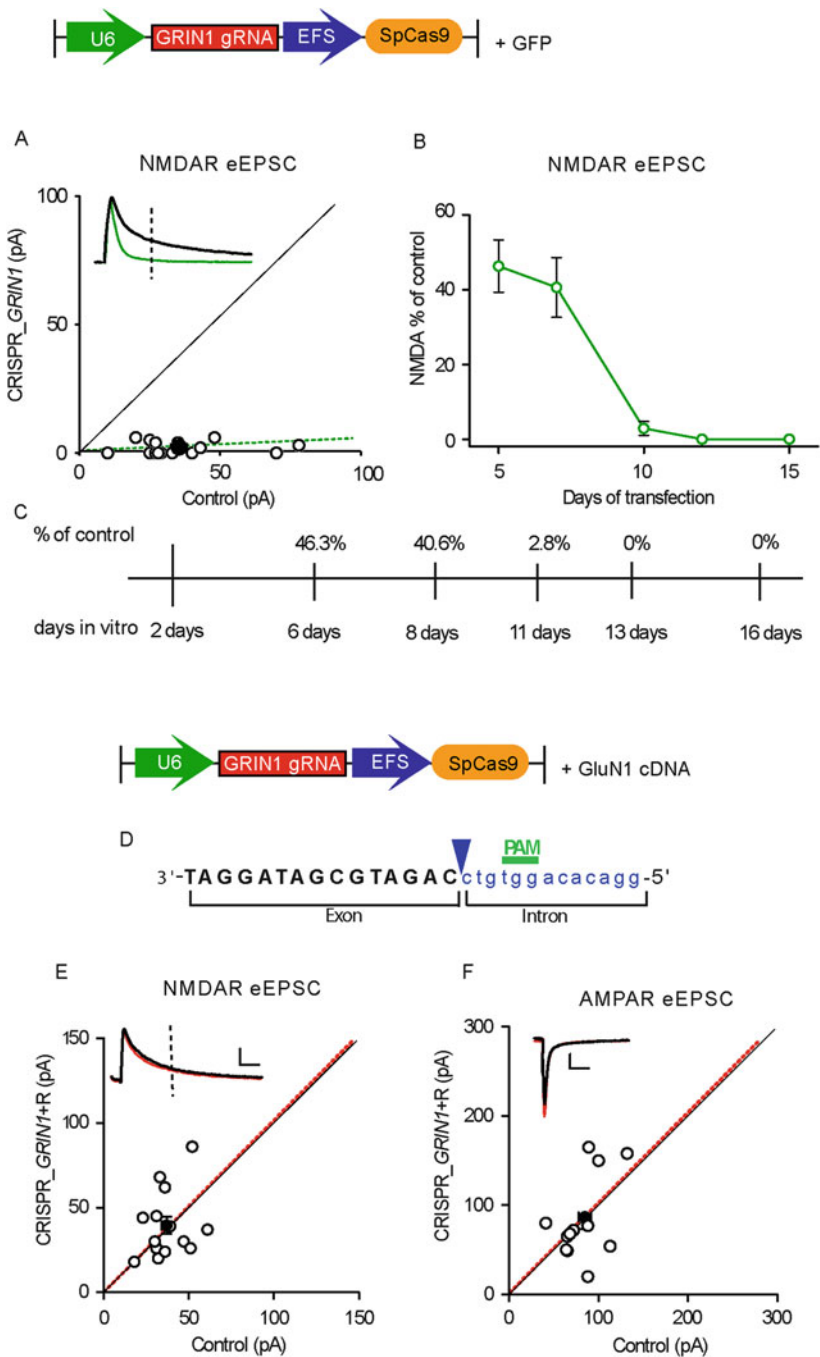


Fig. 2 (a) Scatterplot and sample traces of NMDAR eEPSCs in 14 days transfected CRISPR/Cas9 and neighboring control neurons. *Open circles* represent amplitudes of NMDA EPSCs for single cells; *filled circle* represent the mean. (b) Time course of NMDAR eEPSC 5, 7, 10, and 15 days after transfection. The evoked currents are eliminated after 12 days. (c) Scheme of the time course

becomes possible to assess their significance for the biological process being studied, similar to what has been done with conditional KO animals (Herring et al. 2013).

The Use of CRISPR/Cas9 in Neurons: Proof of Concept

To test the potential of the CRISPR/Cas9 technology in neuroscience, we have performed a proof of concept study aimed at assessing its efficiency to inactivate synaptic proteins. In particular, we focused on two fundamental subunits of the ionotropic glutamate receptors in hippocampal slice cultures: the GluN1 subunit of NMDA receptors and the GluA2 subunit of AMPA receptors. We began by designing two different guide RNAs targeting the extracellular part of the GluN1 subunit, and we selected guide RNAs with a score $>70\%$ corresponding to a low probability of off-target effects according to the MIT online CRISPR design tool. We then co-introduced by biolistic transfection into hippocampal slices a plasmid encoding both Cas9 and one of the guide RNAs together with a plasmid encoding GFP as described previously for cKOs (Adesnik et al. 2008). As the efficiency of this transfection approach is modest, the system as a whole is only minimally perturbed and it becomes possible to directly compare recordings obtained simultaneously from a target, transfected neuron (GFP positive) and a control, untransfected neighbor neuron (GFP negative; Fig. 1a, b). NMDA currents (eEPSCs) were completely abolished in 100% of the pyramidal neurons analyzed (Fig. 1c). Consistent with previous results (Adesnik et al. 2008), we also observed a compensatory increase in AMPA currents. We sequenced the DNA region targeted by Cas9 after PCR amplification of the genomic DNA and found the presence of various small insertions and deletions (indels) creating frameshifts in 90% of the cases (Fig. 1d). This first set of experiments thus suggests that Cas9 is able to efficiently inactivate genes in adult pyramidal neurons by creating double-strand DNA breaks, which are repaired by the non-homologous end joining system. The extreme efficiency that we observed contrasts with the efficiency reported by others using different cell types and is somewhat surprising, but probably reflects the post-mitotic nature of adult pyramidal neurons. In contrast to dividing cells, which rapidly dilute the Cas9 machinery, neurons have the ability to maintain high levels of the CRISPR/Cas9 components for a longer period. Under these conditions, Cas9

Fig. 2 (continued) and percentages of control of NMDAR-evoked EPSCs after CRISPR_*GRIN1* biolistic transfection. **(d)** Scheme of the targeted region in the *GRIN1* gene. The guide RNA not including the PAM region is shown in *bold*; the intronic part of the gene, which includes the PAM region, is shown in *blue*. **(e)** Scatterplot and sample traces of NMDA eEPSCs from a transfected CRISPR_*GRIN1* neuron + GluN1 cDNA and a neighboring control neuron. Scale bar: 50 pA and 50 ms. **(f)** Scatterplot and sample traces of AMPA eEPSCs from a transfected CRISPR_*GRIN1* neuron + the GluN1 cDNA and a neighboring control neuron. Scale bar: 50 pA and 50 ms (Adapted from Figs. 2 and 3 of Incontro et al. 2014)

will presumably have sufficient time to cut the targeted region until it can no longer be properly repaired.

To rule out the existence of off-target effects, we also performed rescue experiments by transfecting a GluN1 cDNA. Re-introduction of the deleted subunit by co-transfection fully rescued the phenotype (Fig. 2).

In a subsequent part of our project, our goal was to target multiplex genes. As a proof of concept we repeated the same experiment with the single GluN1 and GluA2 subunits, this time co-transfecting the two plasmids containing target gRNAs. We observed a complete deletion of both subunits with a complete rectification for AMPA receptors (due to the loss of GluA2) and no NMDA eEPSCs (Fig. 3).

Another issue regards the possibility of studying an effect of a protein deletion *in vivo*. The advent of Cas9 opens a very exciting new concept—we can now inject gRNAs to target potentially any protein in a wild-type (WT) background (co-transfecting with Cas9 plasmids) or in Cas9 knock-in animals (Platt et al. 2014). For example the use of the AMPA receptors triple floxed mouse has been very important for understanding every single subunit's contribution to the structure and function of glutamatergic excitatory synapses. Now we can reproduce these results in a few weeks (compared to years to create Cre-Flox lines and to cross them), optimizing time and cost enormously (Fig. 4).

The possibility of expressing the protein of interest in a KO background enables one to study the function of specific domains in the synaptic context. This approach can be instrumental to the understanding of synaptic proteins that are involved in neurological diseases.

Conclusions and Future Perspectives

The field of biological engineering has seen the rapid development of several novel technologies over the last few years, and neuroscience has embraced many of them to explore the function of synaptic proteins in a more precise and definitive way. Recent development of the CRISPR/Cas9 technology provides a simpler and faster alternative for studying synaptic proteins by removing the time and cost associated with the generation of genetically manipulated animals. Indeed, the approach can be used for the inactivation of target genes, but it also enables one to determine the significance of particular protein domains by performing rescue experiments, as discussed above. In addition, it is possible to place the expression of Cas9 under the control of a neuronal specific promoter for use *in vivo*, similar to what has been done with Cre previously (Gray et al. 2011; Lu et al. 2009; Schnell et al. 2002). Finally, another powerful feature of the CRISPR/Cas9 technology is the ability to easily inactivate several proteins at once using multiplex guide RNAs.

How can one determine that the cleavage has indeed happened? The importance of this validation is best illustrated in a recent short report (Straub et al. 2014) in which the authors performed *in utero* electroporation in mice to inactivate GluN1. Recording from hippocampal slices of 2-week-old mice, they observed a total

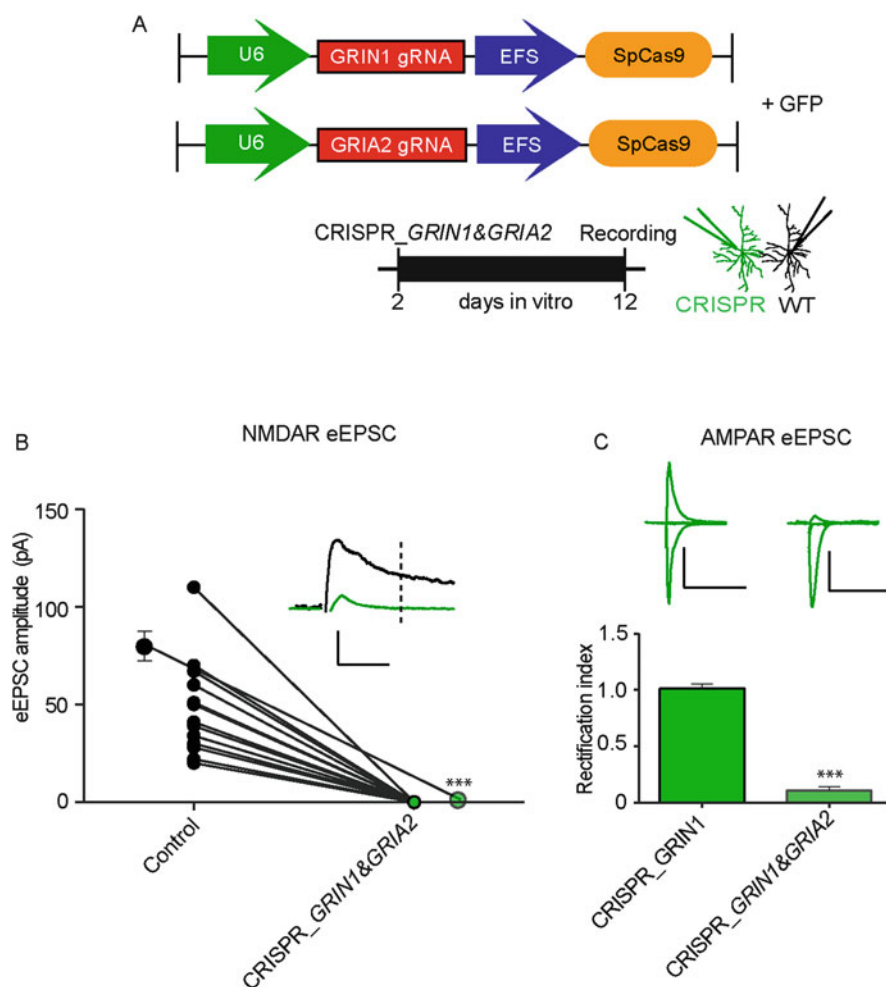


Fig. 3 (a) Scheme of the CRISPR plasmids modified to target specifically *GRIN1* and *GRIA2* and the time course of the transfection period before recording. (b) Sample trace and paired average NMDA-evoked EPSCs of single pairs from control and transfected cells. NMDA currents are completely eliminated after 10 days transfection. (c) AMPAR-evoked EPSCs summary of CRISPR *GRIN1* and double CRISPR *GRIN1*&*GRIA2*. Bar graph indicates the rectification index mean values for the two conditions. The double CRISPR conditions present a fully rectified phenotype typical of GluA1 monomeric receptors

elimination of NMDA currents with one guide RNA whereas the other guide RNA tested had no effect at all. This finding underlines the importance of guide RNA design and the necessity to validate these guide RNAs. In many ways, these considerations are not specific to CRISPR/Cas9 and are also true for Cre-Lox and RNAi approaches.

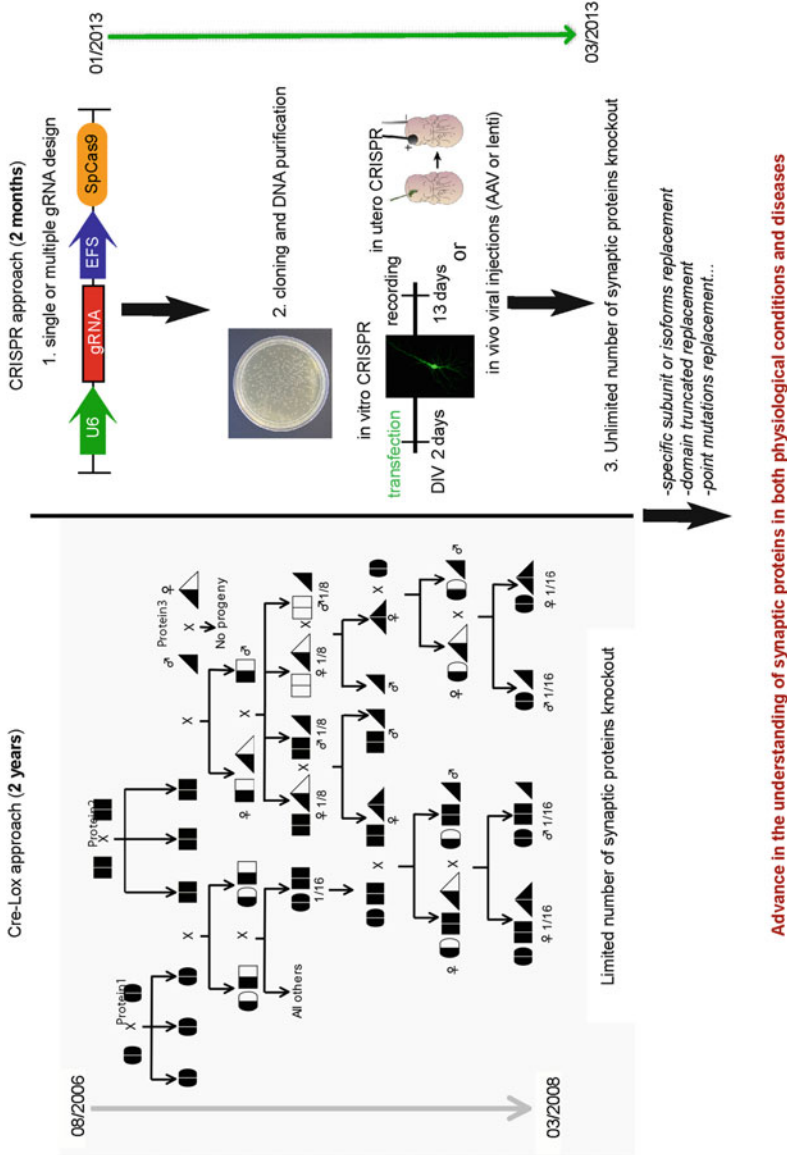


Fig. 4 (a) Scheme of the breeding timeline of triple floxed mice for AMPA receptors. Beyond the time period (2 years) the limit is the number of targets. (b) Example of a CRISPR approach to target a protein. Using different techniques, we can now deliver CRISPR plasmids that can potentially target any synaptic protein

The generation of a Cre-dependent Cas9 knock-in mouse might also become a very useful tool for neuroscientists (Platt et al. 2014). By injecting AAV driving the expression of Cre and of a guide RNA targeting NeuN in the brain of the Cas9 mouse, the authors observed the formation of on-target indels in the infected region accompanied by an 80% reduction in NeuN protein levels. By enabling the inactivation of genes either in vivo or in isolated primary cells, this mouse model will surely serve as a versatile tool and could potentially be used as a platform for genome-wide screens.

Finally, the combination of Cas9 with Sun-TAG technology enables the user to activate the expression of specific genes (Tanenbaum et al. 2014). The system is based on the recruitment of multiple copies of gene regulatory effector domains to a nuclease-deficient CRISPR/Cas9 protein targeted to specific sequences in the genome. CRISPR can thus be used not only to delete synaptic proteins but also to turn on their endogenous expression.

Most recent works on CRISPR/Cas9 systems evidence the importance of an optimized system. In particular modifications are due to the necessity of developing a possible human delivery system containing Cas9. Indeed, the switch to SaCas9 (from *Staphylococcus aureus*), which is much smaller, and the introduction of specific mutations to increase the specificity of Cas9 endonuclease cut, are examples of this race to new drug development (Ran et al. 2015). Furthermore, the introduction of specific mutations in the SpCas9 sequence has significantly enhanced the specificity of this enzyme. Thus, this improvement has reduced to a minimum the possibility of off-target effects, extending the applications of SpCas9 for genome editing (Slaymaker et al. 2016).

The use of CRISPR in neuroscience should be considered simply as a new tool, in particular for the time and cost reduction in the genetic manipulation of synaptic genes. In labs all around the world the introduction of CRISPR may not add anything really new regarding the final result but it can importantly simplify the work (Fig. 4).

Acknowledgments We thank R. H. Edwards, B. E. Herring, and F. Fieni for discussions and comments on the manuscript.

References

- Adesnik H, Li G, During MJ, Pleasure SJ, Nicoll RA (2008) NMDA receptors inhibit synapse unsilencing during brain development. *Proc Natl Acad Sci USA* 105:5597–5602
- Alvarez VA, Ridenour DA, Sabatini BL (2006) Retraction of synapses and dendritic spines induced by off-target effects of RNA interference. *J Neurosci* 26:7820–7825
- Barnes DE (2001) Non-homologous end joining as a mechanism of DNA repair. *Curr Biol* 11: R455–R457
- Bolotin A, Quinquis B, Sorokin A, Ehrlich SD (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151:2551–2561

- Gray JA, Shi Y, Usui H, During MJ, Sakimura K, Nicoll RA (2011) Distinct modes of AMPA receptor suppression at developing synapses by GluN2A and GluN2B: single-cell NMDA receptor subunit deletion in vivo. *Neuron* 71:1085–1101
- Han Y, Kaeser PS, Südhof TC, Schneggenburger R (2011) RIM determines Ca²⁺ channel density and vesicle docking at the presynaptic active zone. *Neuron* 69:304–316
- Hayashi Y, Shi SH, Esteban JA, Piccini A, Poncer JC, Malinow R (2000) Driving AMPA receptors into synapses by LTP and CaMKII: requirement for GluR1 and PDZ domain interaction. *Science* 287:2262–2267
- Herring BE, Shi Y, Suh YH, Zheng CY, Blankenship SM, Roche KW, Nicoll RA (2013) Cornichon proteins determine the subunit composition of synaptic AMPA receptors. *Neuron* 77:1083–1096
- Hun S, Sharma M, Südhof TC, Shen J (2014) Synaptic function of nicastrin in hippocampal neurons. *Proc Natl Acad Sci USA* 111(24):8973–8978
- Incontro S, Asensio CS, Edwards RH, Nicoll RA (2014) Efficient, complete deletion of synaptic proteins using CRISPR. *Neuron* 83:1051–1057
- Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A (1987) Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol* 169:5429–5543
- Jansen R, Embden JD, Gaastra W, Schouls LM (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* 43:1565–1575
- Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* 31:233–239
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816–821
- Lieber MR (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* 79:181–211
- Lu W, Shi Y, Jackson AC, Bjorgan K, During MJ, Sprengel R, Seeburg PH, Nicoll RA (2009) Subunit composition of synaptic AMPA receptors revealed by a single-cell genetic approach. *Neuron* 62:254–268
- Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV (2006) A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct* 16:1–7
- Maximov A, Pang ZP, Tervo DG, Südhof TC (2007) Monitoring synaptic transmission in primary neuronal cultures using local extracellular stimulation. *J Neurosci Methods* 161:75–87
- Nakata A, Amemura M, Makino K (1989) Unusual nucleotide arrangement with repeated sequences in the *Escherichia coli* K-12 chromosome. *J Bacteriol* 171:3553–3556
- Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, Dahlman JE, Parnas O, Eisenhaure TM, Jovanovic M, Graham DB, Jhunjhunwala S, Heidenreich M, Xavier RJ, Langer R, Anderson DG, Hacohen N, Regev A, Feng G, Sharp PA, Zhang F (2014) CRISPR-Cas9 knockin mice for genomic editing and cancer modeling. *Cell* 159:440–455
- Pluck A (1996) Conditional mutagenesis in mice: the Cre/loxP recombination system. *Int J Exp Pathol* 77:269–278
- Pourcel C, Salvignol G, Vergnaud G (2005) CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* 151:653–663
- Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS, Koonin EV, Sharp PA, Zhang F (2015) In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* 520:186–191
- Sauer B, Henderson N (1988) Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc Natl Acad Sci USA* 85:5166–5170

- Schnell E, Sizemore M, Karimzadegan S, Chen L, Bredt DS, Nicoll RA (2002) Direct interactions between PSD-95 and stargazin control synaptic AMPA receptor number. *Proc Natl Acad Sci USA* 99:13902–13907
- Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F (2016) Rationally engineered Cas9 nucleases with improved specificity. *Science* 351:84–88
- Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507:62–67
- Straub C, Granger AJ, Saulnier JL, Sabatini BL (2014) CRISPR/Cas9-mediated gene knock-down in post-mitotic neurons. *PLoS One* 9(8):e105584. doi:[10.1371/journal.pone.0105584](https://doi.org/10.1371/journal.pone.0105584)
- Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD (2014) A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* 159:635–646
- Tsien JZ, Chen DF, Gerber D, Tom C, Mercer EH, Anderson DJ, Mayford M, Kandel ER, Tonegawa S (1996) Subregion- and cell type-restricted gene knockout in mouse brain. *Cell* 87:1317–1326
- Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB, Cheng AW, Trevino AE, Konermann S, Chen S, Jaenisch R, Zhang F, Sharp PA (2014) Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol* 32:670–676

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Recurrently Breaking Genes in Neural Progenitors: Potential Roles of DNA Breaks in Neuronal Function, Degeneration and Cancer

Frederick W. Alt, Pei-Chi Wei, and Bjoern Schwer

Abstract The repair of mammalian DNA double-strand breaks (DSBs) by classical non-homologous end joining (C-NHEJ) suppresses genomic instability and cancer and is required for development of the immune and nervous system. We hypothesize that proper repair of neural DSBs via C-NHEJ or other end-joining pathways is critical for neural functionality and homeostasis over time and that improper DSB repair could contribute to complex psychiatric and neurodegenerative diseases. Here, we summarize various findings made by our laboratory and others over the years that support this hypothesis. This evidence includes, most recently, our discovery of a set of genes, of which most serve neural functions, that can serve as targets of recurrent DSBs in primary neural stem and progenitor cells. We also present a speculative model, based on our findings, of mechanisms by which recurrent DSBs in neural genes can generate neuronal diversity and contribute to neuropsychiatric disease.

Early studies revealed that the lymphocyte-specific V(D)J recombination reaction involves the introduction of DNA double-stranded breaks (DSBs) at the ends of antigen receptor V, D, and J gene segments, followed by the processing of the

F.W. Alt (✉) • P.-C. Wei
Howard Hughes Medical Institute, Boston, MA, USA

Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA, 02115, USA

Department of Genetics, Harvard Medical School, Boston, MA 02115, USA
e-mail: alt@enders.tch.harvard.edu

B. Schwer
Howard Hughes Medical Institute, Boston, MA, USA

Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA, 02115, USA

Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

Department of Neurological Surgery and Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, University of California, San Francisco, CA 94158, USA

generated ends and subsequent fusion of the DSB ends of the different types of gene segments to form V(D)J variable regions exons (Alt and Baltimore 1982). The Baltimore lab discovered the lymphocyte-specific endonuclease (RAG) that generates V(D)J DSBs (Schatz and Swanson 2011). Based on screens of DNA repair-mutant Chinese hamster ovary cell lines, we discovered that the end-joining phase of V(D)J recombination is carried out by a multi-component DSB end-joining pathway (Taccioli et al. 1993). We went on with collaborators to identify many of the various components of the “classical” non-homologous end-joining (C-NHEJ) pathway, including discovering the XRCC4 “core” C-NHEJ factor, based on our finding that this factor restores the ability of a DNA repair-defective Chinese hamster ovary cell line to undergo the joining phase of V(D)J recombination (Li et al. 1995).

To evaluate potential physiological functions of XRCC4 and other C-NHEJ factors newly discovered at the time, or other putative C-NHEJ factors, we inactivated the genes encoding them in mice (Sekiguchi et al. 1999; Ferguson and Alt 2001). Mice in which we inactivated the XRCC4 C-NHEJ factor, or its interaction partner DNA Ligase 4 (Lig4), had essentially identical phenotypes. These phenotypes included, most notably, abrogation of both lymphocyte and neuronal development due to unrepaired DSBs that occurred at the progenitor stage (Frank et al. 1998; Gao et al. 1998). It is striking that the development of lymphocytes and neurons was the most clear-cut defect in these C-NHEJ-deficient mice. As discussed below, XRCC4- or Lig4-deficient mice routinely die late in embryonic development, most likely due to their neuronal developmental defects. At this stage, effects on fetal lymphocyte development can still be assessed.

Lymphocyte development is blocked at the progenitor stages in these core C-NHEJ-deficient backgrounds due to the inability to join V(D)J recombination-associated DSBs generated by the RAG endonuclease in the absence of core C-NHEJ factors (Alt et al. 2013). Thus, progenitor B and T lymphocyte development was completely abrogated due to the inability to, respectively, assemble functional antibody and T cell receptor genes that are needed for further development of the B and T cell lineages. As V(D)J recombination occurs at the G1 cell cycle stage, core C-NHEJ-deficient progenitor lymphocytes correspondingly undergo apoptosis due to a response to their unrepaired V(D)J DSBs that is mediated by the p53 G1 check-point response factor (Frank et al. 2000; Gao et al. 2000; Zhu et al. 2002). In this regard, p53 deficiency, in fact, rescues the embryonic lethality of XRCC4- or Lig4-deficient mice but does not rescue lymphocyte development because V(D)J joining is still abrogated. The alleviation of the p53 response to unrepaired RAG-generated DSBs at antigen receptor genes allows XRCC4- or Lig4-deficient progenitor lymphocytes to survive and enter the cell cycle, resulting in XRCC4/p53-deficient mice that rapidly develop lethal pro-B cell lymphomas (Frank et al. 2000; Gao et al. 2000). These C-NHEJ/p53-deficient pro-B lymphomas all harbor recurrent translocations that fuse RAG-initiated DSBs at the *IgH* locus to DSBs downstream of *c-Myc* (Zhu et al. 2002), with many likely initiated at cryptic RAG off-targets sites in the *c-Myc* downstream region (Hu et al. 2014; Tepsuporn et al. 2014). Notably, however, even though core C-NHEJ-deficient/p53-deficient mice die from recurrent pro-B lymphomas, many of them harbor

medulloblastomas in situ at the time of their death from pro-B lymphoma (Zhu et al. 2002). Finally, conditional inactivation of *Xrcc4* in p53-deficient B cells leads to mature B lymphomas with recurrent translocations involving DSBs initiated by the B cell-specific activation-induced cytidine deaminase (AID) during IgH class switch recombination (CSR, see below) that are joined to upstream regions of the c-Myc gene (Wang et al. 2009).

Our studies demonstrated that XRCC4- or Lig4-deficient neuronal progenitor cells undergo apoptosis throughout the nervous system at a developmental time when particular neuronal progenitor populations differentiate into postmitotic neurons (Gao et al. 1998). Moreover, we implicated p53 checkpoint-initiated apoptosis in response to unrepaired DSBs that occurred in the neuronal progenitors as a mechanism for this death of newly differentiated neurons, as demonstrated by our finding that such neuronal apoptotic death could be rescued by p53 deficiency. In this regard, the postnatal survival of XRCC4-deficient or Lig4-deficient mice conferred by p53 deficiency has been speculated to be due to rescue of newly differentiated neurons with unrepaired DSBs (Sekiguchi et al. 1999). However, the potential effects of such unrepaired DSBs on neuronal functions in these mice could not be assessed due to their rapid death from pro-B cell lymphomas; thus, the potential roles of these implied DSBs in neuronal development and neuronal functions remained speculative. In this regard, a lingering question was the location of the genomic sites of the involved DSBs.

As mentioned above, C-NHEJ/p53 double-deficient mice all develop progenitor B cell lymphomas with recurrent translocations between the IgH and c-Myc genes, whereas p53-deficient mice with *Xrcc4* conditionally inactivated in B-lineage cells develop mature B-lineage tumors with translocations between *IgH* and *c-Myc* but also translocations of other antigen receptor loci (Wang et al. 2008, 2009). Thus, we attempted to identify recurrently breaking genomic sites in neural progenitor cells by conditionally inactivating *Xrcc4* in neuronal stem and progenitor cells in a p53-deficient background. Strikingly, we found that such conditional inactivation of *Xrcc4* in p53-deficient neural progenitors routinely led to medulloblastomas (MBs) with recurrent translocations on several different chromosomes and frequent chromosomal or extrachromosomal amplification of the N-myc gene (Yan et al. 2006). These *N-myc* amplifications were reminiscent of those we found in human neuroblastomas in the process of discovering *N-myc* (Kohl et al. 1983). While the findings supported our original hypothesis that recurrent DSBs in the vicinity of *N-myc* (or other frequently translocated regions in MBs) could predispose to such translocations and amplifications, the resolution available from our studies at that time did not allow mapping of potential fragile break sites.

Together, our prior studies revealed that DSB repair by C-NHEJ in neural stem and progenitor cells (NSPCs) is required for nervous system development and for suppressing childhood brain tumors (Gao et al. 1998; Yan et al. 2006). These studies also raised the interesting possibility of potential parallels between functional outcomes of DSB generation and repair in lymphocytes and neuronal progenitor cells. More recently, studies by others have shown that mature brain cells contain frequent genomic alterations that have been speculated to contribute to

neuronal diversity and disease (McConnell et al. 2013; Poduri et al. 2013; Weissman and Gage 2016). In this regard, beyond inherited germline mutations, somatic, “brain only”, mutations have been implicated in neurodevelopmental and neuropsychiatric disorders (Poduri et al. 2013). However, the potential causes of genomic alterations in brain cells continued to remain largely unexplored and speculative. Based on our observations regarding the effects of C-NHEJ deficiency on neuronal development and neuronal disease, namely cancer, we sought to develop and employ new technologies to test the hypothesis that genomic alterations in mature brain cells and some variations connected to neuropsychiatric diseases might originate from DSBs in NSPCs.

Over the past decade, since our discoveries of the potential roles for DSBs in neuronal diversity and disease, we have developed and enhanced a high-throughput, genome-wide translocation sequencing (HTGTS) approach to rapidly and highly sensitively identify DSBs genome-wide based on their translocation to bait DSBs (Chiarle et al. 2011; Frock et al. 2015; Hu et al. 2016). For this approach, bait DSBs can be introduced ectopically by designer endonucleases (Chiarle et al. 2011; Hu et al. 2014; Meng et al. 2014; Frock et al. 2015) or recurrent endogenous DSBs can be used as bait, including those initiated by AID during IgH CSR (Dong et al. 2015) or by RAG during V(D)J recombination (Zhang et al. 2012; Hu et al. 2015; Zhao et al. 2016).

Our studies have shown that various classes of DSBs, including those induced ectopically by ionizing radiation, show a much greater preference to join to other DSBs within the same topological domain due to proximity effects associated with the spatial genome organization of chromatin domains (Zarrin et al. 2007; Zhang et al. 2012; Alt et al. 2013; Frock et al. 2015). As two random DSBs rarely occur within the relatively short genomic distances within a chromosomal domain, which is often a megabase or less, this phenomenon most greatly impacts the joining of closely linked recurrent DSBs (Alt et al. 2013). Our HTGTS studies provided additional insights into our prior finding (Zarrin et al. 2007; Gostissa et al. 2014) that indicated that CSR joining exploits the predisposition of high frequency DSBs within topological domains to be joined to each other to achieve physiological joining levels (Zarrin et al. 2007; Dong et al. 2015). We also showed that, during V(D)J recombination, RAG exploits chromosomal loop domains to not only achieve high joining frequency but also to developmentally restrict its activity directionally within a loop domain (Hu et al. 2015; Zhao et al. 2016).

To identify the sources and functions of neural DSBs, we applied our HTGTS DSB identification approach to cultured, primary mouse NSPCs. For these HTGTS studies, we employed ectopically generated bait DSBs on several different chromosomes to search for significant, recurrent clusters of DSBs genome-wide that joined to bait DSBs on more than one chromosome. These studies identified 27 recurrent DSB clusters (“RDCs”) in the NSPC genome, all of which were enhanced by mild replication stress via treatment with aphidicolin, a compound that inhibits replication (Wei et al. 2016). Strikingly, all 27 of these RDCs lie within genes, most of which encode surface proteins involved in synaptogenesis and related neural processes (Wei et al. 2016). Moreover, variations of most RDC genes also have been

implicated in neuropsychiatric disorders, including schizophrenia and autism, and many are rearranged in cancers, including brain cancers such as medulloblastoma (Wei et al. 2016; Weissman and Gage 2016). Notably, human counterparts of 9 of the 27 NSPC RDC genes occurred in copy number variations (CNVs) found in individual human frontal cortex neurons (McConnell et al. 2013), suggesting that NSPC RDC DSBs could contribute genomic variations in mature neurons (Wei et al. 2016; Weissman and Gage 2016).

RDC gene transcriptional and replication characteristics suggest that their frequent DSBs could occur during collisions between RNA and DNA polymerases associated with mild replication stress (Wei et al. 2016). RDC gene DSBs appear to occur very frequently across the body of RDC genes, which generally are very long (up to 2 Mb in length) with relatively small exons and which also potentially often lie within topological domains (Wei et al. 2016). As HTGTS maps only those bait DSBs that translocate, local RDC DSB frequency may be much higher than the estimated minimal frequency of 12 RDC translocations per NSPC that we estimated via translocation junction capture via HTGTS (Wei et al. 2016). Indeed, we have estimated that the frequency of DSBs across long RDC genes, while of lower density than CSR DSBs, approach the same order of magnitude in numbers per gene as CSR DSBs in B lymphocytes during IgH CSR (Wei et al. 2016). Notably, because most of the RDC gene sequences are within introns, most of the RDC DSBs also occur within introns as opposed to within exons (Wei et al. 2016).

By analogy to mechanisms of lymphocyte-specific recombination (Dong et al. 2015; Hu et al. 2015), we propose that many DSBs that occur within RDC genes would be joined to other DSBs within the same RDC gene (Wei et al. 2016). Thus, we further propose that frequent RDC gene DSBs, which again mostly occur within introns, may be joined to shuffle exons and, thereby, contribute to neural cell diversity (Fig. 1). Such breakage and joining events may also have the potential of contributing to disease-associated neural gene alterations (Wei et al. 2016; Weissman and Gage 2016).

A number of RDC genes, for example, the neurexins (Treutlein et al. 2014), are thought to produce numerous isoforms via differential RNA processing. Beyond such a diversification mechanism, we propose that RDC-based recombination, by generating exon deletions, might “hard-wire” expression of variant RDC products in NSPCs and, thereby, contribute to neural diversity. Our current findings suggest that such putative activities would occur in NSPCs and the products of recombination events would be carried on into mature neurons; in this regard, the process would be somewhat analogous to V(D)J recombination. However, the actual exon shuffling mechanism we propose would be more similar to IgH CSR, creating different isoforms of the protein rather than creating new exons (Fig. 1). In this scenario, the evolution of long, neural genes that are largely comprised of intronic sequences into which are embedded small exons (Smith et al. 2006) could have evolved to provide large target introns for more random stress-associated DSBs in NSPC development. This would be a different solution to the problem of targeted exon shuffling than that employed by CSR, in which DSBs are introduced into specialized intronic switch region sequences (Fig. 1). Whether or not the processes

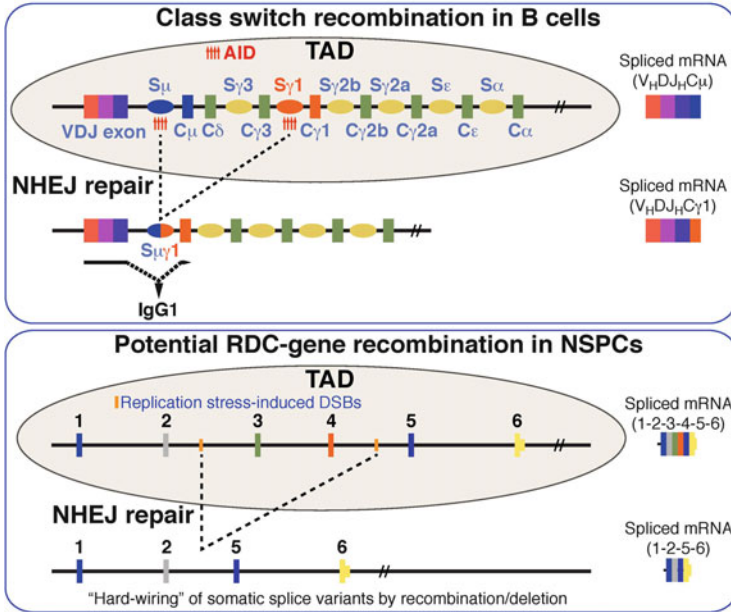


Fig. 1 *Top panel* Diagram of the *IgH* class switch recombination reaction as illustrated by switching from IgM to IgG1. The *IgH* locus is contained within a topological domain (TAD). In activated B cells, switching from IgM to IgG1 results from an exon shuffling process in which the V(D)J exon is first expressed with C_μ to generate IgM but, upon activation, DSBs initiated by AID in repetitive switch (S) regions upstream of C_μ and C_γ1 are joined by C-NHEJ to delete C_μ and replace it with C_γ1. This recombination/deletion exon shuffling process allows the same V(D)J exon to be expressed with a different C exon (For other details, see text or Alt et al. 2013). *Bottom Panel* Diagram of a hypothetical RDC DSB-based exon shuffling mechanism to allow expression of different isoforms of RDC genes to be expressed by “hardwiring” potential somatic splice variants by deletional recombination. This model is based on the finding that at least some RDC genes lie within TADs and that RDC DSB frequency upon replication stress may approach that of *IgH* S regions, allowing ends of different RDC DSBs within the same gene to be frequently joined, based on their proximity within the same topological domain. This model could offer one explanation for why many neural genes are very large and embedded with relatively small exons (Smith et al. 2006): namely, as these genes are mostly comprised of intronic sequences, most “randomly” introduced RDC DSBs across them fall within intronic sequences rather than in exons, providing a basis for a replication stress-associated DSB diversification mechanism. If so, whether or not requisite replication stress is somehow programmed during NSPC development remains to be addressed (See text or Wei et al. 2016 for other details)

that generate RDC genes are specialized to the neural lineage will require further investigation, as will the question of whether enhanced replication stress at the stem and progenitor development stages during neural development could, via an RDC-based mechanism, contribute to neural disease.

RDCs also potentially provide a mechanistic basis for many common fragile sites and certain CNVs, which may result from transcription/replication collisions in generating DSBs or other lesions (Glover and Wilson 2016; Wei et al. 2016). Two

NSPC-RDC genes, *CDH13* and *NRXN3*, are within recurrent CNVs in human MBs (Northcott et al. 2012; Rausch et al. 2012) and several candidate RDCs lie proximal to mouse *N-myc* (Wei, Schwer and Alt, unpublished data). It is possible that RDCs contribute to recurrent genomic variations we and others have found in MBs (Yan et al. 2006), which may offer a mechanism to support the speculation from long ago that proximal, recurrent DSBs during neuroblast differentiation contribute to *N-myc* amplification in human neuroblastomas (Kohl et al. 1983). A number of the 27 identified NSPC RDC-genes undergo somatic genomic rearrangements, including deletions, amplifications, and translocations in various types of cancer (see Wei et al. 2016), and some undergo CNVs in embryonic stem cells and fibroblasts (Wilson et al. 2015; Glover and Wilson 2016). Our HTGTS analysis of additional cell types could identify potential spontaneous or replication stress-induced RDCs in other cell types and, more generally, could shed light on the mechanisms underlying the genetic variations in a range of cancers.

Acknowledgments This work was supported by the Porter Anderson Fund from Boston Children's Hospital and the Howard Hughes Medical Institute. P.W. was supported by a National Cancer Center postdoctoral fellowship. B.S. is a Martin D. Abeloff Scholar of The V Foundation for Cancer Research and is supported by NIA/NIH grant K01AG043630.

References

- Alt FW, Baltimore D (1982) Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D-JH fusions. *Proc Natl Acad Sci USA* 79:4118–4122
- Alt FW, Zhang Y, Meng FL, Guo C, Schwer B (2013) Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell* 152:417–429
- Chiarle R, Zhang Y, Frock RL, Lewis SM, Molinie B, Ho YJ, Myers DR, Choi VW, Compagno M, Malkin DJ, Neuberger D, Monti S, Giallourakis CC, Gostissa M, Alt FW (2011) Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell* 147:107–119
- Dong J, Panchakshari RA, Zhang T, Zhang Y, Hu J, Volpi SA, Meyers RM, Ho YJ, Du Z, Robbiani DF, Meng F, Gostissa M, Nussenzweig MC, Manis JP, Alt FW (2015) Orientation-specific joining of AID-initiated DNA breaks promotes antibody class switching. *Nature* 525:134–139
- Ferguson DO, Alt FW (2001) DNA double strand break repair and chromosomal translocation: lessons from animal models. *Oncogene* 20:5572–5579
- Frank KM, Sekiguchi JM, Seidl KJ, Swat W, Rathbun GA, Cheng HL, Davidson L, Kangaloo L, Alt FW (1998) Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV. *Nature* 396:173–177
- Frank KM, Sharpless NE, Gao Y, Sekiguchi JM, Ferguson DO, Zhu C, Manis JP, Horner J, DePinho RA, Alt FW (2000) DNA ligase IV deficiency in mice leads to defective neurogenesis and embryonic lethality via the p53 pathway. *Mol Cell* 5:993–1002
- Frock RL, Hu J, Meyers RM, Ho YJ, Kii E, Alt FW (2015) Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat Biotechnol* 33:179–186
- Gao Y, Sun Y, Frank KM, Dikkes P, Fujiwara Y, Seidl KJ, Sekiguchi JM, Rathbun GA, Swat W, Wang J, Bronson RT, Malynn BA, Bryans M, Zhu C, Chaudhuri J, Davidson L, Ferrini R,

- Stamato T, Orkin SH, Greenberg ME, Alt FW (1998) A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell* 95:891–902
- Gao Y, Ferguson DO, Xie W, Manis JP, Sekiguchi J, Frank KM, Chaudhuri J, Horner J, DePinho RA, Alt FW (2000) Interplay of p53 and DNA-repair protein XRCC4 in tumorigenesis, genomic stability and development. *Nature* 404:897–900
- Glover TW, Wilson TE (2016) Molecular biology: breaks in the brain. *Nature* 532:46–47
- Gostissa M, Schwer B, Chang A, Dong J, Meyers RM, Marecki GT, Choi VW, Chiarle R, Zarrin AA, Alt FW (2014) IgH class switching exploits a general property of two DNA breaks to be joined in cis over long chromosomal distances. *Proc Natl Acad Sci USA* 111:2644–2649
- Hu J, Tepsuporn S, Meyers RM, Gostissa M, Alt FW (2014) Developmental propagation of V(D)J recombination-associated DNA breaks and translocations in mature B cells via dicentric chromosomes. *Proc Natl Acad Sci USA* 111:10269–10274
- Hu J, Zhang Y, Zhao L, Frock RL, Du Z, Meyers RM, Meng FL, Schatz DG, Alt FW (2015) Chromosomal loop domains direct the recombination of antigen receptor genes. *Cell* 163:947–959
- Hu J, Meyers RM, Dong J, Panchakshari RA, Alt FW, Frock RL (2016) Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat Protoc* 11:853–871
- Kohl NE, Kanda N, Schreck RR, Bruns G, Latt SA, Gilbert F, Alt FW (1983) Transposition and amplification of oncogene-related sequences in human neuroblastomas. *Cell* 35:359–367
- Li Z, Otevrel T, Gao Y, Cheng HL, Seed B, Stamato TD, Taccioli GE, Alt FW (1995) The XRCC4 gene encodes a novel protein involved in DNA double-strand break repair and V(D)J recombination. *Cell* 83:1079–1089
- McConnell MJ, Lindberg MR, Brennand KJ, Piper JC, Voet T, Cowing-Zitron C, Shumilina S, Lasken RS, Vermeesch JR, Hall IM, Gage FH (2013) Mosaic copy number variation in human neurons. *Science* 342:632–637
- Meng FL, Du Z, Federation A, Hu J, Wang Q, Kieffer-Kwon KR, Meyers RM, Amor C, Wasserman CR, Neuberg D, Casellas R, Nussenzweig MC, Bradner JE, Liu XS, Alt FW (2014) Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. *Cell* 159:1538–1548
- Northcott PA, Shih DJ, Peacock J, Garzia L, Morrissy AS, Zichner T, Stütz AM, Korshunov A, Reimand J, Schumacher SE, Beroukhir R, Ellison DW, Marshall CR, Lionel AC, Mack S, Dubuc A, Yao Y, Ramaswamy V, Luu B, Rolider A, Cavalli FM, Wang X, Remke M, Wu X, Chiu RY, Chu A, Chuah E, Corbett RD, Hoad GR, Jackman SD, Li Y, Lo A, Mungall KL, Nip KM, Qian JQ, Raymond AG, Thiessen NT, Varhol RJ, Birol I, Moore RA, Mungall AJ, Holt R, Kawauchi D, Roussel MF, Kool M, Jones DT, Witt H, Fernandez-L A, Kenney AM, Wechsler-Reya RJ, Dirks P, Aviv T, Grajkowska WA, Perek-Polnik M, Haberler CC, Delattre O, Reynaud SS, Doz FF, Pernet-Fattet SS, Cho BK, Kim SK, Wang KC, Scheurlen W, Eberhart CG, Fèvre-Montange M, Jouvét A, Pollack IF, Fan X, Muraszko KM, Gillespie GY, Di Rocco C, Massimi L, Michiels EM, Kloosterhof NK, French PJ, Kros JM, Olson JM, Ellenbogen RG, Zitterbart K, Kren L, Thompson RC, Cooper MK, Lach B, McLendon RE, Bigner DD, Fontebasso A, Albrecht S, Jabado N, Lindsey JC, Bailey S, Gupta N, Weiss WA, Bognár L, Klekner A, Van Meter TE, Kumabe T, Tominaga T, Elbabaa SK, Leonard JR, Rubin JB, Liao LM, Van Meir EG, Fouladi M, Nakamura H, Cinalli G, Garami M, Hauser P, Saad AG, Iolascon A, Jung S, Carlotti CG, Vibhakhar R, Ra YS, Robinson S, Zollo M, Faria CC, Chan JA, Levy ML, Sorensen PH, Meyerson M, Pomeroy SL, Cho YJ, Bader GD, Tabori U, Hawkins CE, Bouffet E, Scherer SW, Rutka JT, Malkin D, Clifford SC, Jones SJ, Korbel JO, Pfister SM, Marra MA, Taylor MD (2012) Subgroup-specific structural variation across 1,000 medulloblastoma genomes. *Nature* 488:49–56
- Poduri A, Evrony GD, Cai X, Walsh CA (2013) Somatic mutation, genomic variation, and neurological disease. *Science* 341:1237758
- Rausch T, Shih DJ, Peacock J, Garzia L, Morrissy AS, Zichner T, Stütz AM, Korshunov A, Reimand J, Schumacher SE, Beroukhir R, Ellison DW, Marshall CR, Lionel AC, Mack S, Dubuc A, Yao Y, Ramaswamy V, Luu B, Rolider A, Cavalli FM, Wang X, Remke M, Wu X, Chiu RY, Chu A, Chuah E, Corbett RD, Hoad GR, Jackman SD, Li Y, Lo A, Mungall

- KL, Nip KM, Qian JQ, Raymond AG, Thiessen NT, Varhol RJ, Birol I, Moore RA, Mungall AJ, Holt R, Kawachi D, Roussel MF, Kool M, Jones DT, Witt H, Fernandez-L A, Kenney AM, Wechsler-Reya RJ, Dirks P, Aviv T, Grajkowska WA, Perek-Polnik M, Haberler CC, Delattre O, Reynaud SS, Doz FF, Pernet-Fattet SS, Cho BK, Kim SK, Wang KC, Scheurlen W, Eberhart CG, Fèvre-Montange M, Jouvét A, Pollack IF, Fan X, Muraszko KM, Gillespie GY, Di Rocco C, Massimi L, Michiels EM, Kloosterhof NK, French PJ, Kros JM, Olson JM, Ellenbogen RG, Zitterbart K, Kren L, Thompson RC, Cooper MK, Lach B, McLendon RE, Bigner DD, Fontebasso A, Albrecht S, Jabado N, Lindsey JC, Bailey S, Gupta N, Weiss WA, Bognár L, Klekner A, Van Meter TE, Kumabe T, Tominaga T, Elbabaa SK, Leonard JR, Rubin JB, Liau LM, Van Meir EG, Fouladi M, Nakamura H, Cinalli G, Garami M, Hauser P, Saad AG, Iolascon A, Jung S, Carlotti CG, Vibhakkar R, Ra YS, Robinson S, Zollo M, Faria CC, Chan JA, Levy ML, Sorensen PH, Meyerson M, Pomeroy SL, Cho YJ, Bader GD, Tabori U, Hawkins CE, Bouffett E, Scherer SW, Rutka JT, Malkin D, Clifford SC, Jones SJ, Korbel JO, Pfister SM, Marra MA, Taylor MD (2012) Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. *Cell* 148:59–71
- Schatz DG, Swanson PC (2011) V(D)J recombination: mechanisms of initiation. *Annu rev Genet* 45:167–202
- Sekiguchi JM, Gao Y, Gu Y, Frank K, Sun Y, Chaudhuri J, Zhu C, Cheng HL, Manis J, Ferguson D, Davidson L, Greenberg ME, Alt FW (1999) Nonhomologous end-joining proteins are required for V(D)J recombination, normal growth, and neurogenesis. *Cold Spring Harb Symp Quant Biol* 64:169–181
- Smith DI, Zhu Y, McAvoy S, Kuhn R (2006) Common fragile sites, extremely large genes, neural development and cancer. *Cancer Lett* 232:48–57
- Taccioli GE, Rathbun G, Oltz E, Stamato T, Jeggo PA, Alt FW (1993) Impairment of V(D)J recombination in double-strand break repair mutants. *Science* 260:207–210
- Tepsuporn S, Hu J, Gostissa M, Alt FW (2014) Mechanisms that can promote peripheral B-cell lymphoma in ATM-deficient mice. *Cancer Immunol res* 2:857–866
- Treutlein B, Gokce O, Quake SR, Sudhof TC (2014) Cartography of neuroligin alternative splicing mapped by single-molecule long-read mRNA sequencing. *Proc Natl Acad Sci USA* 111: E1291–E1299
- Wang JH, Alt FW, Gostissa M, Datta A, Murphy M, Alimzhanov MB, Coakley KM, Rajewsky K, Manis JP, Yan CT (2008) Oncogenic transformation in the absence of Xrcc4 targets peripheral B cells that have undergone editing and switching. *J Exp med* 205:3079–3090
- Wang JH, Gostissa M, Yan CT, Goff P, Hickernell T, Hansen E, Difilippantonio S, Wesemann DR, Zarrin AA, Rajewsky K, Nussenzweig A, Alt FW (2009) Mechanisms promoting translocations in editing and switching peripheral B cells. *Nature* 460:231–236
- Wei PC, Chang AN, Kao J, Du Z, Meyers RM, Alt FW, Schwer B (2016) Long neural genes harbor recurrent DNA break clusters in neural stem/progenitor cells. *Cell* 164:644–655
- Weissman IL, Gage FH (2016) A mechanism for somatic brain mosaicism. *Cell* 164:593–595
- Wilson TE, Arlt MF, Park SH, Rajendran S, Paulsen M, Ljungman M, Glover TW (2015) Large transcription units unify copy number variants and common fragile sites arising under replication stress. *Genome res* 25:189–200
- Yan CT, Kaushal D, Murphy M, Zhang Y, Datta A, Chen C, Monroe B, Mostoslavsky G, Coakley K, Gao Y, Mills KD, Fazeli AP, Tepsuporn S, Hall G, Mulligan R, Fox E, Bronson R, De Girolami U, Lee C, Alt FW (2006) XRCC4 suppresses medulloblastomas with recurrent translocations in p53-deficient mice. *Proc Natl Acad Sci USA* 103:7378–7383
- Zarrin AA, Del Vecchio C, Tseng E, Gleason M, Zarin P, Tian M, Alt FW (2007) Antibody class switching mediated by yeast endonuclease-generated DNA breaks. *Science* 315:377–381
- Zhang Y, McCord RP, Ho YJ, Lajoie BR, Hildebrand DG, Simon AC, Becker MS, Alt FW, Dekker J (2012) Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell* 148:908–921

- Zhao L, Frock RL, Du Z, Hu J, Chen L, Krangel MS, Alt FW (2016) Orientation-specific RAG activity in chromosomal loop domains contributes to Tcrd V(D)J recombination during T cell development. *J Exp med* 213:1921–1936
- Zhu C, Mills KD, Ferguson DO, Lee C, Manis J, Fleming J, Gao Y, Morton CC, Alt FW (2002) Unrepaired DNA breaks in p53-deficient cells lead to oncogenic gene amplification subsequent to translocations. *Cell* 109:811–821

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Neuroscience Research Using Non-human Primate Models and Genome Editing

Noriyuki Kishi and Hideyuki Okano

Abstract The common marmoset (*Callithrix jacchus*) is a small New World non-human primate indigenous to northeastern Brazil. This species has been attracting the attention of biomedical researchers and neuroscientists for its ease of handling and colony maintenance, unique behavioral characteristics, and several human-like traits, such as enriched social vocal communication and strong relationships between parents and offspring. Its high reproductive efficiency makes it particularly amenable for use in the development of transgenic and genome editing technologies in a non-human primate model. Our group has recently generated transgenic marmosets with germ line transmission, opening new avenues in primate research.

In this chapter, we describe recent advances in neuroscience and disease research using common marmosets, and we outline potential uses of genome editing in non-human primates toward the development of knock-in/knock-out marmosets.

Introduction

Rodent models have long played important roles in neuroscience and medical research, made possible in part by the advent of robust genetic technologies. Knock-out/knock-in mouse models have shown particular utility in the neurosciences. There are nonetheless substantial anatomical, physiological, and cognitive differences between rodents and humans. The human brain consists of two major functional domains, one that is evolutionarily conserved and a second that is primate-specific and the locus of many higher cognitive functions. For many human neurological and psychiatric diseases involving higher cognitive dysfunctions, studies using rodent models may thus not be informative with respect to the relevant pathophysiological mechanisms. To gain a better understanding of the

N. Kishi • H. Okano (✉)

Laboratory for Marmoset Neural Architecture, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan

Department of Physiology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan

e-mail: hidokano@a2.keio.jp

pathogenesis of such diseases, we need animal models that exhibit brain functions more closely similar to those in humans.

This need has led to increased interest in the development of genetically engineered non-human primates for use in the study of both functional domains. Our group has recently generated a transgenic common marmoset, a New World monkey (Sasaki et al. 2009). Emerging genome editing techniques are also opening new possibilities for the creation of better non-human primate models for use in the study of neurodegenerative and mental disorders (Izpisua Belmonte et al. 2015).

This chapter is an updated and modified version of previously published review articles on marmosets (Okano et al. 2016; Kishi et al. 2014) and work presented by Hideyuki Okano at the “Genome Editing in Neurosciences” symposium.

Characteristics of the Common Marmoset

Common marmosets (*Callithrix jacchus*) are New World primates native to the Atlantic coastal forests of northeastern Brazil (Abbott et al. 2003; Carrion and Patterson 2012; Mansfield 2003; Okano et al. 2016; Tokuno et al. 2012; Kishi et al. 2014; Izpisua Belmonte et al. 2015). These small monkeys (adult height: 20–30 cm; weight: 350–400 g) have ear tufts and relatively long banded tails, and they are omnivorous, eating plant exudates, lizards, and infant mammals. Common marmosets are monogamous and, unlike many other non-human primates, live in stable families of approximately ten members (Tardif et al. 2003). Females commonly give birth to two babies per litter and are ready to breed again about 10 days after giving birth; they typically have two litters per year. Since mothers need to nurse infants during gestation and the perinatal period, the male partner and other members of the group also provide infant care. This remarkably human-like trait is a focus of attention among neuroscientists and behavioral scientists.

Although common marmosets have been used for biomedical research since the 1960s, macaque monkeys are more widely used in research, due to their closer similarity to humans. The recent rapid advances in genome editing are now calling new attention to the advantages offered by the marmoset because of its size, availability, and high reproductivity.

Macaques are evolutionarily closer to humans than common marmosets, but some marmoset traits are more similar to those of humans, perhaps due either to geographical segregation or convergent evolution. New World primates are estimated to have diverged from Old World primates ~35 mya, and these monkeys have adapted to neotropical environments. Despite this phylogenetic distance, common marmosets, like humans, exhibit strong intergenerational kin relationships and social vocal communications (Dell’Mour et al. 2009; Eliades and Wang 2008; Gordon and Rogers 2010), which may indicate a convergent trajectory in their evolution. The genomic basis of the origins of such traits may be addressable through genome editing studies in the future.

Advantages of Using Common Marmosets for Biomedical Research

Rodents play a crucial role in biomedical investigations in many research fields. Powerful genetic tools, such as knock-out/knock-in mice, have informed the study of gene functions, but the significant anatomical and physiological differences between rodents and humans mean that a more closely similar animal model is needed to advance our understanding of human biology in areas such as the neurosciences.

For biomedical use, the common marmoset offers many advantages. Marmoset endocrinology and metabolism are more similar to those of humans than of rodents, which is important in pharmacological and toxicological studies of new drug candidates. The marmoset is also more closely phylogenetically related to humans (Kitamura et al. 2011; t'Hart et al. 2003, 2012). In Europe, the marmoset is now being used as a non-rodent second species in drug safety tests (Smith et al. 2001).

The common marmoset can be handled with greater ease than many other non-human primates. Along with the appropriateness of the model to the research question, animal welfare and availability are important factors in selecting a model species. Marmosets are readily obtained for laboratory use and, as distinct from macaques, have not been reported to carry herpes b virus (*Macacine herpesvirus 1*), providing a safety benefit to researchers and animal facility staff (Mansfield 2003). The small size of marmosets is also beneficial as it reduces costs and floor space requirements (Smith et al. 2001).

Common marmosets are among the most highly reproductive of all primates. The ovarian cycle is approximately 28 days, similar to that in human (Summers et al. 1985). The gestation period is approximately 145–148 days. Female animals are ready to breed again 10 days after delivery. Usually, female marmosets have two litters per year, which is strongly advantageous when compared to macaques, which require 5 years to sexual maturation and breed only once per year (Austad and Fischer 2011). The remarkable reproductive efficiency of marmosets is extremely well-suited to the development of transgenic and genome editing techniques.

Lastly, a number of basic research tools have been developed for use in marmosets, which is important for encouraging broader adoption by the scientific community. Although the annotated sequencing of its genome has not been completed, a draft sequence with 6× coverage using whole-genome shotgun sequencing is available on GenBank (The Marmoset Genome Sequencing and Analysis Consortium 2014; URL: <https://www.hgsc.bcm.edu/content/marmoset-genome-project>).

Our group has also sequenced the marmoset genome using animals from the colony maintained by the Central Institute for Experimental Animals (CIEA) in Kawasaki, Japan (Sato et al. 2015). Resequencing and assembly of the genome were performed by deep sequencing with high-throughput sequencing technology using a next-generation sequencer, giving approximately 60× coverage. This

enabled us to generate genome assemblies and gene-coding sequence analysis more efficiently and provided a basis for genome editing.

We have also applied non-invasive imaging methods in marmoset research. The use of marmosets in such studies is limited to small numbers due to cost and ethical issues. Magnetic resonance imaging (MRI) is a non-invasive imaging technique to visualize various organs in detail. We have adapted a number of MRI techniques, including diffusion tensor tractography (DTT; Fujiyoshi et al. 2007; Hikishima et al. 2015) and voxel based morphometric (VBM) analysis (Hikishima et al. 2011, 2015), and a new method for the visualization of myelin (Myelin Map; Fujiyoshi et al. 2016).

Transgenic Techniques and Genome Editing Technology for Marmoset Research

One of the strengths of the mouse model is the availability of powerful genetic tools, such as transgenic and knock-in/knock-out animals, that have given the mouse a central place in life sciences research over the past two decades. However, results from mouse genetics are not always directly relevant to humans. Particularly in the neurosciences, there are considerable interspecies differences in brain anatomy and physiology, behavioral control mechanisms, and life span, and some mouse disease models do not recapitulate human symptoms. For example, neurofibrillary tangles, the neuropathological hallmarks of Alzheimer's disease, cannot be recapitulated in mice showing amyloid plaques (Chin 2011; Games et al. 1995; Hsiao et al. 1996; Sturchler-Pierrat et al. 1997; Tanzi and Bertram 2005; Walsh and Selkoe 2004). It is also known that mice in which *parkin*, the gene associated with familial Parkinson's disease in humans, has been knocked out do not show parkinsonism.

Despite the scientific demand for research in non-human primates, efforts to generate transgenic non-human primate animals have been unsuccessful until recently. In 2008, Yang et al. (2008) reported a transgenic rhesus macaque expressing the human huntingtin (*HTT*) gene with a CAG-expansion encoding the poly glutamine as a model of Huntington's disease. However, despite the genomic insertion of the human *HTT*-transgene in the founder monkeys, germline transmission of the transgene has not been confirmed. Our group independently generated transgenic common marmosets expressing the enhanced GFP (*EGFP*) gene and we reported the first germline transmission in a non-human primate (Sasaki et al. 2009).

While the establishment of transgenic marmosets enables the generation of marmoset models of diseases caused by overexpression of a relatively small mutant gene, such as Parkinson's disease, Alzheimer's disease, and amyotrophic lateral sclerosis (ALS), transgenic techniques limited our ability to genetically modify non-human primates. Transgenic technologies available at the time could only randomly insert only <8 kb of exogenous genes into the genome (Sasaki et al. 2009). Moreover,

transgenes were segregated and suppressed across generations, expression levels could not be controlled, and the techniques were only suited to gain-of-function, not loss-of function, studies. Most human genetic diseases are caused by either point mutations or deletions of endogenous genes, which highlighted the need for new gene modification technologies for use against endogenous genes.

Remarkable recent advances in genome editing technology have now made it possible to overcome these previous limitations (Sato et al. 2016). Genome editing tools, i.e., engineered nucleases, bind to a target genome sequence and introduce specific double-strand breaks. Double-strand breaks initiate cell-endogenous repair mechanisms such as homology-directed repair (HDR), non-homologous end-joining (NHEJ), and microhomology-mediated end-joining (MMEJ). Mutagenesis against endogenous genes can be introduced by taking advantage of such mechanisms. Zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and the clustered regularly interspaced short palindromic repeat (CRISPR)/Cas system are mainly used as engineered nucleases. A number of genetically modified animals have already been generated using such restriction enzymes (Bedell et al. 2012; Geurts et al. 2009; Hauschild et al. 2011; Mashimo et al. 2010; Ochiai et al. 2010; Sung et al. 2013; Suzuki et al. 2013; Wang et al. 2013; Yang et al. 2013). Among these, the CRISPR/Cas system was developed the most recently and is particularly promising (Cong et al. 2013; Mali et al. 2013).

Using these genome editing technologies, we recently generated X-linked SCID model marmosets by knock-out of interleukin-2 receptor subunit gamma gene (Sato et al., 2016). Currently, we are now seeking to generate marmoset models of autism spectrum disorders, including Rett syndrome (Chahrour and Zoghbi 2007; Kishi and Macklis 2005) and tuberous sclerosis complex (Ess 2010; Fig. 1). Although a mouse model (male hemizygous *Mecp2* mutation) is available for Rett syndrome, it does not necessarily mimic the critical symptoms. For example, while male hemizygous mice (*Mecp2*^{-/-}) are used as model mice, Rett patients are exclusively female heterozygous in human. It is likely that males with *MECP2* mutations are embryonic lethal in human, but not in mice. Furthermore, phenotypes appear at adult stages in mouse models, whereas symptoms become evident by 1 year of age in human Rett syndrome patients. New primate models that more closely mimic the

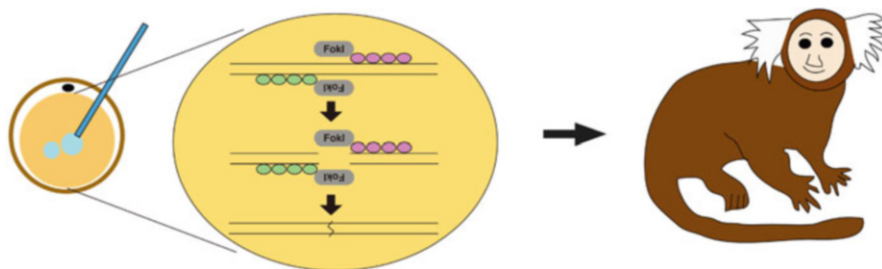


Fig. 1 Generation of a knock-out marmoset by genome editing with ZFN or TALEN

clinical course of human disease may thus contribute to a better understanding of the pathogenesis and future treatments for neurodevelopmental disorders.

Future Perspectives

Genome editing has developed rapidly in recent years, leading to the production of genetically modified animals in many species. This technology has also been applied to non-human primates, and some groups have begun to report genetically modified macaques (Niu et al. 2014; Liu et al. 2014). Macaques offer a number of advantages, but it is difficult to expand colony size within a reasonable research period. We suggest that the common marmoset is thus a highly suitable alternative model primate for many areas of study, and the creation of knock-in/knock-out marmosets would help to introduce the benefits of this model to a larger community of researchers. Since germ-line-competent marmoset embryonic stem cells are not currently available, it is necessary to perform genome editing in one-cell stage embryos (fertilized eggs) to obtain knock-in/knock-out marmosets efficiently. The emergence of more sophisticated genome editing techniques will facilitate and accelerate the development of new gene manipulation technologies in marmoset. Marmoset models of disease generated using genome editing may contribute to the development of new therapeutic strategies for currently incurable neurodegenerative diseases and mental disorders.

Acknowledgments This work was supported by a grant from Brain Mapping by Integrated Neurotechnologies for Disease Studies (Brain/MINDS) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (MEXT) and the Japan Agency for Medical Research and Development (A-MED).

References

- Abbott DH, Barnett DK, Colman RJ, Yamamoto ME, Schultz-Darken NJ (2003) Aspects of common marmoset basic biology and life history important for biomedical research. *Comp Med* 53:339–350
- Austad SN, Fischer KE (2011) The development of small primate models for aging research. *ILAR J/National Research Council, Institute of Laboratory Animal. Resources* 52:78–88
- Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug RG 2nd, Tan W, Penheiter SG, Ma AC, Leung AY, Fahrenkrug SC, Carlson DF, Voytas DF, Clark KJ, Essner JJ, Ekker SC (2012) In vivo genome editing using a high-efficiency TALEN system. *Nature* 491:114–118
- Carrion R Jr, Patterson JL (2012) An animal model that reflects human disease: the common marmoset (*Callithrix jacchus*). *Curr Opin Virol* 2:357–362
- Chahrour M, Zoghbi HY (2007) The story of Rett syndrome: from clinic to neurobiology. *Neuron* 56:422–437
- Chin J (2011) Selecting a mouse model of Alzheimer's disease. *Methods Mol Biol* 670:169–189

- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823
- Dell'Mour V, Range F, Huber L (2009) Social learning and mother's behavior in manipulative tasks in infant marmosets. *Am J Primatol* 71:503–509
- Eliades SJ, Wang X (2008) Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* 453:1102–1106
- Ess KC (2010) Tuberous sclerosis complex: a brave new world? *Curr Opin Neurol* 23:189–193
- Fujiyoshi K, Yamada M, Nakamura M, Yamane J, Katoh H, Kitamura K, Kawai K, Okada S, Momoshima S, Toyama Y, Okano H (2007) In vivo tracing of neural tracts in the intact and injured spinal cord of marmosets by diffusion tensor tractography. *J Neurosci* 27:11991–11998
- Fujiyoshi K, Hikishima K, Nakahara J, Tsuji O, Hata J, Konomi T, Nagai T, Shibata S, Kaneko S, Iwanami A, Momoshima S, Takahashi S, Jinzaki M, Suzuki N, Toyama Y, Nakamura M, Okano H (2016) Application of q-Space diffusion MRI for the visualization of white matter. *J Neurosci* 36:2796–2808
- Games D, Adams D, Alessandrini R, Barbour R, Berthelette P, Blackwell C, Carr T, Clemens J, Donaldson T, Gillespie F, Guido T, Hagopian S, Johnson-Wood K, Khan K, Lee M, Leibowitz P, Lieberburg I, Little S, Masliah E, McConlogue L, Montoya-Zavala M, Mucke L, Paganini L, Penniman E, Power M, Schenk D, Seubert P, Snyder B, Soriano F, Tan H, Vitale J, Wadsworth S, Wolozin B, Zhao J (1995) Alzheimer-type neuropathology in transgenic mice overexpressing V717F beta-amyloid precursor protein. *Nature* 373:523–527
- Geurts AM, Cost GJ, Freyvert Y, Zeitler B, Miller JC, Choi VM, Jenkins SS, Wood A, Cui X, Meng X, Vincent A, Lam S, Michalkiewicz M, Schilling R, Foeckler J, Kalloway S, Weiler H, Menoret S, Anegon I, Davis GD, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Jacob HJ, Buelow R (2009) Knockout rats via embryo microinjection of zinc-finger nucleases. *Science* 325:433
- Gordon DJ, Rogers LJ (2010) Differences in social and vocal behavior between left- and right-handed common marmosets (*Callithrix jacchus*). *J Comp Psychol* 124:402–411
- Hauschild J, Petersen B, Santiago Y, Queisser AL, Carnwath JW, Lucas-Hahn A, Zhang L, Meng X, Gregory PD, Schwinzer R, Cost GJ, Niemann H (2011) Efficient generation of a biallelic knockout in pigs using zinc-finger nucleases. *Proc Natl Acad Sci USA* 108:12013–12017
- Hikishima K, Quallo MM, Komaki Y, Yamada M, Kawai K, Momoshima S, Okano HJ, Sasaki E, Tamaoki N, Lemon RN, Iriki A, Okano H (2011) Population-averaged standard template brain atlas for the common marmoset (*Callithrix jacchus*). *Neuroimage* 54:2741–2749
- Hikishima K, Ando K, Komaki Y, Kawai K, Yano R, Inoue T, Itoh T, Yamada M, Momoshima S, Okano HJ, Okano H (2015) Voxel-based morphometry of the marmoset brain: in vivo detection of volume loss in the substantia nigra of the MPTP-treated Parkinson's disease model. *Neuroscience* 300:585–592
- Hsiao K, Chapman P, Nilsen S, Eckman C, Harigaya Y, Younkin S, Yang F, Cole G (1996) Correlative memory deficits, A β elevation, and amyloid plaques in transgenic mice. *Science* 274:99–102
- Izpisua Belmonte JC, Callaway EM, Caddick SJ, Churchland P, Feng G, Homanics GE, Lee KF, Leopold DA, Miller CT, Mitchell JF, Mitalipov S, Moutri AR, Movshon JA, Okano H, Reynolds JH, Ringach D, Sejnowski TJ, Silva AC, Strick PL, Wu J, Zhang F (2015) Brains, genes, and primates. *Neuron* 86:617–631
- Kishi N, Macklis JD (2005) Dissecting MECP2 function in the central nervous system. *J Child Neurol* 20:753–759
- Kishi N, Sat K, Sasaki E, Okano H (2014) Common marmoset as a new model animal for neuroscience research and genome editing technology. *Dev Growth Differ* 56:53–62
- Kitamura K, Fujiyoshi K, Yamane J, Toyota F, Hikishima K, Nomura T, Funakoshi H, Nakamura T, Aoki M, Toyama Y, Okano H, Nakamura M (2011) Human hepatocyte growth factor promotes functional recovery in primates after spinal cord injury. *PLoS One* 6:e27706

- Liu H, Chen Y, Niu Y, Zhang K, Kang Y, Ge W, Liu X, Zhao E, Wang C, Lin S, Jing B, Si C, Lin Q, Chen X, Lin H, Pu X, Wang Y, Qin B, Wang F, Wang H, Si W, Zhou J, Tan T, Li T, Ji S, Xue Z, Luo Y, Cheng L, Zhou Q, Li S, Sun YE, Ji W (2014) TALEN-mediated gene mutagenesis in rhesus and cynomolgus monkeys. *Cell Stem Cell* 14:323–328
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013) RNA-guided human genome engineering via Cas9. *Science* 339:823–826
- Mansfield K (2003) Marmoset models commonly used in biomedical research. *Comp Med* 53:383–392
- Mashimo T, Takizawa A, Voigt B, Yoshimi K, Hiai H, Kuramoto T, Serikawa T (2010) Generation of knockout rats with X-linked severe combined immunodeficiency (X-SCID) using zinc-finger nucleases. *PLoS One* 5:e8870
- Niu Y, Shen B, Cui Y, Chen Y, Wang J, Wang L, Kang Y, Zhao X, Si W, Li W, Xiang AP, Zhou J, Guo X, Bi Y, Si C, Hu B, Dong G, Wang H, Zhou Z, Li T, Tan T, Pu X, Wang F, Ji S, Zhou Q, Huang X, Ji W, Sha J (2014) Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell* 156:836–843
- Ochiai H, Fujita K, Suzuki K, Nishikawa M, Shibata T, Sakamoto N, Yamamoto T (2010) Targeted mutagenesis in the sea urchin embryo using zinc-finger nucleases. *Genes Cells* 15:875–885
- Okano H, Sasaki E, Yamamori T, Iriki A, Shimogori T, Yamaguchi Y, Kasai K, Miyawaki A (2016) Brain/MINDS: a Japanese National Brain Project for Marmoset Neuroscience. *Neuron* 92:582–590
- Sasaki E, Suemizu H, Shimada A, Hanazawa K, Oiwa R, Kamioka M, Tomioka I, Sotomaru Y, Hirakawa R, Eto T, Shiozawa S, Maeda T, Ito M, Ito R, Kito C, Yagihashi C, Kawai K, Miyoshi H, Tanioka Y, Tamaoki N, Habu S, Okano H, Nomura T (2009) Generation of transgenic non-human primates with germline transmission. *Nature* 459:523–527
- Sato K, Kuroki Y, Kumita W, Fujiyama A, Toyoda A, Kawai J, Iriki A, Sasaki E, Okano H, Sakakibara Y (2015) Resequencing of the common marmoset genome assemblies and gene-coding sequence analysis. *Sci Rep* 5:16894
- Sato K, Oiwa R, Kumita W, Henry R, Sakuma T, Ito R, Nozu R, Inoue T, Katano I, Sato K, Okahara N, Okahara J, Shimizu Y, Yamamoto M, Hanazawa K, Kawakami T, Kametani Y, Ryuji Suzuki R, Takahashi T, Weinstein EJ, Yamamoto T, Sakakibara Y, Habu S, Hata J, Okano H, Sasaki E (2016) Generation of a nonhuman primate model of severe combined immunodeficiency using highly efficient genome editing. *Cell Stem Cell* 19:127–138
- Smith D, Trennery P, Farningham D, Klapwijk J (2001) The selection of marmoset monkeys (*Callithrix jacchus*) in pharmaceutical toxicology. *Lab Anim* 35:117–130
- Sturchler-Pierrat C, Abramowski D, Duke M, Wiederhold KH, Mistl C, Rothacher S, Ledermann B, Burki K, Frey P, Paganetti PA, Waridel C, Calhoun ME, Jucker M, Probst A, Staufenbiel M, Sommer B (1997) Two amyloid precursor protein transgenic mouse models with Alzheimer disease-like pathology. *Proc Natl Acad Sci USA* 94:13287–13292
- Summers PM, Wennink CJ, Hodges JK (1985) Cloprostenol-induced luteolysis in the marmoset monkey (*Callithrix jacchus*). *J Reprod Fertil* 73:133–138
- Sung YH, Baek JJ, Kim DH, Jeon J, Lee J, Lee K, Jeong D, Kim JS, Lee HW (2013) Knockout mice created by TALEN-mediated gene targeting. *Nat Biotechnol* 31:23–24
- Suzuki KT, Isoyama Y, Kashiwagi K, Sakuma T, Ochiai H, Sakamoto N, Furuno N, Kashiwagi A, Yamamoto T (2013) High efficiency TALENs enable F0 functional analysis by targeted gene disruption in *Xenopus laevis* embryos. *Biol Open* 2:448–452
- t'Hart BA, Vervordeldonk M, Heeney JL, Tak PP (2003) Gene therapy in nonhuman primate models of human autoimmune disease. *Gene Ther* 10:890–901
- t'Hart BA, Abbott DH, Nakamura K, Fuchs E (2012) The marmoset monkey: a multi-purpose preclinical and translational model of human biology and disease. *Drug Discov Today* 17:1160–1165
- Tanzi RE, Bertram L (2005) Twenty years of the Alzheimer's disease amyloid hypothesis: a genetic perspective. *Cell* 120:545–555

- Tardif SD, Smucny DA, Abbott DH, Mansfield K, Schultz-Darken N, Yamamoto ME (2003) Reproduction in captive common marmosets (*Callithrix jacchus*). *Comp Med* 53:364–368
- The Marmoset Genome Sequencing and Analysis Consortium (2014) The common marmoset genome provides insight into primate biology and evolution. *Nat Genet* 46:850–857
- Tokuno H, Moriya-Ito K, Tanaka I (2012) Experimental techniques for neuroscience research using common marmosets. *Experimental animals/Japanese Association for. Lab Anim Sci* 61:389–397
- Walsh DM, Selkoe DJ (2004) Deciphering the molecular basis of memory failure in Alzheimer's disease. *Neuron* 44:181–193
- Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153:910–918
- Yang SH, Cheng PH, Banta H, Piotrowska-Nitsche K, Yang JJ, Cheng EC, Snyder B, Larkin K, Liu J, Orkin J, Fang ZH, Smith Y, Bachevalier J, Zola SM, Li SH, Li XJ, Chan AW (2008) Towards a transgenic model of Huntington's disease in a non-human primate. *Nature* 453:921–924
- Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R (2013) One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 154:1370–1379

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Multiscale Genome Engineering: Genome-Wide Screens and Targeted Approaches

Neville E. Sanjana

Abstract New advances in genome engineering technologies, such as efficient programmable CRISPR nucleases, have enabled new advances in forward and reverse genetic studies. Here, I discuss recent work from our group combining top-down approaches like genome-wide loss-of-function screens and bottom-up approaches like disease variant modeling in human stem cells and stem cell-derived cortical neurons.

Introduction

Patient-sequencing studies have yielded large lists of disease-associated gene variants but it has been difficult to establish a causal role based solely on genetics. New methods are needed for rapidly understanding the effects of these variants and ascertaining whether the variants directly influence disease-related phenotypes. At the IPSEN meeting “Genome Editing in Neurosciences,” I presented two approaches (top-down and bottom-up) for harnessing new genome engineering techniques to decipher the roles of genetic variants in human health and disease.

Top-down approaches utilize large-scale pooled libraries of genome-engineering reagents to start with a large, minimally biased hypothesis space and identify relevant variants via a single phenotypic selection (Fig. 1, right). In contrast, bottom-up approaches start with a handful of genetic variants nominated by strong genetic data, e.g., genome-wide association studies, case-control, family linkage studies, etc., and they examine a wide range of phenotypes (Fig. 1, left).

N.E. Sanjana (✉)

New York Genome Center, New York, NY, USA

Department of Biology, New York University, New York, NY, USA

e-mail: neville@sanjanalab.org

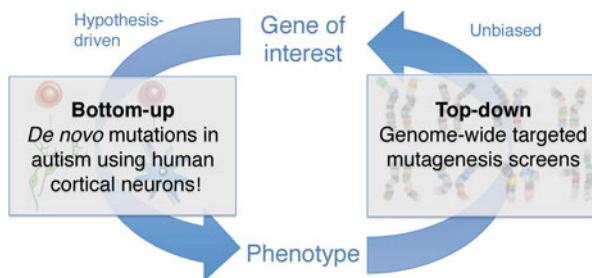


Fig. 1 Schematic diagram of the dynamic interplay between top-down and bottom-up genetic approaches. Top-down approaches (*right*) identify new gene candidates that can shape/reduce the space of hypotheses for more detailed bottom-up (*left*) cellular models and phenotyping. Top-down approaches are unbiased or minimally biased to cast a wide net of possible genetic hypotheses and use phenotypic selection to identify putative disease-associated gene variants. Bottom-up approaches focus on a smaller set of variants but usually provide a more detailed phenotypic analysis of different molecular/cellular/circuit aspects of each genetic variant. Candidate variants for the bottom-up approach can be derived from either genetic evidence (e.g., patient sequencing studies) or from top-down approaches like genome-wide CRISPR screens

Top-Down Approaches Using Genome-Wide CRISPR Screens

The microbial CRISPR-Cas9 nuclease from *S. pyogenes* can be guided to specific DNA sequences using a 20 bp guide sequence. Given the short length of the guide sequence, we have been able to use oligonucleotide array synthesis techniques to create libraries of thousands of guide sequences in a pooled format. By designing pooled libraries to target all genes in a specific genome, we have created a new tool for functional genomic screens (Sanjana 2016). Genome-scale CRISPR knock-out (GeCKO) screens use the consistent phenotypic enrichment of *multiple* CRISPR reagents targeting the same gene to lend evidence to the gene's role in a particular disease. Using a GeCKO library targeting ~18,000 genes with 64,751 guide sequences, we have found loss-of-function mutations that confer resistance to the BRAF inhibitor vemurafenib in human melanoma cells (Shalem et al. 2014). Using a second-generation GeCKO library in a mouse model, we performed an *in vivo* screen to identify driver mutations that trigger metastasis to the lung (Sanjana et al. 2014; Chen et al. 2015). Recently, we have expanded the scope of CRISPR pooled screens to also include noncoding regions of the genome (Sanjana et al. 2016; Wright and Sanjana 2016).

Bottom-Up Approaches Using Exome Sequencing in Autism

Whole exome and whole genome sequencing have ushered in a revolution in identifying rare, disease-associated variants. Several whole exome sequencing studies have pinpointed rare *de novo* variants associated with autism spectrum disorder by examining exomes from autistic individuals and comparing them to parental

exomes (Iossifov et al. 2012; Neale et al. 2012; O’Roak et al. 2012; Sanders et al. 2012). The next logical step is to create relevant cellular models to better understand the mechanisms through which these variants work and to serve as a platform for drug screens and therapeutic testing. There are two major roadblocks for building these kinds of cellular models. The first one concerns gene editing and, over the past 3 years, has become largely historical: Until recently, genome engineering in human stem cells and neurons has been challenging but transfection of CRISPR plasmids or ribonucleoproteins provides an easy, efficient technique for engineering human cells (Peters et al. 2008; Swiech et al. 2015). The second major hurdle has been neural differentiation. Common protocols to differentiate neurons from human stem cells, such as dual SMAD inhibition or embryoid body differentiation, require months to create mature neurons (Zhang et al. 2001; Chambers et al. 2009). Recently, we and others have demonstrated that viral overexpression of Neurogenin 1 or Neurogenin 2 can rapidly drive stem cells into a homogeneous culture of mature cortical neurons (Zhang et al. 2013; Busskamp et al. 2014). These neurons display robust electrophysiological activity within just 2–3 weeks after the start of differentiation, making them ideally suited for synaptic assays, calcium imaging and neurophysiology. We are now moving forward with phenotypic analyses of de novo mutations in autism using the combined CRISPR-Neurogenin platform for rapid mutagenesis and human neuron profiling.

Taken together, these new technologies in genome engineering—enabled in large part by CRISPR nucleases and related transformative methods—have improved our ability to perform forward and reverse genetic assays in relevant model systems. A major challenge for neuroscience is finding clear phenotypes that accurately reflect complex diseases, such as schizophrenia or autism. Despite these challenges, a combination of top-down and bottom-up approaches will pave the way for a clearer understanding of the human brain in healthy and disease states.

References

- Busskamp V, Lewis NE, Guye P, Ng AHM, Shipman SL, Byrne SM, Sanjana NE, Murn J, Li Y, Li S, Stadler M, Weiss R, Church GM (2014) Rapid neurogenesis through transcriptional activation in human stem cells. *Mol Syst Biol* 10:760
- Chambers SM, Fasano CA, Papapetrou EP, Tomishima M, Sadelain M, Studer L (2009) Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol* 27:275–280
- Chen S, Sanjana NE, Zheng K, Shalem O, Lee K, Shi X, Scott DA, Song J, Pan JQ, Weissleder R, Lee H, Zhang F, Sharp PA (2015) Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell* 160:1246–1260
- Iossifov I, Ronemus M, Levy D, Wang Z, Hakker I, Rosenbaum J, Yamrom B, Lee YH, Narzisi G, Leotta A, Kendall J, Grabowska E, Ma B, Marks S, Rodgers L, Stepansky A, Troge J, Andrews P, Bekritsky M, Pradhan K, Ghiban E, Kramer M, Parla J, Demeter R, Fulton LL, Fulton RS, Magrini VJ, Ye K, Darnell JC, Darnell RB, Mardis ER, Wilson RK, Schatz MC, McCombie WR, Wigler M (2012) De novo gene disruptions in children on the autistic spectrum. *Neuron* 74:285–299
- Neale BM, Liu L, Ma’ayan A, Samocha KE, Sabo A, Lin CF, Stevens C, Wang LS, Makarov V, Polak P, Yoon S, Maguire J, Crawford EL, Campbell NG, Geller ET, Valladares O, Schafer C,

- Liu H, Zhao T, Cai G, Lihm J, Dannenfelser R, Jabado O, Peralta Z, Nagaswamy U, Muzny D, Reid JG, Newsham I, Wu Y, Lewis L, Han Y, Voight BF, Lim E, Rossin E, Kirby A, Flannick J, Fromer M, Shakir K, Fennell T, Garimella K, Banks E, Poplin R, Gabriel S, DePristo M, Wimbish JR, Boone BE, Levy SE, Betancur C, Sunyaev S, Boerwinkle E, Buxbaum JD, Cook EH Jr, Devlin B, Gibbs RA, Roeder K, Schellenberg GD, Sutcliffe JS, Daly MJ (2012) Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 485:242–245
- O’Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, Levy R, Ko A, Lee C, Smith JD, Turner EH, Stanaway IB, Vernot B, Malig M, Baker C, Reilly B, Akey JM, Borenstein E, Rieder MJ, Nickerson DA, Bernier R, Shendure J, Eichler EE (2012) Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 485:246–250
- Peters DT, Cowan CA, Musunuru K (2008) Genome editing in human pluripotent stem cells. Harvard Stem Cell Institute, Cambridge. StemBook [Internet] PMID:23785737
- Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ, Ercan-Sencicek AG, DiLullo NM, Parikshak NN, Stein JL, Walker MF, Ober GT, Teran NA, Song Y, El-Fishawy P, Murtha RC, Choi M, Overton JD, Bjornson RD, Carriero NJ, Meyer KA, Bilguvar K, Mane SM, Sestan N, Lifton RP, Günel M, Roeder K, Geschwind DH, Devlin B, State MW (2012) De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 485:237–241
- Sanjana NE (2016) Genome-scale CRISPR pooled screens. *Anal Biochem.* pii:S0003-2697(16)30089-6
- Sanjana NE, Shalem O, Zhang F (2014) Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* 11:783–784
- Sanjana NE, Wright J, Zheng K, Shalem O, Fontanillas P, Joung J, Cheng C, Regev A, Zhang F (2016) High-resolution interrogation of functional elements in the noncoding genome. *Science* 353(6307):1545–1549
- Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, Heckl D, Ebert BL, Root DE, Doench JG, Zhang F (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343:84–87
- Wiech L, Heidenreich M, Banerjee A, Habib N, Li Y, Trombetta J, Sur M, Zhang F (2015) In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat Biotechnol* 33:102–106
- Wright JB, Sanjana NE (2016) CRISPR screens to discover functional noncoding elements. *Trends Genet* 32:526–529
- Zhang SC, Wernig M, Duncan ID, Brüstle O, Thomson JA (2001) In vitro differentiation of transplantable neural precursors from human embryonic stem cells. *Nat Biotechnol* 19:1129–1133
- Zhang Y, Pak C, Han Y, Ahlenius H, Zhang Z, Chanda S, Marro S, Patzke C, Acuna C, Covy J, Xu W, Yang N, Danko T, Chen L, Wernig M, Südhof TC (2013) Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron* 78:785–798

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Using Genome Engineering to Understand Huntington's Disease

Barbara Bailus, Ningzhe Zhang, and Lisa M. Ellerby

Abstract Huntington's disease (HD) is a fatal, dominantly inherited neurodegenerative disorder caused by a CAG trinucleotide expansion in the *Huntingtin* (*HTT*) gene, leading to an expanded polyglutamine (polyQ) region in the encoded protein HTT. We have used homologous recombination (HR) to genetically correct HD patient-derived induced pluripotent stem cells (iPSCs) and found that this reversed HD disease phenotypes. We have utilized exploited genome editing tools including TALENs (Transcription like activator effectors) and CRISPR (Clustered Regulatory Interspaced Short Palindromic Repeats)/Cas9 technology to carry out genetic correction or expansion, and we were able to detect HR without selection in human cells. The overall goal is to use this technology to model HD-relevant cell types and better understand disease progression by leveraging system biology approaches. To understand the disease progression, isogenic iPSC lines were created. We found that the disease phenotypes only manifested in the differentiated neural stem cell (NSC) stage, not in iPSCs. Transcriptomic analysis of HD iPSCs and HD NSCs compared to isogenic controls was utilized to understand the molecular basis for the CAG repeat expansion-dependent disease phenotypes in NSCs. Differential gene expression and pathway analysis identified transforming growth factor β (TGF- β) signaling, netrin-1 signaling and medium spiny neuron (MSNs) maturation and maintenance as the top dysregulated pathways in HD NSCs. The ability to create additional isogenic cell lines through CRISPR-mediated HR will further enhance our understanding of HD progression. These lines can be manipulated with CRISPR to understand the effects of common SNPs (single nucleotide polymorphism) that modulate disease onset in HD, allowing the identification of new pathways and helping to elucidate potential therapeutic targets for HD. Beyond drug discovery, the CRISPR system could eventually be optimized to use in vivo, correcting a patient's disease-causing mutation, in the asymptomatic stages of HD.

B. Bailus • N. Zhang • L.M. Ellerby (✉)

Buck Institute for Research on Aging, 8001 Redwood Blvd, Novato, CA 94945, USA

e-mail: llerby@buckinstitute.org

Huntington' Disease

Huntington's disease (HD) is a devastating, dominantly inherited movement and psychiatric disorder that is caused by expansion of a CAG trinucleotide repeat in the first exon of the Huntingtin gene (*HTT*), resulting in translation of an expanded polyQ repeat in the HTT protein. The production of the abnormal expanded polyQ-containing HTT protein leads to a dramatic loss of striatal and cortical neurons and pro-survival growth factors such as BDNF (brain derived neurotrophic factor) in HD patients. The polyQ expansion in the HTT protein leads to disrupted cellular homeostasis and activation of cellular death pathways (Fig. 1). Since the disease is inherited in an autosomal dominant fashion, each child of an affected parent has a 50% chance of being affected. HD generally manifests in mid-life, with a mean age of onset of 35–45 years of age. The disease begins with cognitive disturbances and progresses to severe and debilitating motor symptoms (chorea) usually accompanied by psychiatric disturbances, with death following in about 15–20 years (Landles and Bates 2004). The current therapeutic approaches in HD focus on normalizing molecular pathways disturbed in HD or on lowering the levels of the mutant HTT protein (Canals et al. 2004; Conforti et al. 2008; Zuccato et al. 2008). To date none of these approaches are approved for use outside of clinical trials and they will not cure the disease.

In this chapter, we discuss the use of gene editing tools to model neurological diseases such as HD as well as the potential to use this technology to treat genetic neurological diseases.

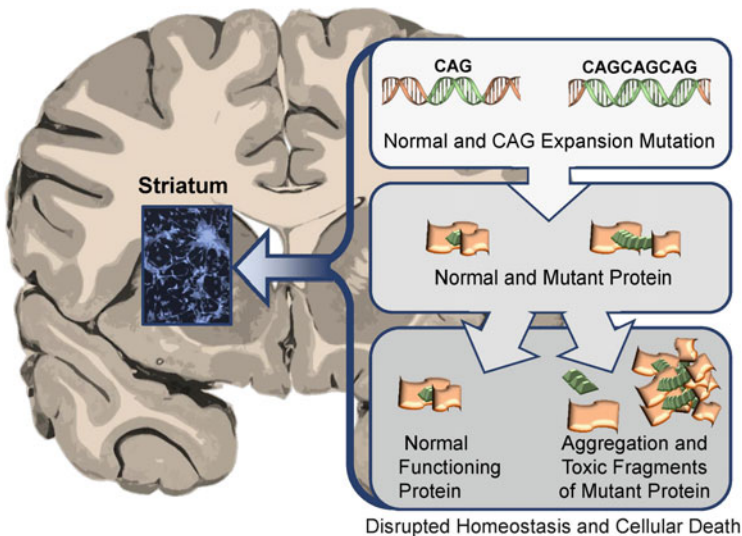


Fig. 1 Illustration on the neuronal changes occurring in the striatum of a Huntington's disease patient. The exon 1 CAG expansion in the *HTT* allele results in a mutant protein being formed; the mutant protein aggregates and is also cleaved into toxic fragments. The aggregates and the toxic fragments result in a disrupted cellular homeostasis and eventual neuronal cellular death in the striatum

Gene Editing Enzymes

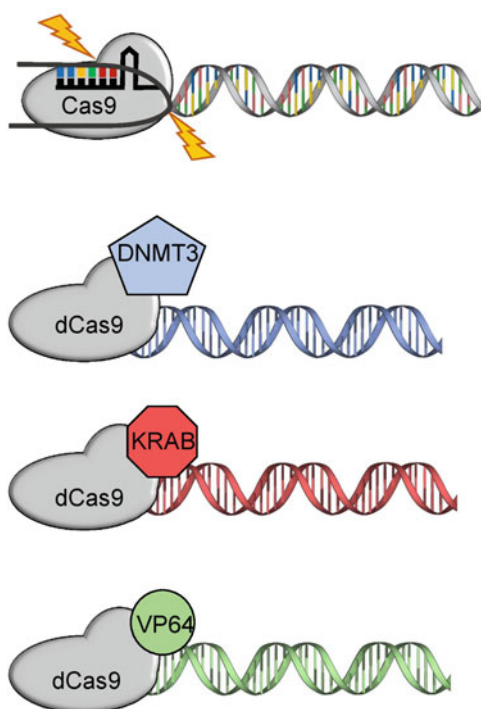
Targeted gene editing has evolved dramatically in the last 25 years. While originally a technique that a handful of laboratories had mastered, it is now a common tool used in hundreds of laboratories around the world. One family of gene editing proteins is the customized zinc finger proteins (Segal and Barbas 2000; Wolfe et al. 2000; Pabo et al. 2001; Nagaoka and Sugiura 2000). These proteins were adapted for targeted use in the late 1990s (Liu et al. 1997; Segal et al. 1999; Dreier et al. 2001). Each zinc finger protein could be designed to recognize three different base pairs on DNA through various interactions between the proteins alpha helix amino acids and the DNA base pairs (Segal and Barbas 2001). To recognize a specific sequence of DNA, the zinc fingers could be attached to each other, with six zinc fingers recognizing a unique 18-base pair sequence in an organism's genome. The zinc finger proteins could have effector or nuclease domains attached, allowing for gene regulation or gene replacement. The effector domains included VP64 for gene activation, KRAB for gene silencing and DNMT1 for methylation (Beerli et al. 1998; Rivenbark et al. 2012). The nuclease domain could cut targeted genomic sites and allow for mutagenesis or homologous recombination at enhanced efficiency. Zinc finger proteins have been successfully used in human cells, animal organs and have reached Phase II human clinical trials (Geurts et al. 2009; Urnov et al. 2005; SangamoBiosciences 2001; Eisenstein 2012). Although promising, zinc fingers presented several challenges for researchers. Their targeting ability was limited, they required specialized design techniques and they exhibited a frequent incidence of off-target events (Cornu and Cathomen 2010; Gupta et al. 2010; Gabriel et al. 2011). Some advances have been made to reduce the off-target potential and increase detection of these events (Zykovich et al. 2009; Cornu et al. 2008). The therapeutic potential of zinc fingers for a variety of diseases, including HD, continues to be explored by the biotechnology company Sangamo (Cornu et al. 2008; Wolffe 2016).

In 2009, a new gene editing protein was described, transcription activator-like effectors (TALEs; Boch et al. 2009; Moscou and Bogdanove 2009). These proteins were originally characterized in *Xanthomonas* bacteria and represented a major advance for DNA regulating proteins. TALEs, unlike zinc fingers, made contact with individual DNA base pairs, which greatly expanded the sequences that could be targeted in the genome (Moscou and Bogdanove 2009). They were also much easier to design and assemble. Much like zinc fingers, TALEs could have effector or nuclease domains attached to the DNA binding domain, allowing for the DNA to be cut or for genes to be regulated (Christian et al. 2010; Maeder et al. 2013a, b; Cong et al. 2012). Promising experiments in a variety of organisms have validated the efficacy of TALEs, although no human clinical trials have begun. A recent publication has shown the ability of TALEs to specifically silence the mutant HTT allele in cell culture models or to engineer an allelic series into the HTT locus (Fink et al. 2016; Wang et al. 2013). The TALEs still exhibit off-target effects and may have potential immune issues (Guilinger et al. 2014).

Gene editing became a widely accessible technology in 2012 with the characterization of the CRISPR system and its implications for targeted gene editing and

regulation. The CRISPR system is composed of a Cas9 nuclease and a gRNA complex. To cut the DNA, Cas9 attaches to the guide RNA (gRNA), which targets a specific site in the organism's DNA (Jinek et al. 2012; Wiedenheft et al. 2012). This system is found in archaea and bacteria and is used as a natural defense mechanism against bacteriophages. The system has been characterized and adapted for mammalian-targeted genome editing. The gRNA has one targeting requirement, a PAM motif (typically a NGG) at the 3' end of the DNA targeting site; this sequence is common in DNA and thus almost any gene can be targeted with the CRISPR system (Gilbert et al. 2013; Qi et al. 2013). As with previous gene editing proteins, the Cas9 can be modified to either silence or activate gene transcription (Fig. 2; Sander and Joung 2014; Larson et al. 2013). Due to some initial off-target cleavage events, the Cas9 nuclease was modified to become a Cas9 nickase (Cas9n; Ran et al. 2013). This modification drastically increased targeting specificity, as the binding of two Cas9n proteins targeting two different DNA sites was required to make a double strand break in the DNA and encouraged homologous recombination (HR) with a potential donor DNA strand. Overall the off-target effects of Cas9n could be reduced to background levels (O'Geen et al. 2015; Wu et al. 2014). The modified Cas9n was found to have similar cleavage efficiency when two gRNAs were used, one targeted on each strand of the DNA, resulting in a double strand break. The technique has been widely adopted to create disease-modeling cell lines, rodent and non-human primate models and in non-viable human embryos

Fig. 2 Illustrations of different CRISPR/Cas9 uses with variable effector domains. The wild-type Cas9 nuclease can be used to initiate double strand breaks, encouraging homologous recombination. The inactive Cas9 (dCas9) attached to a DNMT3 can be used for site-specific methylation, resulting in semi-permanent gene repression. A dCas9 can have a KRAB domain attached for temporary gene repression or a VP64 domain for activation



(Liang et al. 2015; Chen et al. 2015). What has made the CRISPR system so accessible is that, unlike the zinc fingers and TALEs, the same core protein, Cas9, is used to target any sequence, whereas the targeting portion of the CRISPR system, the gRNA, is what varies. The gRNA can be designed and synthesized either in a standard lab or by an outside company. This separation of the targeting portion (gRNA) of the CRISPR system from the modifying portion (Cas9 or other effectors) allows for targeting multiple genes in one experiment (Wang et al. 2013). The ability to target multiple genes in a single experiment drastically reduces the time needed to model complex genetic disorders in which more than one gene is involved. All of these unique characteristics have resulted in a rapid popularization of the CRISPR system in research labs, with thousands of papers having been published in the last five years.

Uses for Gene Editing to Understand Human Diseases

Due to their ability to precisely target a gene or regulatory element, genome editing tools have been widely utilized to model human diseases both in cells and in animals. Neurodegenerative diseases such as Parkinson's disease and HD have been modeled by introducing disease-causing mutations into human induced pluripotent stem cells (iPSCs) facilitated by genome editing tools (O'Brien et al. 2015; Soldner et al. 2011). CRISPR/Cas9 or TALENs can also be injected into zygotes or embryos to get genetically modified animals. Researchers have injected TALEN-expressing mRNAs into zebrafish embryos to target the gene *glucocerebrosidase 1*, which is mutated in the lysosomal storage disorder Gaucher's disease. The introduction of these TALENs caused a deletion mutation of the protein Glucocerebrosidase 1, and characteristics of the Gaucher's disease were present in this zebrafish model (Keatinge et al. 2015). Duchenne muscular dystrophy (DMD) is a neuromuscular disorder caused by a loss-of-function mutation of the gene *dmd*. A DMD rat model was generated by delivering CRISPR system into rat zygotes to target the *dmd* gene (Nakamura et al. 2014). These disease models are valuable tools for the exploration of disease mechanisms and for the pursuit of therapeutics.

When combined with human pluripotent stem cells, genome editing tools can provide some unique advantages in disease modeling and mechanism study. Human pluripotent stem cells, including iPSCs and embryonic stem cells, can be directed to any cell types of the human body with the correct differentiation conditions. Thus relevant cell types for the disease and changes in this development can be studied in these models. When genome editing tools are used to add or remove a mutation at the pluripotent stem cell stage, isogenic cell lines with an almost identical genetic background are obtained. As cells are differentiated into more restricted stem cells and terminally differentiated cells, the isogenic background will persist. Phenotypic changes of these cells are most likely a result of the mutation, as they have an identical genetic background. However, one may still have to consider epigenetic changes and mitochondrial mutations that may remain harbored in the patient's iPSCs' background

(Chinnery et al. 2012; Calvanese et al. 2009). These isogenic cell lines can be subjected to systematic approaches including DNA microarray, RNA-seq and mass spectrometry for transcriptomic and proteomic information. Bioinformatic analysis can identify interesting gene/protein targets or signaling pathways that have distinct disease-associated patterns. The cleaner background of isogenic cell models should result in more relevant and reliable hits. After proper validation, these potentially important disease targets may lead to discovery of new mechanisms or drugs.

Recent advances in stem cell research suggest that iPSCs may provide novel models of disease and new treatments for diseases. An isogenic iPSC line was established in the Ellerby lab through traditional means of HR on a human HD patient iPSC line. This isogenic line introduced a corrected donor strand for the CAG expansion and corrected the disease allele to a wild type allele (An et al. 2012). The isogenic corrected line had the exact same genetic background as the patient, reducing the genetic variables that are present when one compares disease phenotypes across multiple different patients to matched wild type individuals. One of the first questions we addressed was whether we could take HD patient-derived iPSCs and, through genetic correction of the disease allele, reverse disease phenotypes. Interestingly, we did not detect phenotypes in the undifferentiated HD iPSCs but only observed disease phenotypes in the differentiated neural stem cell (NSC) state, and these phenotypes were reversible upon genetic correction of the patient mutation.

To understand the molecular basis for the CAG repeat expansion-dependent disease phenotypes in iPSCs and NSCs, RNA-Seq was performed comparing the isogenic corrected lines to HD iPSCs and HD NSCs. We observed that there were few phenotypic differences between HD and wild type iPSCs, but there were substantial differences—over 2000 dysregulated genes—in the NSCs. Some of the key pathways that were dysregulated included TGF- β , netrin-1 signaling and development of the striatum (Fig. 3; Ring et al. 2015). Particularly important, our isogenic HD-iPSCs with corrected alleles identified the maturation or maintenance of medium spiny neurons (MSNs) as being dysregulated (Ring et al. 2015). We showed that the pathways or factors that were involved in this process were therapeutic targets for HD (Ring et al. 2015). A subsequent publication from another group emphasized the de-differentiation of MSNs or loss of MSN identity in HD is a major source of dysfunction (Langfelder et al. 2016). These pathways offer new options for therapeutic treatments and drug targets. Using genetic engineering, we generated an isogenic allelic HD iPSC series for HD modeling (CAG repeat of 21, 45, 72, 100). By creating additional isogenic lines, the contribution of the CAG expansion to the disease phenotypes can be elucidated from background variation; this information can help guide researchers towards additional treatment targets (O'Brien et al. 2015).

Besides directly modifying the disease gene, genome editing tools can also be used to engineer cells to facilitate disease research by making reporter cell lines. In an effort to investigate the roles of a gene encoding a sodium channel subunit in epilepsy, a tdTomato fluorescence protein gene cassette was inserted into iPSCs under a GABAergic neuron-specific promoter with CRISPR/Cas9. When differentiated into GABAergic neurons, these cells were red fluorescently labeled and could be readily followed for electrophysiological studies (Liu et al. 2016). Another example is in the

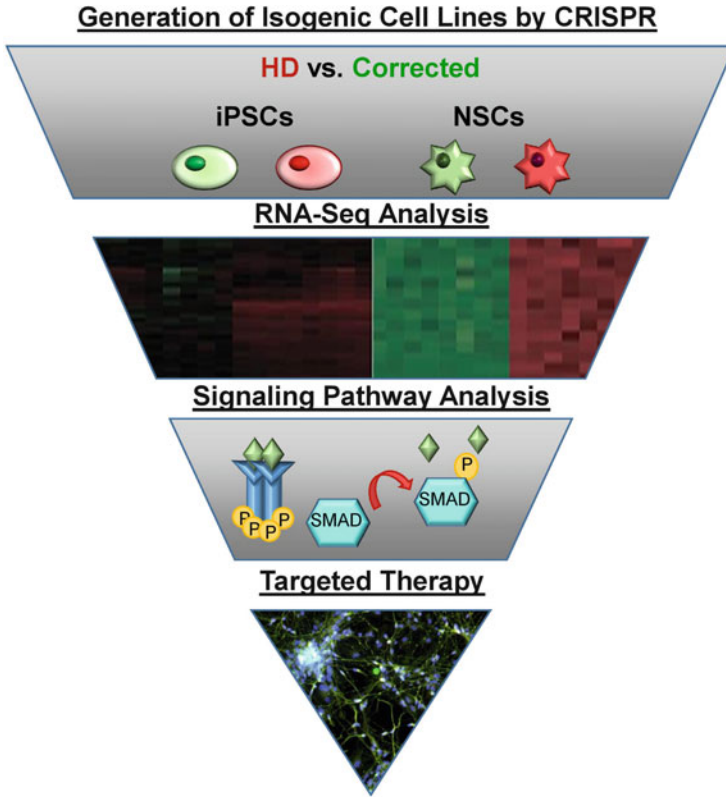


Fig. 3 A flow chart comparing the donor Huntington disease (HD) and genetically corrected isogenic iPSCs and NSCs. Transcriptomic analysis was performed on the cell lines in which significant differences were found in multiple signaling pathways. These newly identified pathways could result in additional drug targets

peripheral neuropathy Charcot–Marie–Tooth disease, type 1A. With TALENs a bio-luminescent reporter was integrated under the regulation of the disease causing gene *pmp22*, which allowed high throughput screening for reagents that can decrease expression of this gene (Inglese et al. 2014). In an effort to better track the recombination repair efficiency in HD cells, the Ellerby lab has designed a myc-tagged donor strand that, when incorporated into the cell, is detectable by both Western blot and PCR amplification; these methods are so sensitive that recombination efficiencies can be detected at levels as low as 5% (Fig. 4). For polyglutamine disease, it is also possible to detect the prevalence of the polyglutamine expansion through the use of specific antibodies, which detect the expanded polyglutamine region (Fig. 4; An et al. 2014). The ability to qualitatively assess how many cells have been corrected will increase the field's understanding of what may be a therapeutic level of correction for the disease. Having specific tags to monitor genetic correction rates and resulting phenotypic improvements will advance the field's understanding toward designing genetic correction and optimize treatment conditions.

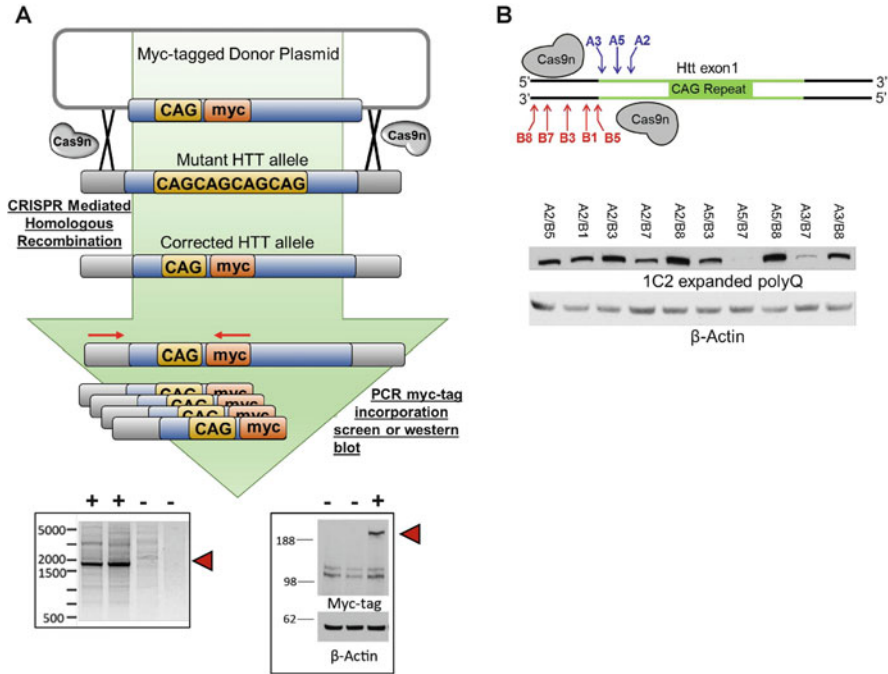


Fig. 4 (a) Use of myc tag in corrected donor plasmid allows for both insertion screening at the DNA level by PCR (*left*) and at the protein level by Western blot (*right*); red triangles indicate expected band size. (b) Use of 1C2 antibody screening with an expanded donor plasmid, a rapid method to optimize different gRNA combinations for homologous recombination efficiency

Gene Editing In Vivo to Treat Genetic Diseases

With its extreme ease of use and targeting, the CRISPR system is being studied extensively with a goal of in vivo correction of genetic mutations. Recent advances have shown that it takes about 15 h for Cas9-mediated double strand breaks to be repaired; this is potentially due to Cas9 remaining bound to the DNA for an extended period of time and because it asymmetrically releases the target strand (Richardson et al. 2016). This asymmetric release of the strand has given researchers the ability to rationally design the donor strands in an effort to increase gene correction percentages; it also provides additional insight as to how to target and design the donor strands. The guide RNAs have also continued to evolve since the first characterization of the CRISPR system. Initially there were two components to the guide RNA, a crRNA and a gRNA, and these were able to be fused creating a simpler method in which the gRNA could be delivered already assembled. Multiple assembled gRNAs could be placed on the same plasmid, allowing for multiple gene targeting with minimal plasmids (Wang et al. 2013; Hsu et al. 2013). A couple of new CRISPR variants have been characterized that offer even lower off-target binding levels and are smaller (Ran et al. 2015). Both of these new

characteristics may be useful in eventual patient treatment, as a smaller CRISPR protein could be more easily packaged for delivery and lower off-target binding increases the specificity of the CRISPR protein, restricting the effects to the target site.

The most exciting application of genome editing tools in human genetic diseases is genetic correction and normalization of those disease mutations. These have been done in cells. For example, in Myotonic dystrophy type 1, a genetic modification has been introduced by TALEN in a NSC model and this modification has shown some restoration of disease phenotypes (Xia et al. 2015). More encouragingly, genetic correction has been achieved in adult animals. Recently several groups published genetic correction in a mouse DMD model. Adeno-associated virus-delivered CRISPR/Cas9 was used to remove a mutation from the gene *dmd*. Partial phenotypic recovery has been observed in these studies (Xu et al. 2016; Nelson et al. 2016; Tabebordbar et al. 2016). The use of CRISPR in vivo to ablate the rhodopsin gene carrying the dominant S334ter mutation in rats with severe autosomal dominant retinitis pigmentosa also highlights the use of genetic correction in disease (Bakondi et al. 2016). These proof-of-principle experiments may be the first steps towards overcoming many currently incurable genetic diseases. CRISPR technology is already being used in human cells and disease models with the eventual goal of patient treatment. A recent study conducted in China has even used CRISPR technology on non-viable human embryos (Liang et al. 2015). As this technology has advanced so rapidly, the scientific community has held a summit meeting to discuss the potential future of CRISPR technology, much in the same way the Asilomar Conference discussed recombinant DNA over 40 years ago (Baltimore et al. 2015; Berg et al. 1975a, b).

In HD, it is possible that a variety of CRISPR tools could prove beneficial for treatment. Previous studies have shown that a reduction in mutant HTT levels can ameliorate symptoms of the disease (Canals et al. 2004; Conforti et al. 2008; Zuccato et al. 2008). A recent study has shown reduction of mutant Huntingtin in cells by using TALE-ATFs (artificial transcription factors) to specifically target the mutant allele by targeting SNPs common on that allele. The TALE-ATF has a KRAB domain attached that represses transcription of the mutant Huntingtin allele (Fink et al. 2016). This technique has yet to be tried in Huntington model mice; however, previous studies have used ATFs to repress transcription in the brains of mice (Bailus et al. 2016). Another approach using CRISPR would involve increasing transcription of genes that could be neuroprotective in HD; BDNF could be a potential target for this type of therapy (Pollock et al. 2016). As screening studies are further refined using more genetically engineered isogenic cell lines, it will be possible to uncover additional gene regulation targets.

The ideal therapy for HD would involve gene replacement therapy, where the mutant allele would be replaced by a corrected donor allele. Using the CRISPR system, it will eventually be possible to do this correction in vivo. When designing the donor strand, it is possible to detect site-specific insertion by PCR if a small tag is added to the donor strand, allowing for optimization of different CRISPR components (Fig. 4). After design and condition optimization, there are still several

issues that need to be addressed to develop CRISPR into an in vivo therapy. One area to examine is the immune response, as Cas9 is not an endogenous protein in mammals, although there are mouse models that constitutively express Cas9 from birth (Platt et al. 2014). Previous studies in humans with zinc finger proteins have shown minimal immune response. Cas9 is not endogenous to animals and may elicit an immune response if given over an extended period of time. A second major concern for gene correction in vivo is the delivery of the CRISPR system to the desired organ or tissue. For certain diseases, it may be possible to directly inject the organ and correct only a subpopulation of the cells; for other diseases, especially those that effect the brain, delivery is more difficult (Liu et al. 2016; Yin et al. 2016). Direct injection into the brain is possible, and packaging the CRISPR system into an appropriately pseudotyped viral vector could allow for additional coverage beyond the injection point. The CRISPR system has been packaged into both AAV and lentivirus and used successfully in several mouse studies (Yin et al. 2016; Senis et al. 2014; Wang et al. 2015; Graham 2016). Nanoparticles and purified proteins are additional methods that have been used to successfully deliver CRISPR into cells and tissues (Wang et al. 2016; Ramakrishna et al. 2014). Each of these delivery methods has advantages and disadvantages, but with additional optimization successful gene replacement therapy in vivo should be possible. Since early HD diagnosis is possible, genetic correction therapy could be performed during the asymptomatic stage, potentially preventing onset of the disease.

Conclusion

Genome engineering is providing neuroscientists with new methods to address critical questions in the field and offers the hope for new treatments of neurological genetic diseases. The application of genetic engineering to disease modeling is accelerating efforts to understand the molecular mechanism of these diseases and offers new approaches to identifying therapeutic targets and drugs. The recent advances in genetic engineering allow for better modeling and understanding the role of SNPs in diseases with complex genetic alterations. These new genomic engineering technologies, which precisely alter the genome, are already offering insights into the complexity of the nervous system, its normal function and alterations in disease. Eventually these genome engineering technologies may correct the disease allele in human patients (in vivo) before symptoms manifest, resulting in therapy at the DNA level.

Acknowledgments Support for this work comes from NIH R01s NS094422 and NS100529.

References

- An MC, Zhang N, Scott G, Montoro D, Wittkop T, Mooney S, Melov S, Ellerby LM (2012) Genetic correction of Huntington's disease phenotypes in induced pluripotent stem cells. *Cell Stem Cell* 11:253–263
- An MC, O'Brien RN, Zhang N, Patra BN, De La Cruz M, Ray A, Ellerby LM (2014) Polyglutamine disease modeling: epitope based screen for homologous recombination using CRISPR/Cas9 System. *PLoS Curr* 6. doi: [10.1371/currents.hd.0242d2e7ad72225efa72f6964589369a](https://doi.org/10.1371/currents.hd.0242d2e7ad72225efa72f6964589369a)
- Bailus BJ, Pyles B, McAlister MM, O'Geen H, Lockwood SH, Adams AN, Nguyen JTT, Yu A, Berman RF, Segal DJ (2016) Protein delivery of an artificial transcription factor restores widespread Ube3a expression in an Angelman syndrome mouse brain. *Mol Ther* 24:548–555
- Bakondi B, Lv W, Lu B, Jones MK, Tsai Y, Kim KJ, Levy R, Akhtar AA, Breunig JJ, Svendsen CN, Wang S (2016) In vivo CRISPR/Cas9 gene editing corrects retinal dystrophy in the S334ter-3 rat model of autosomal dominant retinitis pigmentosa. *Mol Ther* 24:556–563
- Baltimore D, Berg P, Botchan M, Carroll D, Charo RA, Church G, Corn JE, Daley GQ, Doudna JA, Fenner JA, Greely HT, Jinek M, Martin GS, Penhoet E, Puck J, Sternberg SH, Weissman JS, Yamamoto KR (2015) Biotechnology. A prudent path forward for genomic engineering and germline gene modification. *Science* 348:36–38
- Beerli RR, Segal DJ, Dreier B, Barbas CF III (1998) Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc Natl Acad Sci USA* 95:14628–14633
- Berg P, Baltimore D, Brenner S, Roblin RO 3rd, Singer MF (1975a) Asilomar conference on recombinant DNA molecules. *Science* 188:991–994
- Berg P, Baltimore D, Brenner S, Roblin RO, Singer MF (1975b) Summary statement of the Asilomar conference on recombinant DNA molecules. *Proc Natl Acad Sci USA* 72:1981–1984
- Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326:1509–1512
- Calvanese V, Lara E, Kahn A, Fraga MF (2009) The role of epigenetics in aging and age-related diseases. *Ageing Res Rev* 8:268–276
- Canals JM, Pineda JR, Torres-Peraza JF, Bosch M, Martín-Ibañez R, Muñoz MT, Mengod G, Ernfors P, Alberch J (2004) Brain-derived neurotrophic factor regulates the onset and severity of motor dysfunction associated with enkephalinergic neuronal degeneration in Huntington's disease. *J Neurosci* 24:7727–7739
- Chen Y, Zheng Y, Kang Y, Yang W, Niu Y, Guo X, Tu Z, Si C, Wang H, Xing R, Pu X, Yang SH, Li S, Ji W, Li XJ (2015) Functional disruption of the dystrophin gene in rhesus monkey using CRISPR/Cas9. *Hum Mol Genet* 24:3764–3774
- Chinnery PF, Elliott HR, Hudson G, Samuels DC, Relton CL (2012) Epigenetics, epidemiology and mitochondrial DNA diseases. *Int J Epidemiol* 41:177–187
- Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ, Voytas DF (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186:757–761
- Conforti P, Ramos C, Apostol BL, Simmons DA, Nguyen HP, Riess O, Thompson LM, Zuccato C, Cattaneo E (2008) Blood level of brain-derived neurotrophic factor mRNA is progressively reduced in rodent models of Huntington's disease: restoration by the neuroprotective compound CEP-1347. *Mol Cell Neurosci* 39:1–7
- Cong L, Zhou R, Kuo YC, Cunniff M, Zhang F (2012) Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat Commun* 3:968
- Cornu TI, Cathomen T (2010) Quantification of zinc finger nuclease-associated toxicity. *Methods Mol Biol* 649:237–245
- Cornu TI, Thibodeau-Beganny S, Guhl E, Alwin S, Eichinger M, Joung JK, Cathomen T (2008) DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol Ther* 16:352–358

- Dreier B, Beerli RR, Segal DJ, Flippin JD, Barbas CF 3rd (2001) Development of zinc finger domains for recognition of the 5'-ANN-3' family of DNA sequences and their use in the construction of artificial transcription factors. *J Biol Chem* 276:29466–29478
- Eisenstein M (2012) Sangamo's lead zinc-finger therapy flops in diabetic neuropathy. *Nat Biotechnol* 30:121–123
- Fink KD, Deng P, Gutierrez J, Anderson JS, Torrest A, Komarla A, Kalomoiris S, Cary W, Anderson JD, Gruenloh W, Duffy A, Tempkin T, Annett G, Wheelock V, Segal DJ, Nolte JA (2016) Allele-specific reduction of the mutant Huntingtin allele using transcription activator-like effectors in human Huntington's disease fibroblasts. *Cell Transplant* 25:677–686
- Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaeppl C, Nowrouzi A, Bartholomae CC, Wang J, Friedman G, Holmes MC, Gregory PD, Glimm H, Schmidt M, Naldini L, von Kalle C (2011) An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol* 29:816–823
- Geurts AM, Cost GJ, Freyvert Y, Zeitler B, Miller JC, Choi VM, Jenkins SS, Wood A, Cui X, Meng X, Vincent A, Lam S, Michalkiewicz M, Schilling R, Foeckler J, Kalloway S, Weiler H, Ménoret S, Anegón I, Davis GD, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Jacob HJ, Buelow R (2009) Knockout rats via embryo microinjection of zinc-finger nucleases. *Science* 325:433
- Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, Lim WA, Weissman JS, Qi LS (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154:442–451
- Graham DM (2016) CRISPR/Cas9 corrects retinal dystrophy in rats. *Lab Anim (NY)* 45:85
- Guilinger JP, Pattanayak V, Reyon D, Tsai SQ, Sander JD, Joung JK, Liu DR (2014) Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity. *Nat Methods* 11:429–435
- Gupta A, Meng X, Zhu LJ, Lawson ND, Wolfe SA (2010) Zinc finger protein-dependent and -independent contributions to the in vivo off-target activity of zinc finger nucleases. *Nucleic Acids Res* 39:381–392
- Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, Cradick TJ, Marraffini LA, Bao G, Zhang F (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* 31:827–832
- Inglese J, Dranchak P, Moran JJ, Jang SW, Srinivasan R, Santiago Y, Zhang L, Guha R, Martinez N, MacArthur R, Cost GJ, Svaren J (2014) Genome editing-enabled HTS assays expand drug target pathways for Charcot-Marie-Tooth disease. *ACS Chem Biol* 9:2594–2602
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816–821
- Keatinge M, Bui H, Menke A, Chen YC, Sokol AM, Bai Q, Ellett F, Da Costa M, Burke D, Gegg M, Trollope L, Payne T, McTighe A, Mortiboys H, de Jager S, Nuthall H, Kuo MS, Fleming A, Schapira AH, Renshaw SA, Highley JR, Chacinska A, Panula P, Burton EA, O'Neill MJ, Bandmann O (2015) Glucocerebrosidase 1 deficient Danio rerio mirror key pathological aspects of human Gaucher disease and provide evidence of early microglial activation preceding alpha-synuclein-independent neuronal cell death. *Hum Mol Genet* 24:6640–6652
- Landles C, Bates GP (2004) Huntingtin and the molecular pathogenesis of Huntington's disease. Fourth in molecular medicine review series. *EMBO Rep* 5:958–963
- Langfelder P, Cantle PJ, Chatzopoulou D, Wang N, Gao F, Al-Ramahi I, XH L, Ramos EM, El-Zein K, Zhao Y, Deverasetty S, Tebbe A, Schaab C, Lavery DJ, Howland D, Kwak S, Botas J, Aaronson JS, Rosinski J, Coppola G, Horvath S, William Yang X (2016) Integrated genomics and proteomics define Huntingtin CAG length-dependent networks in mice. *Nat Neurosci* 19:623–633
- Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS (2013) CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc* 8:2180–2196

- Liang P, Xu Y, Zhang X, Ding C, Huang R, Zhang Z, Lv J, Xie X, Chen Y, Li Y, Sun Y, Bai Y, Songyang Z, Ma W, Zhou C, Huang L (2015) CRISPR/Cas9-mediated gene editing in human triprenuclear zygotes. *Protein Cell* 6:363–372
- Liu Q, Segal DJ, Ghiara JB, Barbas CF III (1997) Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *Proc Natl Acad Sci USA* 94:5525–5530
- Liu J, Gao C, Chen W, Ma W, Li X, Shi Y, Zhang H, Zhang L, Long Y, Xu H, Guo X, Deng S, Yan X, Yu D, Pan G, Chen Y, Lai L, Liao W, Li Z (2016) CRISPR/Cas9 facilitates investigation of neural circuit disease using human iPSCs: mechanism of epilepsy caused by an SCN1A loss-of-function mutation. *Transl Psychiatry* 6:e703
- Maeder ML, Angstman JF, Richardson ME, Linder SJ, Cascio VM, Tsai SQ, Ho QH, Sander JD, Reyon D, Bernstein BE, Costello JF, Wilkinson MF, Joung JK (2013a) Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nat Biotechnol* 31:1137–1142
- Maeder ML, Linder SJ, Reyon D, Angstman JF, Fu Y, Sander JD, Joung JK (2013b) Robust, synergistic regulation of human gene expression using TALE activators. *Nat Methods* 10:243–245
- Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. *Science* 326:1501
- Nagaoka M, Sugiura Y (2000) Artificial zinc finger peptides: creation, DNA recognition, and gene regulation. *J Inorg Biochem* 82:57–63
- Nakamura K, Fujii W, Tsuboi M, Tanihata J, Teramoto N, Takeuchi S, Naito K, Yamanouchi K, Nishihara M (2014) Generation of muscular dystrophy model rats with a CRISPR/Cas system. *Sci Rep* 4:5635
- Nelson CE et al (2016) In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* 351:403–407
- O'Brien R, DeGiacomo F, Holcomb J, Bonner A, Ring KL, Zhang N, Zafar K, Weiss A, Lager B, Schilling B, Gibson BW, Chen S, Kwak S, Ellerby LM (2015) Integration-independent transgenic Huntington Disease fragment mouse models reveal distinct phenotypes and life span in vivo. *J Biol Chem* 290:19287–19306
- O'Geen H, Henry IM, Bhakta MS, Meckler JF, Segal DJ (2015) A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. *Nucleic Acids Res* 43:3389–3404
- Pabo CO, Peisach E, Grant RA (2001) Design and selection of novel Cys2His2 zinc finger proteins. *Annu Rev Biochem* 70:313–340
- Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, Dahlman JE, Parnas O, Eisenhaure TM, Jovanovic M, Graham DB, Jhunjhunwala S, Heidenreich M, Xavier RJ, Langer R, Anderson DG, Hacohen N, Regev A, Feng G, Sharp PA, Zhang F (2014) CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* 159:440–455
- Pollock L, Dahlenburg H, Nelson H, Fink KD, Cary W, Hendrix K, Annett G, Torrest A, Deng P, Gutierrez J, Nacey C, Pepper K, Kalomoiris S, D Anderson J, McGee J, Gruenloh W, Fury B, Bauer G, Duffy A, Tempkin T, Wheelock V, Nolte JA (2016) Human mesenchymal stem cells genetically engineered to overexpress brain-derived neurotrophic factor improve outcomes in Huntington's disease mouse models. *Mol Ther* 24:965–977
- Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152:1173–1183
- Ramakrishna S, Kwaku Dad AB, Beloor J, Gopalappa R, Lee SK, Kim H (2014) Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Res* 24:1020–1027
- Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154:1380–1389

- Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS, Koonin EV, Sharp PA, Zhang F (2015) In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* 520:186–191
- Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE (2016) Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol* 34:339–344
- Ring KL, An MC, Zhang N, O'Brien RN, Ramos EM, Gao F, Atwood R, Bailus BJ, Melov S, Mooney SD, Coppola G, Ellerby LM (2015) Genomic analysis reveals disruption of striatal neuronal development and therapeutic targets in human Huntington's disease neural stem cells. *Stem Cell Rep* 5:1023–1038
- Rivenbark AG, Stolzenburg S, Beltran AS, Yuan X, Rots MG, Strahl BD, Blancafort P (2012) Epigenetic reprogramming of cancer cells via targeted DNA methylation. *Epigenetics* 7:350–360
- Sander JD, Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol* 32:347–355
- SangamoBioSciences (2001) Sangamo BioSciences announces presentation of first preclinical data using ZFP transcription factors. <http://www.sangamo.com/053101giordano.htm>
- Segal DJ, Barbas CF III (2000) Design of novel sequence-specific DNA-binding proteins. *Curr Opin Chem Biol* 4:34–39
- Segal DJ, Barbas CF III (2001) Custom DNA-binding proteins come of age: polydactyl zinc-finger proteins. *Curr Opin Biotechnol* 12:632–637
- Segal DJ, Dreier B, Beerli RR, Barbas CF III (1999) Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. *Proc Natl Acad Sci USA* 96:2758–2763
- Senis E, Fatouros C, Große S, Wiedtke E, Niopek D, Mueller AK, Börner K, Grimm D (2014) CRISPR/Cas9-mediated genome engineering: an adeno-associated viral (AAV) vector toolbox. *Biotechnol J* 9:1402–1412
- Soldner F, Laganière J, Cheng AW, Hockemeyer D, Gao Q, Alagappan R, Khurana V, Golbe LI, Myers RH, Lindquist S, Zhang L, Guschin D, Fong LK, BJ V, Meng X, Urnov FD, Rebar EJ, Gregory PD, Zhang HS, Jaenisch R (2011) Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell* 146:318–331
- Tabebordbar M, Zhu K, Cheng JK, Chew WL, Widrick JJ, Yan WX, Maesner C, Wu EY, Xiao R, Ran FA, Cong L, Zhang F, Vandenbergh LH, Church GM, Wagers AJ (2016) In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* 351:407–411
- Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* 435:646–651
- Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153:910–918
- Wang X, Wang Y, Wu X, Wang J, Wang Y, Qiu Z, Chang T, Huang H, Lin RJ, Yee JK (2015) Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat Biotechnol* 33:175–178
- Wang M, Zuris JA, Meng F, Rees H, Sun S, Deng P, Han Y, Gao X, Pouli D, Wu Q, Georgakoudi I, Liu DR, Xu Q (2016) Efficient delivery of genome-editing proteins using bio-reducible lipid nanoparticles. *Proc Natl Acad Sci USA* 113:2868–2873
- Wiedenheft B, Sternberg SH, Doudna JA (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482:331–338
- Wolfe SA, Ramm EI, Pabo CO (2000) Combining structure-based design with phage display to create new Cys(2)His(2) zinc finger dimers. *Struct Fold Des* 8:739–750
- Wolffe E (2016) Corporate profile: Sangamo BioSciences, Inc. *Regen Med* 11:375–379

- Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB, Cheng AW, Trevino AE, Konermann S, Chen S, Jaenisch R, Zhang F, Sharp PA (2014) Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol* 32:670–676
- Xia G, Gao Y, Jin S, Subramony SH, Terada N, Ranum LP, Swanson MS, Ashizawa T (2015) Genome modification leads to phenotype reversal in human myotonic dystrophy type 1 induced pluripotent stem cell-derived neural stem cells. *Stem Cells* 33:1829–1838
- Xu L, Park KH, Zhao L, Xu J, El Refaey M, Gao Y, Zhu H, Ma J, Han R (2016) CRISPR-mediated genome editing restores dystrophin expression and function in mdx mice. *Mol Ther* 24:564–569
- Yin H, Song CQ, Dorkin JR, Zhu LJ, Li Y, Wu Q, Park A, Yang J, Suresh S, Bizhanova A, Gupta A, Bolukbasi MF, Walsh S, Bogorad RL, Gao G, Weng Z, Dong Y, Koteliensky V, Wolfe SA, Langer R, Xue W, Anderson DG (2016) Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. *Nat Biotechnol* 34:328–333
- Zuccato C, Marullo M, Conforti P, MacDonald ME, Tartari M, Cattaneo E (2008) Systematic assessment of BDNF and its receptor levels in human cortices affected by Huntington's disease. *Brain Pathol* 18:225–238
- Zykovich A, Korf I, Segal DJ (2009) Bind-n-Seq: high-throughput analysis of in vitro protein-DNA interactions using massively parallel sequencing. *Nucleic Acids Res* 37:e151

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Therapeutic Gene Editing in Muscles and Muscle Stem Cells

Mohammadsharif Tabebordbar, Jason Cheng, and Amy J. Wagers

Abstract Duchenne muscular dystrophy (DMD) is a devastating, degenerative muscle disease that affects ~1 in every 3500 male births. DMD arises from mutations in the *DMD* gene that prevent expression of its encoded protein, Dystrophin (Burghes et al. Nature 328:434–437, 1987). Interestingly, patients with *Dmd* mutations that delete certain segments of the Dystrophin coding region, but maintain protein reading frame, have a much milder form of the disease, known as Becker Muscular Dystrophy (BMD). This observation has spurred interest in developing “exon skipping” strategies in which certain mutation-containing or mutation-adjacent *Dmd* exons are intentionally removed in order to restore protein reading frame, and thereby Dystrophin expression, in DMD patients (Beroud et al. Hum Mutat 28:196–202, 2007; Yokota et al. Expert Opin Biol Ther 7:831–842, 2007).

Recently our lab (Tabebordbar et al. 2015) and others (Long et al. 2015; Nelson et al. 2015) reported a novel strategy to accomplish permanent sequence-specific modification of the *Dmd* gene in vivo in the *mdx* mouse model of DMD. This strategy utilizes the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-Cas9 RNA guided gene editing system, delivered using adeno-associated virus (AAV) vectors. We showed that, when administered systemically, AAV-*Dmd*-CRISPR enables sequence-targeted genome modification in each of the key affected cell types and organs of DMD model *mdx* mice, including cardiomyocytes, skeletal muscle fibers and endogenous muscle stem cells (Tabebordbar et al. 2015). Gene editing in these cells restores Dystrophin protein reading frame and expression, recovers muscle contractile function, enhances muscle resilience in the face of controlled muscle damage, and establishes a pool of therapeutically modified progenitors that can participate in subsequent muscle regenerative events.

These studies provided a critical advance in allowing programmable genome editing that can irreversibly modify disease-causing mutations in the affected tissues of dystrophic individuals. Moreover, the results represent critical proof-of-concept evidence demonstrating the feasibility of systemic gene editing in vivo, which has the potential

M. Tabebordbar • J. Cheng • A.J. Wagers (✉)

Department of Stem Cell and Regenerative Biology, Harvard University and Harvard Stem Cell Institute, Cambridge, MA 02138, USA

e-mail: amy_wagers@harvard.edu

to recover Dystrophin expression in up to 80% of patients with DMD (Beroud et al. 2007; Yokota et al. 2007). Yet, important challenges remain for the future therapeutic application of *Dmd*-CRISPR gene editing, including enhancing the efficiencies with which gene editing may be accomplished in muscle fibers and satellite cells and circumventing the possible emergence of a host immune response to the bacterial Cas9 endonuclease, which could interfere with gene editing and/or lead to elimination of gene-edited cells. Overcoming these challenges will be crucial for developing clinically relevant strategies to accomplish safe, efficient and durable in vivo gene editing for DMD.

Duchenne Muscular Dystrophy

Duchenne muscular dystrophy (DMD) is one of the most common X-linked genetic disorders in humans; it arises from point mutations, deletions or duplications in the *DMD* gene that prevent expression of its encoded protein, Dystrophin (Burghes et al. 1987; Koenig et al. 1987). Dystrophin is an essential structural protein in skeletal and cardiac muscle (Ervasti and Campbell 1991). Its primary function is to link the cytoskeleton of muscle fibers to the extracellular matrix and thereby stabilize the muscle fiber membrane (Straub et al. 1992). Absence of functional Dystrophin protein increases the susceptibility of dystrophic muscle fibers to contraction-induced injury (Campbell and Kahl 1989). Increased cytosolic calcium following mechanical stress, activation of proteases (particularly calpains), destruction of membrane constituents and ultimately muscle fiber necrosis occur frequently in dystrophic muscles (reviewed in Tabebordbar et al. 2013). Due to continual myofiber destruction in dystrophic muscle, the resident pool of regenerative muscle stem cells (known as “satellite cells”) must support repeated rounds of activation and regeneration in an attempt to compensate for ongoing damage. As the disease advances, satellite cells show reduced capacity to regenerate muscle, possibly due to cell-intrinsic defects (Dumont et al. 2015) or proliferation-induced reductions in telomere length (Sacco et al. 2010). Absent an adequate regenerative response, fat and fibrotic tissue replace muscle fibers, leading to further weakening and wasting (Wallace and McNally 2009).

Current Gene-Targeted Therapeutic Strategies for DMD

Current treatment options for DMD are disappointingly limited and focus mainly on managing symptoms and suppressing the immune and inflammatory response (Muir and Chamberlain 2009; Partridge 2011). Patients are typically diagnosed at 3–5 years of age, they are wheelchair-bound in their second decade, and they have an average life expectancy of only about 30 years. In contrast, a related group of patients with mutations that impact this same gene but maintain its open reading frame produce an internally deleted but still partially functional Dystrophin protein that results in a markedly less severe disease known as Becker Muscular Dystrophy

(BMD; England et al. 1990; Nakamura et al. 2008; Taglia et al. 2015). Many BMD patients are not diagnosed until adolescence or even adulthood and some enjoy a normal life span. These observations have provided motivation for the generation of rationally modified, truncated versions of Dystrophin for therapeutic application in DMD, including engineered “microdystrophins” and endogenous exon “skipped” *DMD* mRNAs.

The extremely large size of the *DMD* gene (2.4 Mb) and its encoded mRNA (14 kb) makes it very difficult, if not impossible, to package full-length dystrophin expression cassettes into clinically relevant viral vectors such as Adeno-associated viruses (AAVs), which have a packaging capacity of <5 kb. This limitation has propelled the generation of truncated mini- (6–8 kb) and microdystrophin (<4 kb) genes (Harper et al. 2002), which reduce the Dystrophin protein to its most essential functional elements. These rationally designed microgenes delete large regions of the internal Rod domain of Dystrophin, which contains 24 spectrin-like repeats and comprises 80% of the overall protein (Chamberlain 2002), while maintaining much of its functional integrity. Microdystrophin genes can be packaged into viral vectors for exogenous delivery and ectopic expression from ubiquitous or muscle-specific promoters (Fabb et al. 2002; Gregorevic et al. 2004), and delivery by AAVs results in effective expression of protein products that correctly localize to the sarcolemma and recruit other dystrophin glycoprotein complex (DGC) proteins. Importantly, while microdystrophins are not equivalent in function to full-length Dystrophin, they have been shown to ameliorate DMD pathologies in *mdx* mice (Harper et al. 2002; Wang et al. 2000) and dystrophin-deficient canine models (Shin et al. 2013). A related approach—exon skipping—similarly generates a modified Dystrophin protein product, but in this case the endogenous *Dmd* pre-mRNA transcript is targeted to remove mutation-carrying and/or mutation-adjacent exons from the mRNA. By choosing specific exons for removal, exon skipping approaches are able to generate *Dmd* mRNAs with restored reading frame.

In both gene complementation by microdystrophin and exon skipping approaches, the overall goal is to convert a severe DMD mutation, lacking Dystrophin protein expression entirely, into a milder BMD-like one, via expression of a truncated but still partially functional protein. It has been estimated that exon skipping strategies for *Dmd* could ultimately provide significant therapeutic benefit to the majority (~80%) of existing DMD patients (Beroud et al. 2007; Yokota et al. 2007), while complementation by ectopic expression of microdystrophin could in theory be useful for any mutation that abrogates Dystrophin protein production.

Challenges for Therapeutic Exon Skipping and Microdystrophin Delivery Strategies

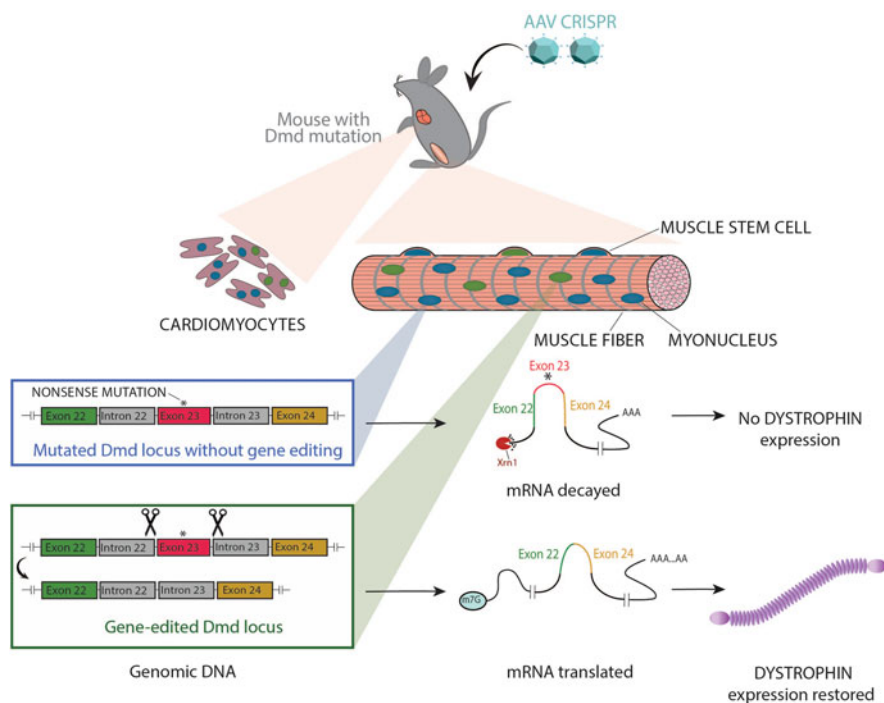
Clinical application of exon skipping approaches to date has relied on antisense oligonucleotides (AONs) designed to mask splice donor and acceptor sequences in mutation-affected or mutation-adjacent exons. However, for many of the therapeutically relevant *Dmd* exons, exon-skipping AONs have not yet been developed or have

not progressed to clinical trials (<https://www.sarepta.com/our-pipeline>). In addition, in a recent clinical trial, AON-mediated skipping of *Dmd* exon 51 failed to achieve sufficient rescue of Dystrophin protein to meet predetermined clinical endpoints (Lu et al. 2014). Importantly, even with relatively stable chemistries (Goyenvalle et al. 2015), AONs have a defined half-life (Goyenvalle et al. 2015; Vila et al. 2015) and require repeated (weekly) administration. This need for recurrent treatment increases the cost and potential side effects of AON therapy. Also, delivery of AONs to cardiac muscle has been more challenging than delivery to skeletal muscle, and delivery to resident muscle stem cells, if it occurs, is unlikely to be effective due to the dilution of AONs that occurs during cell proliferation. Thus, any benefit from AON therapy in satellite cells would be lost during muscle regenerative responses, which require proliferation of satellite cells and their progeny. Strategies in which AONs are delivered virally, by embedding within small nuclear RNAs, appear to suffer from similar progressive loss of the viral genome and its encoded AONs from dystrophic muscles (Vulin et al. 2012; Le Hir et al. 2013).

Relatedly, exogenous gene supplementation therapies using partially functional engineered microdystrophin constructs have encountered some challenges in clinical application. An initial Phase I clinical trial of microdystrophin gene therapies in human DMD patients yielded suboptimal transgene expression despite continued presence of vector genomes, possibly due to pre-existing or acquired T cell-mediated immune responses to dystrophin epitopes or AAV capsid proteins (see below), disease-associated inflammatory responses, CMV promoter silencing, or low AAV tropism (Bowles et al. 2012; Mendell et al. 2010). Additional Phase I trials of microdystrophin therapies utilizing different gene regulatory elements and AAV serotypes are currently underway (ClinicalTrials.gov Identifier: NCT02376816), and may mitigate these concerns; however, similar to AON delivery, delivery of AAV-microdystrophin to muscle satellite cells is unlikely to result in sustained transgene production, as the episomal AAV genome will be diluted with successive cell divisions. These challenges that have been encountered in the development of effective AON and microdystrophin therapies highlight the need for further evaluation of alternative strategies that could provide an efficient, permanent, one-time, systemic treatment to restore expression of Dystrophin in skeletal and cardiac muscles, as well as muscle satellite cells, of DMD patients.

Gene-Editing Approaches to Restore Dystrophin Function in DMD

In a recent report (Tabebordbar et al. 2015), we described a novel genome-targeted editing approach (Fig. 1), based on *Dmd* exon skipping approaches, that was designed to accomplish irreversible removal of a mutated segment of the *Dmd* gene in the affected tissues of *mdx* mice, an animal model of human DMD (Sicinski et al. 1989). We further showed that this approach resulted in production of functional Dystrophin protein and improved muscle stability and contractility (Tabebordbar et al. 2015). Our



Tabebordbar, Zhu et al., Science, 2015

Fig. 1 Gene-editing strategy for recovery of Dystrophin expression in DMD model mice. *Mdx* mice with a mutation in the *Dmd* gene were injected with AAV particles carrying clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 endonucleases and paired guide RNAs targeting the mutated *Dmd* exon23. This procedure led to excision of the targeted DNA and restored *Dmd* gene reading frame and Dystrophin expression in gene-edited skeletal muscle fibers, cardiomyocytes and muscle stem cells following local delivery or delivery via the bloodstream, in dystrophic mice. Gene-edited nuclei are shown in green and non-edited nuclei are shown in blue. The mutated *Dmd* mRNA is degraded and Dystrophin expression is lacking in the dystrophic tissues of untreated mice (graphical summary describes data reported in Tabebordbar et al. 2015. See text for details)

approach made use of the CRISPR-Cas9 gene editing system, which allows the introduction of user-defined “cuts” in the genome. Each CRISPR-Cas9 gene editing complex consists of a Cas9 endonuclease and a programmable guide RNA (gRNA) that probes the genome for protospacer-adjacent motifs (PAM) [e.g., –NGG (Ran et al. 2013a) or –NNGRR(T) (Ran et al. 2015)]. Upon PAM recognition and base-pairing of the gRNA with an adjacent complementary DNA sequence, Cas9 creates a double-strand break (DSB) in the genomic DNA. Introduction of DSBs at two sites in the same linear stretch of DNA favors excision of the intervening sequence, and repair of this lesion by non-homologous end joining (NHEJ) juxtaposes the remaining 5' and 3' sequences (Canver et al. 2014; Tabebordbar et al. 2015). Alternatively, inclusion of a homologous donor template enables repair by homology directed recombination

(HDR), leading to incorporation of precise nucleotide changes, encoded in the donor template, at the site of the DSB. Changes introduced by HDR can range from a single base pair to insertions of entire genes or even large cassettes of multiple genes (Urnov et al. 2005; Ding et al. 2013; Voit et al. 2014). Significantly, the relative activity of NHEJ and HDR repair mechanisms can vary with cell type, cell cycle and developmental stage, which can have important ramifications for the efficacy and outcome of therapeutic genome modification (Yang et al. 2016).

CRISPR-Cas9 RNA guided endonucleases (RGENs) have been used to target both expressed and non-expressed genes in multiple cell types from multiple organisms both in vitro (Cho et al. 2013; Cong et al. 2013; DiCarlo et al. 2013; Ding et al. 2013; Friedland et al. 2013; Hwang et al. 2013; Mali et al. 2013; Wu et al. 2016) and in vivo (Ding et al. 2014; Xue et al. 2014; Yin et al. 2014; Ran et al. 2015; Yang et al. 2016). Published data demonstrate the utility of this system for multi-organ gene targeting of many distinct cell lineages, including hepatocytes, muscle fibers, cardiomyocytes, and muscle regenerative stem cells (Long et al. 2015; Nelson et al. 2015; Ran et al. 2015; Tabebordbar et al. 2015; Yang et al. 2016). We adapted the CRISPR-Cas9 system for *Dmd* editing in cardiac and skeletal muscle in vivo by utilizing a smaller Cas9 ortholog from *Staphylococcus aureus* (SaCas9), which could be packaged into recombinant AAV particles using the muscle-tropic serotype 9 (Zincarelli et al. 2008). Our strategy (Fig. 1) employed a dual AAV system (termed “AAV-*Dmd*-CRISPR”), which, due to AAV packaging limitations, was superior to single vector systems in terms of gene editing efficiency (Tabebordbar et al. 2015). In the dual system (Tabebordbar et al. 2015), the first AAV delivers SaCas9, driven by a strong CMV promoter, whereas the second AAV carries two gRNAs that target sequences in the introns flanking mouse *Dmd* exon 23 (“*Dmd*23 gRNAs”), each driven by a U6 promoter. This targeting of intronic sequences is important because it allows for tolerance of small insertions and deletions that are common with NHEJ-mediated repair of DNA DSBs (Symington and Gautier 2011). When injected intramuscularly or systemically into adult (P42) or early postnatal (P3) recipient mice, which carry a nonsense mutation (*mdx*) in *Dmd* exon 23, AAV-*Dmd*-CRISPR caused excision of exon 23 in heart cells (cardiomyocytes), skeletal muscle fibers and muscle stem cells [satellite cells, marked by transgenic expression of the fluorescent zsGreen protein from the Pax7 promoter (Bosnakovski et al. 2008)], producing an exon 23-deleted Dystrophin mRNA that, when translated, generated a truncated but functional Dystrophin protein (Tabebordbar et al. 2015; Fig. 1). Dystrophin protein restoration in AAV-*Dmd*-CRISPR treated *mdx* mice improved structural and functional aspects of the muscle, increased muscle strength and improved resistance to eccentric contraction-induced damage. Importantly, AAV-*Dmd*-CRISPR gene editing complexes could be disseminated systemically and were functional in both neonatal and adult mice. Exon-deleted transcripts represented almost 50% of total *Dmd* mRNA in muscle after intramuscular delivery in adults and 5–15% in skeletal and cardiac muscles after systemic delivery in neonates (Tabebordbar et al. 2015).

Importantly, and emphasizing the robustness and reproducibility of these results, similar outcomes were reported simultaneously by two other groups (Long et al. 2015; Nelson et al. 2015) using different Cas9 proteins and regulatory elements (Long et al. 2015), different AAV serotypes (Nelson et al. 2015), different routes of systemic

administration (Long et al. 2015), and different gRNAs (Long et al. 2015; Nelson et al. 2015). All three groups reported gene editing in skeletal muscle fibers and cardiomyocytes, with efficiencies in skeletal muscle reported by Long et al. and Nelson et al. to vary from 1 to 67% Dystrophin + fibers, depending on the delivery approach used (local vs. systemic), dose of virus and age of the recipient animals. Long et al. also documented *Dmd* modification in vascular smooth muscle cells but not in brain, and our group, as discussed above, demonstrated detectable editing in endogenous muscle satellite cells (Tabebordbar et al. 2015). Finally, by analyzing treated muscle tissues at 4, 8, and 12 weeks after AAV injection, Long et al. ascertained that the percentage of dystrophin-positive myofibers might increase over time, and Nelson et al. observed that dystrophin restoration could be maintained for at least 6 months after treatment, indicating the potential long-term efficacy of AAV-*Dmd*-CRISPR therapies. Promisingly, differences in experimental design among the three studies and the varying efficiencies obtained suggest that multiple parameters may be adjusted and optimized to enhance genomic editing and increase dystrophin protein expression levels for more effective treatment of disease phenotypes by *Dmd*-CRISPR.

In summary, published work from our lab and others provides strong evidence supporting the efficacy of in vivo genome editing to correct disruptive mutations in DMD in a relevant dystrophic mouse model (Long et al. 2015; Nelson et al. 2015; Tabebordbar et al. 2015). These data indicate that programmable CRISPR complexes can be delivered locally and systemically to terminally differentiated skeletal muscle fibers, cardiomyocytes and smooth muscle cells, as well as regenerative muscle satellite cells, in neonatal and adult mice, where they mediate targeted gene modification, restore Dystrophin expression and partially recover functional deficiencies of dystrophic muscle. As prior studies in mice and humans indicate that Dystrophin levels as low as 3–15% of wild type are sufficient to ameliorate pathologic symptoms in the heart and skeletal muscle (van Putten et al. 2012, 2013, 2014; Long et al. 2014), and levels as low as 30% can completely suppress the dystrophic phenotype (Neri et al. 2007), the level of Dystrophin expression that is potentially achievable by one-time administration of AAV-*Dmd*-CRISPR urges further development of this system, which could be used independently or together with other therapies, including AON-mediated exon skipping (Aartsma-Rus et al. 2009) and AAV-mediated delivery of engineered “microdystrophins” (Harper et al. 2002; Ramos and Chamberlain 2015), as discussed above.

Remaining Challenges for Therapeutic Development of *DMD*-CRISPR

Taken together, the rodent studies described above provide strong pre-clinical proof-of-concept data that should inspire further evaluation and optimization of AAV-CRISPR as a new therapeutic option for DMD patients, either as a stand-alone intervention or in conjunction with other existing DMD therapies. Below, we discuss a number of

challenges that remain to be overcome before realizing the potential of this approach in human patients.

Challenges of *DMD-CRISPR* Delivery

Engineered recombinant AAVs are particularly attractive vectors for both local and systemic delivery of gene editing complexes due to their general non-pathogenicity in human populations, their relatively low immunogenicity, and their inability to integrate efficiently into the genome (Gao et al. 2004; Boutin et al. 2010). Because of these traits, AAVs are currently in use in several human clinical trials (Mingozzi and High 2011; Kotterman and Schaffer 2014), and the immune response to AAV vectors has been extensively studied in both animal models and humans. Because engineered AAV vectors do not replicate and do not encode viral proteins, immune responses to AAVs are directed solely at the viral capsid and exhibit a relatively low pro-inflammatory profile (Mingozzi and High 2011). While pre-existing and acquired immunity to AAV remains a challenge for systemic, and repeated, administration of AAV vectors in human populations, these issues have been investigated for several decades and promising pharmacologic and physical strategies have emerged (Mingozzi and High 2011). In addition, clinical responses to AAV administration have been monitored in hundreds of human subjects, with little evidence as yet of acute adverse events (Mingozzi and High 2011). Thus, the successful application of AAV-mediated therapy in multiple human trials suggests that the immune response to AAV itself is unlikely to preclude gene editing therapies based on AAV delivery.

Still, a clear limitation of current AAV systems is that levels of gene targeting achieved in mouse models by AAV-mediated delivery of CRISPR-Cas9 to muscle satellite cells are rather low (<5% of satellite cells targeted; Tabebordbar et al. 2015), suggesting a need to investigate additional AAV serotypes to identify those with optimal tropism for satellite cells. Directed evolution and in vivo selection have been used recently to engineer novel AAV capsids with high tropism for tissues that are difficult to transduce with naturally occurring AAVs, such as human hepatocytes in a xenograft liver model (Lisowski et al. 2014) and the outer retina after injection into the eye's vitreous humor (Dalkara et al. 2013). In addition, transduction rates of blood-forming hematopoietic stem cells have been improved through incorporation of novel amino acid substitutions in capsids (Song et al. 2013a, b). Thus, the application of directed evolution and in vivo selection strategies for generating novel AAV serotypes with high tropism for satellite cells represents an exciting future direction for increasing gene-editing efficiencies in these cells in vivo.

On the other hand, the development of alternative delivery strategies that enable transient expression of *DMD-CRISPR* may hold some advantages, particularly since the therapeutic effect of gene-editing approaches does not depend on persistent expression of Cas9 and gRNAs. Transient expression of CRISPR components could mitigate several of the possible adverse effects associated with prolonged Cas9 exposure, including potential genomic toxicity and immunogenicity (Wang et al. 2015). Indeed,

in vitro experiments indicate that transient expression of Cas9 does produce lower off-target effects (Kim et al. 2014; Zuris et al. 2015).

Recent advances in lipid nanoparticle-mediated delivery of Cas9:gRNA complexes in vitro (Kim et al. 2014; Woo et al. 2015; Zuris et al. 2015) and Cas9 mRNA in vivo (Yin et al. 2016) provide additional promising avenues that may circumvent the challenges of AAV immunity. Delivering Cas9 and gRNAs conjugated with cell penetrating peptides (CPPs) has also been useful in targeting gene-editing complexes to human cell lines in culture (Ramakrishna et al. 2014), and combining this approach with incorporation of novel muscle-homing peptides (Gao et al. 2014) could potentially be effective for in vivo delivery of *DMD*-CRISPR.

Potential Immune Response to Restored Dystrophin Protein

A possible immune response to the repaired *DMD* protein is also of potential concern for clinical application of *DMD*-CRISPR-mediated gene editing; however, due to large variations in the types of *DMD* mutations seen in patients (Aartsma-Rus et al. 2009), it is likely that the nature of individual immune responses to Dystrophin protein will vary as well and will depend at least in part on the nature of the mutation and the frequency with which “natural” exon skipping, which gives rise to revertant fibers in both *DMD* patients and *mdx* mice (Hoffman et al. 1990; Burrow et al. 1991; Klein et al. 1992; Nicholson et al. 1993; Fanin et al. 1995; Uchino et al. 1995; Lu et al. 2000), may allow for endogenous exposure and tolerance to near-full length Dystrophin. Interestingly, in gene therapy trials for hemophilia B, in which AAV vectors were used to deliver Factor IX (F. IX), no subjects developed immune responses against the F.IX transgene, even though some carried null mutations in the F.IX gene (Manno et al. 2006; Nathwani et al. 2011). Similarly, promising results from studies using “microdystrophin” in mice and primates suggest that this protein is effectively expressed for up to 5 months without overt T cell or cytokine responses (Rodino-Klapac et al. 2010). These data argue that acquired immunity against the therapeutic protein also may not be therapy limiting. On the other hand, results from a clinical trial using intramuscular AAV-mediated delivery of microdystrophin, expressed under the control of a ubiquitous CMV promoter, revealed the presence in some patients of T cells recognizing self and non-self Dystrophin epitopes (Mendell et al. 2010). Interestingly, these T cells were present both before and after vector injection in two of the six patients, raising the possibility that screening for pre-existing immunity to Dystrophin protein in larger cohorts of *DMD* patients could provide useful information relevant to patient inclusion and exclusion criteria in future trials. Anti-Dystrophin antibodies were not detected in any of the treated patients; however, detection of Dystrophin-specific T cells and a lack of transgene expression in muscles of patients injected with AAV-microdystrophin (with the exception of two patients analyzed 6 weeks after injection) may suggest a cytotoxic response against fibers expressing microdystrophin. Thus, currently available data

point to a compelling need for further studies to investigate more deeply the potential immune response to restored Dystrophin expression in dystrophic muscle.

Pre-existing and Acquired Immunity to Cas9

Potential immunity to the Cas9 endonuclease is also a significant consideration for therapeutics development in humans. An essential component of the CRISPR-based gene-editing machinery, Cas9 is a bacterially derived protein whose expression in transduced cells can evoke both humoral and cellular responses (Wang et al. 2015; and see below). Additionally, about 20% of individuals in the human population are persistent carriers of *Staphylococcus aureus* and another 60% have been periodic carriers at some point in their lives (Kluytmans et al. 1997). Thus, a significant fraction of potential patients is likely to have been exposed to the Cas9 protein from this species, raising the possibility that a pre-existing anti-Cas9 immune response could modulate the efficacy of CRISPR-mediated gene editing for recovery of Dystrophin expression in dystrophic muscles. Moreover, as emerging data suggest that the immune system and its products can modulate the expression of AAV-encoded transgenes (Mingozzi and High 2011), as well as components of cellular DNA damage response pathways (Jackson and Bartek 2009; Calvo et al. 2012), Cas9-induced immune responses could potentially alter both the degree of on-target *DMD* editing and the frequency and types of off-target modifications induced. Thus, while studies in our lab and others (Long et al. 2015; Nelson et al. 2015; Tabebordbar et al. 2015) clearly demonstrate that anti-Cas9 immunity does not preclude gene editing in vivo, Cas9 immune responses could, nevertheless, have profound implications for the persistence of therapeutic benefit in the muscle and other tissues. We therefore believe that it is particularly important at this juncture to begin to assess the nature and consequences of the immune response to the foreign Cas9 protein itself and to determine whether preventing or ameliorating this response might improve the efficiency, durability, repeatability and/or safety of Cas9-mediated therapeutic gene editing.

Assessing Mutagenic Events at On-Target and Off-Target Sites

Off-target modifications pose a potential threat for gene editing approaches because the unintended activity of CRISPR-Cas9 at these locations can cause pathogenic modifications that impair cellular function or promote tumorigenesis. Furthermore, because in general Cas9-induced DNA DSBs can be repaired by either HDR or NHEJ, editing can result in different outcomes, depending on the number of alleles affected and the type of modification introduced. Thus, it is critical to develop tools

that enable facile assessment of mutagenic potential in an un-biased genome-wide manner, since such evaluations are likely to show patient- and gRNA-specific variation.

Recent advances have developed several different strategies to reduce genome-wide off-target mutations of *Streptococcus pyogenes* Cas9 (SpCas9). These strategies include use of paired SpCas9 nickases (Ran et al. 2013b), gRNAs with reduced length of the guide sequence (Fu et al. 2014) and the engineering of SpCas9 variants with amino acid substitutions in the DNA binding domain that reduce off-target rates (Kleinstiver et al. 2016; Slaymaker et al. 2016). Yet, there is still need for improving the specificity of SpCas9 and its smaller orthologs (e.g., SaCas9), and this issue is particularly important for the targeting of muscle stem cells, which have substantial proliferative capacity. The risk of generating undesired and deleterious mutations at proto-oncogene loci or at loci critical to stem cell function by CRISPR-Cas9 transduction of these cells must be rigorously analyzed before proceeding further with clinical translation of gene editing for DMD.

Enabling HR for Precise Repair of *Dmd*

Prior work in mice demonstrates that DMD pathology in skeletal muscle can be reversed by transplantation of sorted muscle stem cells isolated from wild-type animals (carrying a normal copy of the *Dmd* gene; Cerletti et al. 2008; Sacco et al. 2010). However, muscle stem cells are extremely rare, cannot be expanded effectively ex vivo, must be delivered by intramuscular injection (as they fail to migrate to muscle tissue when injected intravenously), and do not engraft cardiac muscle, which also is affected by *Dmd* mutation. These significant complications have limited the application of stem cell transplantation therapy to DMD, despite promising results in individually injected muscle groups.

Likewise, as discussed above, recent reports document the feasibility of AAV-based delivery of gene-editing complexes into cardiac and skeletal muscle in vivo and demonstrate that this system could be used to specifically excise a mutated segment of the *Dmd* gene in *mdx* mice to restore *Dmd* reading frame and allow production of a partially functional Dystrophin protein that improves muscle stability and contractility (Long et al. 2015; Nelson et al. 2015; Tabebordbar et al. 2015). However, it is important to note that the “first-generation” gene-editing strategies applied in these studies do not produce a full length Dystrophin. Instead, these approaches generate an internally truncated protein analogous to that seen in patients with BMD. While BMD is a markedly less severe disease compared to DMD, BMD patients still experience muscle pathology, and so, while clearly providing a potential clinical benefit, this approach is not fully curative for DMD. For this reason, future studies should be aimed at achieving full restitution of Dystrophin protein expression through precise gene editing to restore the normal *Dmd* gene sequence. Importantly, as conventional wisdom holds that HDR is limited to proliferative cells and DSBs introduced into post-mitotic cells (e.g., muscle fibers) will be repaired instead by

NHEJ, it is likely that achieving precise repair of the *Dmd* gene will require efficient co-delivery into muscle satellite cells of CRISPR/Cas9, *Dmd*-targeting guide RNAs, and donor DNA template to direct HDR. Such a feat will, in turn, likely necessitate the identification of novel or optimized delivery vehicles that exhibit high satellite cell tropism (see above). While certainly challenging, success in such an approach would represent a very promising treatment strategy for DMD.

Gene-Editing Therapy in Combination with AONs or Microdystrophin

AAV-mediated delivery of expression constructs encoding for AONs has been shown to enable widespread exon skipping, restoration of Dystrophin protein production and improvement of muscle function in short-term animal studies (Goyenvall et al. 2004, 2012; Denti et al. 2006; Le Guiner et al. 2014); however, long-term studies in a more severe mouse model (Le Hir et al. 2013) and also in the Golden retriever model of DMD (Vulin et al. 2012) revealed that vector genomes are lost from dystrophic muscle upon muscle damage and also over time. This observation can be explained by injury-induced loss or degeneration of muscle fibers that previously were transduced by the AAV vector and subsequent incorporation of new satellite cell-derived nuclei to the muscle. Importantly, as noted above, the low rate of satellite cell transduction with the AAV serotypes tested thus far, together with the likelihood that these cells and their progeny proliferate prior to incorporation into muscle fibers, makes it doubtful that additional vector genomes are delivered to muscle via fusion of satellite cell progeny in this system. Consistent with this, acute muscle damage by cardiotoxin injury of AAV injected mouse dystrophic muscle results in rapid loss of vector genome from the muscle (Le Hir et al. 2013). Irreversible gene correction of regenerating satellite cells and their progeny, achieved by gene editing, has the potential to overcome this challenge. Moreover, to avoid immune response complications related to re-administration of AAV, non-viral delivery of *DMD*-CRISPR to dystrophic muscle could potentially be used to complement viral delivery of AONs or microdystrophin to achieve long-term and persistent Dystrophin restoration.

Possible Application of CRISPR-mediated gene editing Strategies in Other Diseases

The reprogrammable targeting of the Cas9 endonuclease via easily constructed gRNAs presents the exciting possibility of utilizing this system to treat a wide range of genetic diseases. Results from *Dmd* targeting by AAV-CRISPR in *mdx* mice are most immediately pertinent to other muscle disorders that are likewise

amenable to mRNA splicing modulation, i.e., exon skipping or exon retention strategies, conventionally achieved by AONs. These disorders include primary dysferlinopathies, such as limb-girdle muscular dystrophy type 2B, resulting from mutations in the large dysferlin protein coding region that may be skipped inconsequentially if contained in redundant C2 domains (Wein et al. 2010). AONs also have been used in spinal muscular atrophy (SMA) to interrupt the function of an intronic splicing silencer that would otherwise result in the omission of exon 7 of the survival motor neuron 2 (SMN2) protein product, thereby allowing compensation for the loss-of-function of its paralog, survival motor neuron 1 (SMN1) in SMA patients (Burghes and McGovern 2010). Furthermore, the use of AAV-CRISPR as an AON alternative is suitable for non-muscle specific diseases like Leber congenital amaurosis (Maeder and Gersbach 2016).

Aside from complementing and potentially superseding the use of AONs in exon exclusion strategies, NHEJ-mediated DNA excision is applicable more generally for the targeted removal of specific genomic elements associated with disease. For example, chemokine receptor 5 (CCR5) is a critical human immunodeficiency virus type 1 (HIV-1) co-receptor that is necessary for the fusion to and infection of cells by CCR5-tropic virions (Broder and Collman 1997). Mutations in the CCR5 gene can confer immunity to HIV-1 infection, and transplantation of hematopoietic stem cells carrying the same mutated gene has been aggressively pursued as a possible curative treatment (Allers et al. 2011). By using Cas9 and paired gRNAs, researchers have been able recently to selectively mutate the CCR5 gene and thereby provide resistance of immune cells to HIV-1 infection (Kang et al. 2015; Mandal et al. 2014). Moreover, CRISPR-Cas9 can be used to directly target and disrupt integrated proviral genomes (Vulin et al. 2012; Ebina et al. 2013; Kennedy and Cullen 2015; Wang et al. 2015). Other uses may include the removal of excess nucleotides in trinucleotide repeat disorders (Park et al. 2015) and the knock-out of proprotein convertase subtilisin/kexin type 9 (PCSK9) involved in hypercholesterolemia (Ding et al. 2014; Ran et al. 2015; Wang et al. 2016). Finally, approaches utilizing co-delivery of CRISPR components with a donor DNA template to correct mutations via activation of the HDR pathway are also currently under development to treat cystic fibrosis (Schwank et al. 2013), hemophilia A (Park et al. 2015), hereditary tyrosinemia (Yin et al. 2014), sickle cell disease (Orkin 2016), severe combined immunodeficiency (Booth et al. 2016), and other, predominantly loss-of-function genetic diseases.

Conclusions and Perspective

Three independent studies have provided evidence for AAV-mediated delivery of CRISPR components targeting *Dmd* and restoring Dystrophin expression in dystrophic cardiac and skeletal muscle (Long et al. 2015; Nelson et al. 2015; Tabebordbar et al. 2015). One study (Tabebordbar et al. 2015) also showed *Dmd*

gene targeting in dystrophic muscle stem cells. Correction of *Dmd* in dystrophic satellite cells provides a critical reservoir of myogenic progenitors capable of producing Dystrophin-expressing muscle fibers and represents a potential advantage compared to conventional transgene-mediated gene therapy. Transgenes delivered by AAV are generally maintained as non-replicating episomes and thus are diluted during expansion of satellite cells and their myoblast progeny. In contrast, CRISPR-mediated gene editing allows for irreversible modification of *Dmd* in satellite cells and their progeny, a result that is even more advantageous if the gene-corrected cells are selected for, or enriched, in dystrophic tissue. Expansion of clusters of naturally occurring Dystrophin-expressing revertant fibers in *mdx* muscle, which depends on muscle regeneration, suggests that such a selective advantage may exist for Dystrophin-expressing satellite cells in dystrophic muscle (Yokota et al. 2006). It would be interesting to test if gene-corrected satellite cells are selectively enriched in dystrophic muscles after induced muscle degeneration and regeneration. Furthermore, it would be informative to examine whether permanent gene correction of dystrophic satellite cells (and their progeny) prevents the loss of Dystrophin-expressing nuclei in muscle fibers, which is typically seen with traditional gene therapy approaches (Vulin et al. 2012; Le Hir et al. 2013). Minimizing off-target activity of Cas9 nuclease, analyzing potential immune responses against CRISPR components and therapeutic gene products and developing non-viral delivery approaches for transient expression of *DMD*-CRISPR in dystrophic muscle will also be important to help to move gene editing technology towards clinical application for DMD. In addition, it is important to keep in mind that the efficacy and safety of this approach in non-rodent dystrophy models is yet to be studied. Canine models of DMD, including the golden retriever muscular dystrophy (GRMD) model, exhibit more severe dystrophic phenotypes that show greater similarity to human DMD phenotypes than the *mdx* mouse model (Kornegay et al. 2012). Therefore, preclinical studies in dog models might better indicate the therapeutic potential of in vivo gene editing for DMD. The recently developed human muscle xenograft model also provides a unique and informative opportunity for studying the efficacy of *DMD*-CRISPR in correcting mutations in human dystrophic muscle fibers and satellite cells in vivo (Zhang et al. 2014). Finally, to assess the likelihood of vertical transfer of gene-editing events to the next generation after systemic gene editing, germline and also transplacental transmission of AAV-CRISPR should be rigorously analyzed. AAV9 has been shown to penetrate the placenta (Picconi et al. 2014) in mice, a finding that should be taken into consideration for planning clinical application of this technology. Still, the possibility to directly modify the human genome to correct deleterious mutations that lead to devastating human diseases, such as DMD, presents unprecedented promise for the future of regenerative medicine.

References

- Aartsma-Rus A, Fokkema I, Verschuuren J, Ginjaar I, van Deutekom J, van Ommen GJ, den Dunnen JT (2009) Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy mutations. *Hum Mutat* 30:293–299
- Allers K, Hutter G, Hofmann J, Loddenkemper C, Rieger K, Thiel E, Schneider T (2011) Evidence for the cure of HIV infection by CCR5Delta32/Delta32 stem cell transplantation. *Blood* 117:2791–2799
- Beroud C, Tuffery-Giraud S, Matsuo M, Hamroun D, Humbertclaude V, Monnier N, Moizard MP, Voelckel MA, Calemard LM, Boisseau P, Blayau M, Philippe C, Cossee M, Pages M, Rivier F, Danos O, Garcia L, Claustres M (2007) Multiexon skipping leading to an artificial DMD protein lacking amino acids from exons 45 through 55 could rescue up to 63% of patients with Duchenne muscular dystrophy. *Hum Mutat* 28:196–202
- Booth C, Gaspar HB, Thrasher AJ (2016) Treating immunodeficiency through HSC gene therapy. *Trends Mol Med* 22:317–327
- Bosnakovski D, Xu Z, Li W, Thet S, Cleaver O, Perlingeiro RC, Kyba M (2008) Prospective isolation of skeletal muscle stem cells with a Pax7 reporter. *Stem Cells* 26:3194–3204
- Boutin S, Monteilhet V, Veron P, Leborgne C, Benveniste O, Montus MF, Masurier C (2010) Prevalence of serum IgG and neutralizing factors against adeno-associated virus (AAV) types 1, 2, 5, 6, 8, and 9 in the healthy population: implications for gene therapy using AAV vectors. *Hum Gene Ther* 21:704–712
- Bowles DE, McPhee SW, Li C, Gray SJ, Samulski JJ, Camp AS, Li J, Wang B, Monahan PE, Rabinowitz JE, Grieger JC, Govindasamy L, Agbandje-McKenna M, Xiao X, Samulski RJ (2012) Phase 1 gene therapy for Duchenne muscular dystrophy using a translational optimized AAV vector. *Mol Therapy J Am Soc Gene Therapy* 20:443–455
- Broder CC, Collman RG (1997) Chemokine receptors and HIV. *J Leukoc Biol* 62:20–29
- Burghes AH, Logan C, Hu X, Belfall B, Worton RG, Ray PN (1987) A cDNA clone from the Duchenne/Becker muscular dystrophy gene. *Nature* 328:434–437
- Burghes AH, McGovern VL (2010) Antisense oligonucleotides and spinal muscular atrophy: skipping along. *Genes Dev* 24:1574–1579
- Burrow KL, Coovert DD, Klein CJ, Bulman DE, Kissel JT, Rammohan KW, Burghes AH, Mendell JR (1991) Dystrophin expression and somatic reversion in prednisone-treated and untreated Duchenne dystrophy. CIDD Study Group. *Neurology* 41:661–666
- Calvo JA, Meira LB, Lee CY, Moroski-Erkul CA, Abolhassani N, Taghizadeh K, Eichinger LW, Muthupalani S, Nordstrand LM, Klungland A, Samson LD (2012) DNA repair is indispensable for survival after acute inflammation. *J Clin Invest* 122:2680–2689
- Campbell KP, Kahl SD (1989) Association of dystrophin and an integral membrane glycoprotein. *Nature* 338:259–262
- Canver MC, Bauer DE, Dass A, Yien YY, Chung J, Masuda T, Maeda T, Paw BH, Orkin SH (2014) Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. *J Biol Chem* 289:21312–21324
- Cerletti M, Jurga S, Witczak CA, Hirshman MF, Shadrach JL, Goodyear LJ, Wagers AJ (2008) Highly efficient, functional engraftment of skeletal muscle stem cells in dystrophic muscles. *Cell* 134:37–47
- Chamberlain JS (2002) Gene therapy of muscular dystrophy. *Hum Mol Genet* 11:2355–2362
- Cho SW, Kim S, Kim JM, Kim JS (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* 31:230–232
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823
- Dalkara D, Byrne LC, Klimczak RR, Visel M, Yin L, Merigan WH, Flannery JG, Schaffer DV (2013) In vivo-directed evolution of a new adeno-associated virus for therapeutic outer retinal gene delivery from the vitreous. *Sci Transl Med* 5:189ra176

- Denti MA, Rosa A, D'Antona G, Sthandier O, De Angelis FG, Nicoletti C, Allocca M, Pansarasa O, Parente V, Musaro A, Auricchio A, Bottinelli R, Bozzoni I (2006) Body-wide gene therapy of Duchenne muscular dystrophy in the mdx mouse model. *Proc Natl Acad Sci USA* 103:3758–3763
- DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM (2013) Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res* 41:4336–4343
- Ding Q, Regan SN, Xia Y, Oostrom LA, Cowan CA, Musunuru K (2013) Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell* 12:393–394
- Ding Q, Strong A, Patel KM, Ng SL, Gosis BS, Regan SN, Rader DJ, Musunuru K (2014) Permanent alteration of PCSK9 with in vivo CRISPR-Cas9 genome editing. *Circ Res* 115:488–492
- Dumont NA, Wang YX, von Maltzahn J, Pasut A, Bentzinger CF, Brun CE, Rudnicki MA (2015) Dystrophin expression in muscle stem cells regulates their polarity and asymmetric division. *Nat Med* 21:1455–1463
- Ebina H, Misawa N, Kanemura Y, Koyanagi Y (2013) Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci Rep* 3:2510
- England SB, Nicholson LV, Johnson MA, Forrest SM, Love DR, Zubrzycka-Gaarn EE, Bulman DE, Harris JB, Davies KE (1990) Very mild muscular dystrophy associated with the deletion of 46% of dystrophin. *Nature* 343:180–182
- Ervasti JM, Campbell KP (1991) Membrane organization of the dystrophin-glycoprotein complex. *Cell* 66:1121–1131
- Fabb SA, Wells DJ, Serpente P, Dickson G (2002) Adeno-associated virus vector gene transfer and sarcolemmal expression of a 144 kDa micro-dystrophin effectively restores the dystrophin-associated protein complex and inhibits myofibre degeneration in nude/mdx mice. *Hum Mol Genet* 11:733–741
- Fanin M, Danieli GA, Cadaldini M, Miorin M, Vitiello L, Angelini C (1995) Dystrophin-positive fibers in Duchenne dystrophy: origin and correlation to clinical course. *Muscle Nerve* 18:1115–1120
- Friedland AE, Tzur YB, Esvelt KM, Colaiacovo MP, Church GM, Calarco JA (2013) Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat Methods* 10:741–743
- Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK (2014) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol* 32:279–284
- Gao G, Vandenberghe LH, Alvira MR, Lu Y, Calcedo R, Zhou X, Wilson JM (2004) Clades of Adeno-associated viruses are widely disseminated in human tissues. *J Virol* 78:6381–6388
- Gao X, Zhao J, Han G, Zhang Y, Dong X, Cao L, Wang Q, Moulton HM, Yin H (2014) Effective dystrophin restoration by a novel muscle-homing peptide-morpholino conjugate in dystrophin-deficient mdx mice. *Mol Ther* 22:1333–1341
- Goyenvallé A, Vulin A, Fougère F, Leturcq F, Kaplan JC, Garcia L, Danos O (2004) Rescue of dystrophic muscle through U7 snRNA-mediated exon skipping. *Science* 306:1796–1799
- Goyenvallé A, Babbs A, Wright J, Wilkins V, Powell D, Garcia L, Davies KE (2012) Rescue of severely affected dystrophin/utrophin-deficient mice through scAAV-U7snRNA-mediated exon skipping. *Hum Mol Genet* 21:2559–2571
- Goyenvallé A, Griffith G, Babbs A, El Andaloussi S, Ezzat K, Avril A, Dugovic B, Chaussonnet R, Ferry A, Voit T, Amthor H, Buhr C, Schurch S, Wood MJ, Davies KE, Vaillend C, Leumann C, Garcia L (2015) Functional correction in mouse models of muscular dystrophy using exon-skipping tricyclo-DNA oligomers. *Nat Med* 21:270–275
- Gregorevic P, Blankinship MJ, Allen JM, Crawford RW, Meuse L, Miller DG, Russell DW, Chamberlain JS (2004) Systemic delivery of genes to striated muscles using adeno-associated viral vectors. *Nat Med* 10:828–834
- Harper SQ, Hauser MA, DelloRusso C, Duan D, Crawford RW, Phelps SF, Harper HA, Robinson AS, Engelhardt JF, Brooks SV, Chamberlain JS (2002) Modular flexibility of dystrophin: implications for gene therapy of Duchenne muscular dystrophy. *Nat Med* 8:253–261

- Hoffman EP, Morgan JE, Watkins SC, Partridge TA (1990) Somatic reversion/suppression of the mouse mdx phenotype in vivo. *J Neurol Sci* 99:9–25
- Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JR, Joung JK (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* 31:227–229
- Jackson SP, Bartek J (2009) The DNA-damage response in human biology and disease. *Nature* 461:1071–1078
- Kang H, Minder P, Park MA, Mesquitta WT, Torbett BE, Slukvin II (2015) CCR5 disruption in induced pluripotent stem cells using CRISPR/Cas9 provides selective resistance of immune cells to CCR5-tropic HIV-1 virus. *Mol Therapy Nucleic acids* 4:e268
- Kennedy EM, Cullen BR (2015) Bacterial CRISPR/Cas DNA endonucleases: a revolutionary technology that could dramatically impact viral research and treatment. *Virology* 479–480:213–220
- Kim S, Kim D, Cho SW, Kim J, Kim JS (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res* 24:1012–1019
- Klein CJ, Coover DD, Bulman DE, Ray PN, Mendell JR, Burghes AH (1992) Somatic reversion/suppression in Duchenne muscular dystrophy (DMD): evidence supporting a frame-restoring mechanism in rare dystrophin-positive fibers. *Am J Hum Genet* 50:950–959
- Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z, Joung JK (2016) High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529:490–495
- Kluytmans J, van Belkum A, Verbrugh H (1997) Nasal carriage of *Staphylococcus aureus*: epidemiology, underlying mechanisms, and associated risks. *Clin Microbiol Rev* 10:505–520
- Koenig M, Hoffman EP, Bertelson CJ, Monaco AP, Feener C, Kunkel LM (1987) Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell* 50:509–517
- Kornegay JN, Bogan JR, Bogan DJ, Childers MK, Li J, Nghiem P, Detwiler DA, Larsen CA, Grange RW, Bhavaraju-Sanka RK, Tou S, Keene BP, Howard JF Jr, Wang J, Fan Z, Schatzberg SJ, Styner MA, Flanigan KM, Xiao X, Hoffman EP (2012) Canine models of Duchenne muscular dystrophy and their use in therapeutic strategies. *Mamm Genome* 23:85–108
- Kotterman MA, Schaffer DV (2014) Engineering adeno-associated viruses for clinical gene therapy. *Nat Rev Genet* 15:445–451
- Le Guiner C, Montus M, Servais L, Cherel Y, Francois V, Thibaud JL, Wary C, Matot B, Larcher T, Guigand L, Dutilleul M, Domenger C, Allais M, Beuvin M, Moraux A, Le Duff J, Devaux M, Jaulin N, Guilbaud M, Latournerie V, Veron P, Boutin S, Leborgne C, Desgue D, Deschamps JY, Moullec S, Fromes Y, Vulin A, Smith RH, Laroudie N, Barnay-Toutain F, Rivière C, Bucher S, Le TH, Delaunay N, Gasmi M, Kotin RM, Bonne G, Adjali O, Masurier C, Hogrel JY, Carlier P, Moullier P, Voit T (2014) Forelimb treatment in a large cohort of dystrophic dogs supports delivery of a recombinant AAV for exon skipping in Duchenne patients. *Mol Ther* 22:1923–1935
- Le Hir M, Goyenvallé A, Peccate C, Precigout G, Davies KE, Voit T, Garcia L, Lorain S (2013) AAV genome loss from dystrophic mouse muscles during AAV-U7 snRNA-mediated exon-skipping therapy. *Mol Ther* 21:1551–1558
- Lisowski L, Dane AP, Chu K, Zhang Y, Cunningham SC, Wilson EM, Nygaard S, Grompe M, Alexander IE, Kay MA (2014) Selection and evaluation of clinically relevant AAV variants in a xenograft liver model. *Nature* 506:382–386
- Long C, McAnally JR, Shelton JM, Mireault AA, Bassel-Duby R, Olson EN (2014) Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. *Science* 345:1184–1188
- Long C, Amoasii L, Mireault AA, McAnally JR, Li H, Sanchez-Ortiz E, Bhattacharyya S, Shelton JM, Bassel-Duby R, Olson EN (2015) Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* 351:400–403
- Lu QL, Morris GE, Wilton SD, Ly T, Artem'yeva OV, Strong P, Partridge TA (2000) Massive idiosyncratic exon skipping corrects the nonsense mutation in dystrophic mouse muscle and produces functional revertant fibers by clonal expansion. *J Cell Biol* 148:985–996

- Lu QL, Cirak S, Partridge T (2014) What can we learn from clinical trials of exon skipping for DMD? *Mol Ther Nucl Acids* 3:e152
- Maeder ML, Gersbach CA (2016) Genome-editing technologies for gene and cell therapy. *Mol Therapy J Am Soc Gene Therapy* 24:430–446
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013) RNA-guided human genome engineering via Cas9. *Science* 339:823–826
- Mandal PK, Ferreira LM, Collins R, Meissner TB, Boutwell CL, Friesen M, Vrbanac V, Garrison BS, Stortchevoi A, Bryder D, Musunuru K, Brand H, Tager AM, Allen TM, Talkowski ME, Rossi DJ, Cowan CA (2014) Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. *Cell Stem Cell* 15:643–652
- Manno CS, Pierce GF, Arruda VR, Glader B, Ragni M, Rasko JJ, Ozelo MC, Hoots K, Blatt P, Konkle B, Dake M, Kaye R, Razavi M, Zajko A, Zehnder J, Rustagi PK, Nakai H, Chew A, Leonard D, Wright JF, Lessard RR, Sommer JM, Tigges M, Sabatino D, Luk A, Jiang H, Mingozzi F, Couto L, Ertl HC, High KA, Kay MA (2006) Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nat Med* 12:342–347
- Mendell JR, Campbell K, Rodino-Klapac L, Sahenk Z, Shilling C, Lewis S, Bowles D, Gray S, Li C, Galloway G, Malik V, Coley B, Clark KR, Li J, Xiao X, Samulski J, McPhee SW, Samulski RJ, Walker CM (2010) Dystrophin immunity in Duchenne's muscular dystrophy. *N Engl J Med* 363:1429–1437
- Mingozzi F, High KA (2011) Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nat Rev Genet* 12:341–355
- Muir LA, Chamberlain JS (2009) Emerging strategies for cell and gene therapy of the muscular dystrophies. *Expert Rev Mol Med* 11:e18
- Nakamura A, Yoshida K, Fukushima K, Ueda H, Urasawa N, Koyama J, Yazaki Y, Yazaki M, Sakai T, Haruta S, Takeda S, Ikeda S (2008) Follow-up of three patients with a large in-frame deletion of exons 45–55 in the Duchenne muscular dystrophy (DMD) gene. *J Clin Neurosci* 15:757–763
- Nathwani AC, Tuddenham EG, Rangarajan S, Rosales C, McIntosh J, Linch DC, Chowdary P, Riddell A, Pie AJ, Harrington C, O'Beirne J, Smith K, Pasi J, Glader B, Rustagi P, Ng CY, Kay MA, Zhou J, Spence Y, Morton CL, Allay J, Coleman J, Sleep S, Cunningham JM, Srivastava D, Basner-Tschakarjan E, Mingozzi F, High KA, Gray JT, Reiss UM, Nienhuis AW, Davidoff AM (2011) Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *N Engl J Med* 365:2357–2365
- Nelson CE, Hakim CH, Ousterout DG, Thakore PI, Moreb EA, Rivera RM, Madhavan S, Pan X, Ran FA, Yan WX, Asokan A, Zhang F, Duan D, Gersbach CA (2015) In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* 351:403–407
- Neri M, Torelli S, Brown S, Ugo I, Sabatelli P, Merlini L, Spitali P, Rimessi P, Gualandi F, Sewry C, Ferlini A, Muntoni F (2007) Dystrophin levels as low as 30% are sufficient to avoid muscular dystrophy in the human. *Neuromuscul Disord* 17:913–918
- Nicholson LV, Johnson MA, Bushby KM, Gardner-Medwin D (1993) Functional significance of dystrophin positive fibres in Duchenne muscular dystrophy. *Arch Dis Child* 68:632–636
- Orkin SH (2016) Recent advances in globin research using genome-wide association studies and gene editing. *Ann N Y Acad Sci* 1368:5–10
- Park CY, Kim DH, Son JS, Sung JJ, Lee J, Bae S, Kim JH, Kim DW, Kim JS (2015) Functional correction of large factor VIII gene chromosomal inversions in hemophilia a patient-derived iPSCs using CRISPR-Cas9. *Cell Stem Cell* 17:213–220
- Partridge TA (2011) Impending therapies for Duchenne muscular dystrophy. *Curr Opin Neurol* 24:415–422
- Picconi JL, Muff-Luett MA, Wu D, Bunchman E, Schaefer F, Brophy PD (2014) Kidney-specific expression of GFP by in-utero delivery of pseudotyped adeno-associated virus 9. *Mol Ther Methods Clin Dev* 1:14014

- Ramakrishna S, Kwaku Dad AB, Beloor J, Gopalappa R, Lee SK, Kim H (2014) Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Res* 24:1020–1027
- Ramos J, Chamberlain JS (2015) Gene therapy for Duchenne muscular dystrophy. *Expert Opin Orphan Drugs* 3:1255–1266
- Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F (2013a) Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 8:2281–2308
- Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F (2013b) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154:1380–1389
- Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS, Koonin EV, Sharp PA, Zhang F (2015) In vivo genome editing using Staphylococcus aureus Cas9. *Nature* 520:186–191
- Rodino-Klapac LR, Montgomery CL, Bremer WG, Shontz KM, Malik V, Davis N, Sprinkle S, Campbell KJ, Sahenk Z, Clark KR, Walker CM, Mendell JR, Chicoine LG (2010) Persistent expression of FLAG-tagged micro dystrophin in nonhuman primates following intramuscular and vascular delivery. *Mol Ther* 18:109–117
- Sacco A, Mourikioti F, Tran R, Choi J, Llewellyn M, Kraft P, Shkreli M, Delp S, Pomerantz JH, Artandi SE, Blau HM (2010) Short telomeres and stem cell exhaustion model Duchenne muscular dystrophy in mdx/mTR mice. *Cell* 143:1059–1071
- Schwank G, Koo BK, Sasselli V, Dekkers JF, Heo I, Demircan T, Sasaki N, Boymans S, Cuppen E, van der Ent CK, Nieuwenhuis EE, Beekman JM, Clevers H (2013) Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* 13:653–658
- Shin JH, Pan X, Hakim CH, Yang HT, Yue Y, Zhang K, Terjung RL, Duan D (2013) Microdystrophin ameliorates muscular dystrophy in the canine model of duchenne muscular dystrophy. *Mol Therapy J Am Soc Gene Therapy* 21:750–757
- Sicinski P, Geng Y, Ryder-Cook AS, Barnard EA, Darlison MG, Barnard PJ (1989) The molecular basis of muscular dystrophy in the mdx mouse: a point mutation. *Science* 244:1578–1580
- Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F (2016) Rationally engineered Cas9 nucleases with improved specificity. *Science* 351:84–88
- Song L, Li X, Jayandharan GR, Wang Y, Aslanidi GV, Ling C, Zhong L, Gao G, Yoder MC, Ling C, Tan M, Srivastava A (2013a) High-efficiency transduction of primary human hematopoietic stem cells and erythroid lineage-restricted expression by optimized AAV6 serotype vectors in vitro and in a murine xenograft model in vivo. *PLoS One* 8:e58757
- Song L, Kauss MA, Kopin E, Chandra M, Ul-Hasan T, Miller E, Jayandharan GR, Rivers AE, Aslanidi GV, Ling C, Li B, Ma W, Li X, Andino LM, Zhong L, Tarantal AF, Yoder MC, Wong KK Jr, Tan M, Chatterjee S, Srivastava A (2013b) Optimizing the transduction efficiency of capsid-modified AAV6 serotype vectors in primary human hematopoietic stem cells in vitro and in a xenograft mouse model in vivo. *Cytherapy* 15:986–998
- Straub V, Bittner RE, Leger JJ, Voit T (1992) Direct visualization of the dystrophin network on skeletal muscle fiber membrane. *J Cell Biol* 119:1183–1191
- Symington LS, Gautier J (2011) Double-strand break end resection and repair pathway choice. *Annu Rev Genet* 45:247–271
- Tabebordbar M, Wang ET, Wagers AJ (2013) Skeletal muscle degenerative diseases and strategies for therapeutic muscle repair. *Annu Rev Pathol* 8:441–475
- Tabebordbar M, Zhu K, Cheng JK, Chew WL, Widrick JJ, Yan WX, Maesner C, EY W, Xiao R, Ran FA, Cong L, Zhang F, Vandenbergh LH, Church GM, Wagers AJ (2015) In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* 351:407–411
- Taglia A, Petillo R, D'Ambrosio P, Picillo E, Torella A, Orsini C, Ergoli M, Scutifero M, Passamano L, Palladino A, Nigro G, Politano L (2015) Clinical features of patients with dystrophinopathy sharing the 45-55 exon deletion of DMD gene. *Acta Myol* 34:9–13
- Uchino M, Tokunaga M, Mita S, Uyama E, Ando Y, Teramoto H, Miike T, Ando M (1995) PCR and immunocytochemical analyses of dystrophin-positive fibers in Duchenne muscular dystrophy. *J Neurol Sci* 129:44–50

- Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* 435:646–651
- van Putten M, Hulsker M, Nadarajah VD, van Heiningen SH, van Huizen E, van Iterson M, Admiraal P, Messemaker T, den Dunnen JT, t Hoen PA, Aartsma-Rus A (2012) The effects of low levels of dystrophin on mouse muscle function and pathology. *PLoS One* 7:e31937
- van Putten M, Hulsker M, Young C, Nadarajah VD, Heemskerk H, van der Weerd L, t Hoen PA, van Ommen GJ, Aartsma-Rus AM (2013) Low dystrophin levels increase survival and improve muscle pathology and function in dystrophin/utrophin double-knockout mice. *FASEB J* 27:2484–2495
- van Putten M, van der Pijl EM, Hulsker M, Verhaart IE, Nadarajah VD, van der Weerd L, Aartsma-Rus A (2014) Low dystrophin levels in heart can delay heart failure in mdx mice. *J Mol Cell Cardiol* 69:17–23
- Vila MC, Klimek MB, Novak JS, Rayavarapu S, Uaesoontrachoon K, Boehler JF, Fiorillo AA, Hogarth MW, Zhang A, Shaughnessy C, Gordish-Dressman H, Burki U, Straub V, Lu QL, Partridge TA, Brown KJ, Hathout Y, van den Anker J, Hoffman EP, Nagaraju K (2015) Elusive sources of variability of dystrophin rescue by exon skipping. *Skelet Muscle* 5:1–12
- Voit RA, Hendel A, Pruett-Miller SM, Porteus MH (2014) Nuclease-mediated gene editing by homologous recombination of the human globin locus. *Nucleic Acids Res* 42:1365–1378
- Vulin A, Barthelemy I, Goyenvallé A, Thibaud JL, Beley C, Griffith G, Benchouir R, le Hir M, Unterfinger Y, Lorain S, Dreyfus P, Voit T, Carlier P, Blot S, Garcia L (2012) Muscle function recovery in golden retriever muscular dystrophy after AAV1-U7 exon skipping. *Mol Ther* 20:2120–2133
- Wallace GQ, McNally EM (2009) Mechanisms of muscle degeneration, regeneration, and repair in the muscular dystrophies. *Annu Rev Physiol* 71:37–57
- Wang B, Li J, Xiao X (2000) Adeno-associated virus vector carrying human minidystrophin genes effectively ameliorates muscular dystrophy in mdx mouse model. *Proc Natl Acad Sci USA* 97:13714–13719
- Wang D, Mou H, Li S, Li Y, Hough S, Tran K, Li J, Yin H, Anderson DG, Sontheimer EJ, Weng Z, Gao G, Xue W (2015) Adenovirus-mediated somatic genome editing of Pten by CRISPR/Cas9 in mouse liver in spite of Cas9-specific immune responses. *Hum Gene Ther* 26:432–442
- Wang X, Raghavan A, Chen T, Qiao L, Zhang Y, Ding Q, Musunuru K (2016) CRISPR-Cas9 targeting of PCSK9 in human hepatocytes in vivo. *Arterioscler Thromb Vasc Biol* 36:783–786
- Wein N, Avril A, Bartoli M, Beley C, Chaouch S, Laforet P, Behin A, Butler-Browne G, Mouly V, Krahn M, Garcia L, Levy N (2010) Efficient bypass of mutations in dysferlin deficient patient cells by antisenseinduced exon skipping. *Hum Mutat* 31:136–142
- Woo JW, Kim J, Kwon SI, Corvalan C, Cho SW, Kim H, Kim SG, Kim ST, Choe S, Kim JS (2015) DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat Biotechnol* 33:1162–1164
- Wu J, Hunt SD, Xue H, Liu Y, Darabi R (2016) Generation and characterization of a MYF5 reporter human iPS cell line using CRISPR/Cas9 mediated homologous recombination. *Sci Rep* 6:18759
- Xue W, Chen S, Yin H, Tammela T, Papagiannakopoulos T, Joshi NS, Cai W, Yang G, Bronson R, Crowley DG, Zhang F, Anderson DG, Sharp PA, Jacks T (2014) CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature* 514:380–384
- Yang Y, Wang L, Bell P, McMenamin D, He Z, White J, Yu H, Xu C, Morizono H, Musunuru K, Batshaw ML, Wilson JM (2016) A dual AAV system enables the Cas9-mediated correction of a metabolic liver disease in newborn mice. *Nat Biotechnol* 34:334–338
- Yin H, Xue W, Chen S, Bogorad RL, Benedetti E, Grompe M, Koteliensky V, Sharp PA, Jacks T, Anderson DG (2014) Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat Biotechnol* 32:551–553
- Yin H, Song CQ, Dorkin JR, Zhu LJ, Y L, Wu Q, Park A, Yang J, Suresh S, Bizhanova A, Gupta A, Bolukbasi MF, Walsh S, Bogorad RL, Gao G, Weng Z, Dong Y, Koteliensky V, Wolfe SA, Langer R, Xue W, Anderson DG (2016) Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. *Nat Biotechnol* 34:328–333

- Yokota T, QL L, Morgan JE, Davies KE, Fisher R, Takeda S, Partridge TA (2006) Expansion of revertant fibers in dystrophic mdx muscles reflects activity of muscle precursor cells and serves as an index of muscle regeneration. *J Cell Sci* 119:2679–2687
- Yokota T, Pistilli E, Duddy W, Nagaraju K (2007) Potential of oligonucleotide-mediated exon-skipping therapy for Duchenne muscular dystrophy. *Expert Opin Biol Ther* 7:831–842
- Zhang Y, King OD, Rahimov F, Jones TI, Ward CW, Kerr JP, Liu N, Emerson CP Jr, Kunkel LM, Partridge TA, Wagner KR (2014) Human skeletal muscle xenograft as a new preclinical model for muscle disorders. *Hum Mol Genet* 23:3180–3188
- Zincarelli C, Soltys S, Rengo G, Rabinowitz JE (2008) Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection. *Mol Ther* 16:1073–1080
- Zuris JA, Thompson DB, Shu Y, Guilinger JP, Bessen JL, JH H, Maeder ML, Joung JK, Chen ZY, Liu DR (2015) Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat Biotechnol* 33:73–80

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

