An Exploration of the Properties of Repair Template DNA that Promote Precision Genome Editing

An Exploration of the Properties of Repair Template DNA that Promote Precision Genome Editing

A Dissertation Presented

By

Krishna Sumanth Ghanta

Submitted to the Faculty of the

University of Massachusetts Graduate School of Biomedical Sciences, Worcester

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

(AUGUST 03, 2021)

Interdisciplinary Graduate Program

An Exploration of the Properties of Repair Template DNA that Promote Precision Genome Editing

A Dissertation Presented

By

Krishna Sumanth Ghanta

This work was undertaken in the Graduate School of Biomedical Sciences

(Interdisciplinary Graduate Program)

Under the mentorship of

Craig C. Mello, Ph.D., Thesis Advisor

Allan Jacobson, Ph.D., Member of Committee

Sean Ryder, Ph.D., Member of Committee

Jonathan K. Watts, Ph.D., Member of Committee

Dustin Updike, Ph.D., External Member of Committee

Eric Baehrecke, Ph.D., Chair of Committee

Mary Ellen Lane, Ph.D.,

Dean of the Graduate School of Biomedical Sciences

August 03, 2021

Acknowledgements

The time I have spent during my Ph.D. studies has been incredible and memorable. This dissertation would not be possible without the support of the people in my personal and professional lives. These incredible people have helped shape my graduate research work with their mentorship, advice, and support.

I am profoundly grateful to my mentor Dr. Craig Mello for his support and guidance. When I first met Craig to discuss the opportunity to rotate in his lab, he was analyzing the phenotypes of *C. elegans* dead embryos. My initial thought was that he may be helping a student for a few minutes, but I quickly realized that he was working on his own project and did the bench work himself. Craig's excitement for discoveries and his passion to solve scientific mysteries is contagious and inspiring. He motivated me to attack the problems from all angles and to be persistent. Craig has given me complete freedom to pursue any ideas that excite me and has been there guiding and supporting me. I am thankful for everything he has given me and grateful for the opportunity to pursue my graduate studies under his mentorship.

I would like to thank Dr. Jonathan Watts and Dr. Erik Sontheimer for their support during my thesis studies. Jon has always been incredibly supportive to me and generous with his time. He always listened to my ideas, good and bad, and has provided me the resources to test those ideas. I am thankful to Erik for

all the guidance he has provided to me. I could walk into his office anytime to discuss my project goals or seek career advice.

I am thankful to Dr. Eric Baehrecke who has served as the chair of my thesis committee. Eric has always been there whenever I needed his guidance. He connected me with other scientists who could help me with my projects. He allowed me to test my genome editing protocols in drosophila in his lab. Eric has provided me guidance for my career related queries that helped me find direction.

I would like to thank Dr. Victor Ambros for advice and guidance throughout my Ph.D. Victor's insightful questions always forced me to think critically. I am thankful to Dr. Sean Ryder and Dr. Allen Jacobson for their guidance and encouragement during the committee meetings. I would also like to thank Dr. Dustin Updike for accepting my request to be on the dissertation committee and for his time.

I would like to thank my peers and colleagues who collaborated with me and guided me over the years. Masaki Shirayama and Takao Ishidate have been great mentors in the Mello lab whose guidance helped me in my projects. I would like to thank Greg Dokshin, Aamir Mir and Zexiang Chen with whom I have worked closely while developing CRISPR protocols and this work would not be possible without their collaborations as well as Gina Caldas and Meetu Seth for their guidance and mentorship during the initial years of my graduate studies. I would like to thank Heesun Kim, Yuehe Ding, Wen Tang for their guidance on

experiments. I am thankful to have great lab members who made the lab a fun and enjoyable place: Ahmet Ozturk, Dan Durning, Johan Girgenrath, Katya Makeyeva and Yuchen Yang, for their collaborations and discussions. I am very thankful to Nikki Grace and Deb Bordne for managing the lab function smoothly and helping with all the administrative work.

I would like to thank Dr. Jaime Rivera, Yeonsoo Yoon, Judith Gallant and Ping Xu for their guidance and collaboration on mouse genome editing experiments.

I am indebted to my family without whom any of this work would not be possible. My parents have always encouraged me to pursue my interests and be persistent to achieve my goals. They are the key figures in shaping my life and I could reach this stage in life because of them. My sister has always been there whenever I needed her. My grandparents were instrumental in shaping my childhood. My loving wife, Mona, has made me a better person and a better scientist. Her support allows me to pursue opportunities without the fear of failure. We will continue to be there for each other and achieve our aspirations!

Table of Contents

ACKNOWLEDGEMENTS III

TABLE OF CONTENTS VI

ABSTRACT IX

LIST OF TABLES XI

LIST OF FIGURES XII

ABBREVIATIONS XIV

LIST OF PUBLICATIONS XV

CHAPTER I: INTRODUCTION 17

INTRODUCTION TO GENOME EDITING 17

1.1 ZINC FINGER NUCLEASES AND TALENS 18

1.2 PROKARYOTIC ANTI-PHAGE DEFENSE SYSTEMS- CRISPR 20

1.3 CRISPR/ CAS9 TECHNOLOGY 22

1.4 MECHANISMS OF DOUBLE STRAND BREAK REPAIR 24

1.5 FACTORS AFFECTING EFFICIENCY OF GENOME EDITING 31

1.6 FACTORS INFLUENCING HDR AND METHODS TO IMPROVE EFFICIENCIES 40

1.7 CRISPR IN CLINIC 48

1.8 GENOME EDITING IN *C. ELEGANS* 51

PREFACE TO CHAPTER II 54

CHAPTER II: MICROINJECTION FOR PRECISION GENOME EDITING IN *CAENORHABDITIS ELEGANS*. 55

ABSTRACT 56

CULTURING WORMS 60

PREPARATION OF INJECTION PADS 61

PREPARATION OF NEEDLES 62

PREPARATION OF STOCK SOLUTIONS 67

GUIDE RNA DESIGN 68

TEMPLATED PRECISE EDITING- DONOR DESIGN 69

STEP-BY-STEP METHOD DETAILS 77

LIMITATIONS 93

ACKNOWLEDGMENTS 97

AUTHOR CONTRIBUTIONS 98

DECLARATION OF INTERESTS 98

PREFACE TO CHAPTER III 99

CHAPTER III: ROBUST GENOME EDITING WITH SHORT SINGLE-STRANDED AND LONG, PARTIALLY SINGLE-STRANDED DNA DONORS IN *CAENORHABDITIS ELEGANS* 100

ABSTRACT 101

INTRODUCTION 102

MATERIAL AND METHODS 103

RESULTS 110

DISCUSSION 125

PREFACE TO CHAPTER IV 134

CHAPTER IV: MELTING THE DSDNA DONORS POTENTIATES PRECISION GENOME EDITING IN *CAENORHABDITIS ELEGANS* 135

ABSTRACT 136

INTRODUCTION 137

MATERIAL AND METHODS 139

RESULTS 149

DISCUSSION 172

PREFACE TO CHAPTER V 179

CHAPTER V: 5' MODIFICATIONS IMPROVE POTENCY AND EFFICACY OF DNA DONORS FOR PRECISION GENOME EDITING 180

ABSTRACT 181

INTRODUCTION 182

MATERIAL AND METHODS 184

RESULTS 197

DISCUSSION 230

CHAPTER VI: DISCUSSION 236

FUTURE DIRECTIONS FOR GENOME EDITING IN *C. ELEGANS* 243

FUTURE DIRECTIONS FOR GENOME EDITING IN VERTEBRATE SYSTEMS 245

CONCLUSIONS 247

REFERENCES 248

Abstract

CRISPR/Cas9 induced DNA breaks can be precisely repaired by cellular homology-directed repair (HDR) pathways using exogenously provided template DNA (donor). However, the full potential of precision editing is hindered in many model systems by low cutting efficiencies, low HDR efficiencies and, cytotoxicity related to Cas9 and donor DNA. In this thesis, I address these challenges and present methods that we developed to increase HDR efficiencies in multiple model organisms.

In *Caenorhabditis elegans,* we show that by reducing toxicity high editing efficiencies can be achieved with single stranded oligonucleotide (ssODN) donors. We demonstrate that melting dsDNA donors dramatically improves the knock-in efficiencies of longer (1kb) edits. In addition, we describe 5′-terminal modifications to the donor molecules that further increase the frequency of precision editing. With our methodology a single optimally injected animal can yield more than 100 Green Fluorescent Protein (GFP) positive progeny, dramatically enhancing efficiency of genome editing.

Next, we demonstrate the generality of 5′ modified donors by extending our studies to human cell cultures and mice zygotes. In mammalian models, 2′OMe-RNA modifications consistently increase HDR efficiencies by several fold over unmodified donors. Furthermore, end-modified donors exhibited a striking reduction in end-joining reactions including reduced concatemer formation and

reduced direct ligation into the host genome. Our study demonstrates that HDR can be improved without inhibiting competing end-joining pathways and provides a platform to identify new chemical modifications that could further increase the potency and efficacy of precision genome editing.

List of Tables

Table 2.1 Key Resources Table 58

Table 3.1 List of *C. elegans* Strains 103

Table 3.2 Efficiencies of FLAG tag insertions at argonaute loci using ssODNs 130

Table 3.3 HDR efficiencies of GFP insertion with blunt ended PCRs as donors 131

Table 3.4 Comparison of HDR Efficiencies with donors of different homology arms 132

Table 4.1 List of *C. elegans* strains 140

Table 4.2 Sequences of crRNAs 140

Table 4.3 Sequences of oligos 141

Table 5.1 Sequences of guide RNA spacers 194

Table 5.2 Sequences of oligos 195

List of Figures

Figure 1.1 Double strand break repair pathways..... 27

Figure 1.2 Strategies to Improve HDR efficiencies..... 50

Figure 2.1 Graphical abstract of CRISPR/Cas genome editing protocol 57

Figure 2.2 Donor design strategy..... 70

Figure 2.3 Mounting worms..... 86

Figure 2.4 Microinjection 91

Figure 3.1 Determining optimal Cas9 RNP concentrations..... 114

Figure 3.2. Efficient integration of 3XFLAG ssODN donor at fourteen of the *C. elegans* argonaute genes using pRF4::rol-6(su1006) co-injection marker 122

Figure 3.3 Alternative targeting strategy for *sago-2* 123

Figure 3.4 Efficient editing with long, partially single-stranded dsDNA donors..... 124

Figure 4.1. Melting dsDNA donors potentiates HDR in *C. elegans*..... 151

Figure 4.2: Melted dsDNA donors promote homology directed repair 152

Figure 4.3. High doses of donor DNA reduce the number of Rollers..... 154

Figure 4.4. HDR efficiencies are improved with melted donors 155

Figure 4.5. Melted donors increase HDR efficiencies 158

Figure 4.6. Precise genome editing using Cas12a and melted dsDNA donors..... 160

Figure 4.7. Editing occurs later in the brood after roller cohort 164

Figure 4.8. Purity of donor DNA is crucial for best HDR efficacy 168

Figure 4.9. Short PCR contaminants outcompete dsDNA donors and integrate into the target locus efficiently 171

Figure 4.10 Quickly cooled donors act as better repair templates than slowly cooled donors 176

Figure 5.1 5′ end-modified donors promote HDR in HEK293T/TLR cells 199
Figure 5.2 2′OMe-RNA at 5′ ends of donors promote HDR in mammalian cells 200
Figure 5.3 End-modified donors promote HDR at endogenous loci in mammalian cell cultures 203
Figure 5.4 RNA::TEG donors with short (90bp) homology arms are more potent than unmodified donors at EMX1 locus 205
Figure 5.5 End-modifications increase potency of ssODN donors 208
Figure 5.6 Effects of terminal and non-terminal modifications of ssODN donors on HDR efficacy 209
Figure 5.7 Modified donors promote precise editing in *C. elegans* 213
Figure 5.8 Modified donors promote precise editing at the *csr-1* locus in *C. elegans* 215
Figure 5.9 2′OMe-RNA-TEG donors promote precise editing at *Tyr* locus in mouse zygotes 218
Figure 5.10 Microinjection information for editing at the Tyr locus in Albino mice 220
Figure 5.11 5′ modified donors improve knock-in efficiencies of V5 tag at Sox2 locus 223
Figure 5.12 Microinjection information for editing at the Sox2 locus in mouse zygotes 224
Figure 5.13 End-modifications suppress formation of donor concatemers 226
Figure 5.14 End-modified donors suppress donor integration at off-target loci 228

Abbreviations

AAV	Adeno-Associated Virus
AGO	Argonaute
BFP	Blue Fluorescent Protein
bp	Base Pair
CAR	Chimeric Antigen Receptor
CAS	Crispr-Associated
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNA	Crispr RNA
D-loop	Displacement Loop
dCas9	Dead Cas9
DNA	Deoxyribonucleic acid
dsDNA	Double Stranded DNA
EJ	End Joining
GFP	Green Fluorescent Protein
gRNA	Guide RNA
HEK293	Human Embryonic Kidney 293
Indels	Insertions and Deletions
IR	Ionizing Radiation
Kb	kilobase
MMEJ	Microhomology-Mediated End Joining
MosSCI	Mos1-mediated single copy insertion
NHEJ	Non-Homologous End Joining
NLS	Nulcear Localization Signal
Nt	Nucleotides
PCR	Polymerase Chain Reaction
PNA	Peptide Nucleic Acid
RNA	Ribonucleic Acid
RNAi	RNA interference
RNP	Ribonucleic Protein
ROL-6	ROLler-6
sgRNA	Single Guide RNA
SNP	Single Nucleotide Polymorphism
ssODN	Short Single-stranded Oligodeoxynucleotide
SSTR	Single strand Template Repair
SSA	Single strand Annealing
TALEN	Transcription Activator-Like Effector Nucleases
WT	Wildtype
WAGOs	Worm Argonaute proteins
ZFN	Zinc-Finger Nucleases

List of Publications

Parts of this dissertation have been published or will be submitted for publication as:

- **Ghanta, K.S**., Chen, Z., Mir, A., Dokshin, G.A., Krishnamurthy, P.M., Yoon, Y., Gallant, J., Xu, P., Zhang, X.-O., Ozturk, A., et al. (2021). 5′ Modifications Improve Potency and Efficacy of DNA Donors for Precision Genome Editing. bioRxiv, 354480.
- **Ghanta, K.S**., Ishidate, T., and Mello, C.C (2021). Microinjection for Precision Genome Editing in *Caenorhabditis elegans*. STAR Protocols (2021).
- **Ghanta, K.S**., and Mello, C.C. (2020). Melting dsDNA Donor Molecules Greatly Improves Precision Genome Editing in *Caenorhabditis elegans*. Genetics 216, 643-650.
- Dokshin, G.A., **Ghanta, K.S**., Piscopo, K.M., and Mello, C.C. (2018). Robust Genome Editing with Short Single-Stranded and Long, Partially Single-Stranded DNA Donors in *Caenorhabditis elegans*. Genetics 210, 781-787.
- Craig Cameron Mello, **Krishna Sumanth Ghanta**, Gregoriy Dokshin, Aamir Mir, Hassan Gneid, Jonathan Kenneth Watts, Erik Joseph Sontheimer. Compositions and methods for improved gene editing. US Patent application number 16/384,612.

Publications that are not part of this thesis

- Kim, H., Ishidate, T., **Ghanta, K.S**., Seth, M., Conte, D., Jr., Shirayama, M., and Mello, C.C. (2014). A co-CRISPR strategy for efficient genome editing in *Caenorhabditis elegans*. Genetics 197, 1069-1080.

Chapter I: Introduction

Introduction to genome editing

Modulating gene expression and studying the associated phenotypic changes in an organism are central to the understanding of biological processes. For several decades, researchers have relied on random chemical mutagenesis to select for the desired traits and study the associated genotypic changes[1-3]. With the advent of recombinant DNA technology and improvements in DNA transformation techniques transgenesis was made possible[4-6]. Using these technologies, transgenes were introduced into eukaryotic germ cells and their transmission to off-spring was shown[5,7]. These advancements helped in understanding some of the fundamental questions such as how mutant proteins cause diseases or how gene expression is regulated during animal development. Despite great advances, these techniques lacked the ability to perform targeted DNA editing *in vivo*.

In-vivo genome editing is one of the highly sought-after techniques that could help us study and understand the functions of gene products in organisms. The ability to precisely edit any locus in the genome has great promise to treat diseases. The pursuit to edit the genome in a targeted manner led to the development of techniques such as of zinc-finger nucleases (ZFN), transcription activator-like effector nucleases (TALENs) and clustered regularly interspersed short palindromic repeats (CRISPR). Here, I review targeted genome editing

techniques with a focus on CRISPR/Cas9 technology. Specifically, this work focuses on advancements in the field of precision genome editing by Homology Directed Repair (HDR) and methods to improve HDR efficiencies in various model systems.

1.1 Zinc Finger Nucleases and TALENs

A long-standing goal of biologists is to edit the genomes in a targeted manner—to be able to induce Double Stranded Breaks (DSB) in the host genome at a defined locus to delete or precisely insert new sequences. This goal was first achieved by the development of Zinc Finger Nucleases (ZFN)[8]. ZFNs are artificial chimeric proteins consisting of tandem zinc finger domains fused to the cleavage domain of Fok1 restriction enzyme. DNA binding peptides of Zinc fingers generally share the consensus sequence containing Cys_2His_2 motifs. Each zinc finger functions independently and binds a specific triplet of nucleotides[9]. This specificity can be increased by fusing several fingers in tandem. In theory, by altering the combinations in the array, all the 64 triplets in the genetic code can be targeted. Although these peptide arrays can bind DNA in a sequence specific manner, they lack the ability to cleave the DNA strands which is a requirement for genome editing. Therefore, by fusing the nuclease domain of Fok1 protein to the zinc finger arrays ZFNs were shown to be capable of cleaving DNA in a site-specific manner *in vitro*[8]. Efficient DNA cleavage is achieved only with the dimerization of Fok1 domains, therefore ZFNs are used in the pairs with ZF

targeting both the strands in an inverted orientation such that Fok1 domains dimerize in the middle. Once zinc fingers recognize the target DNA, Fok1 dimer induces DNA breaks which will be repaired by the cell's repair mechanisms. Modular peptide arrays containing at least 6 zinc fingers per ZFN targeting 18-24 nucleotides have been shown to provide good specificity and binding efficiencies[10]. Later, innovative strategies were developed to design and select for efficient zinc finger arrays[11,12]. Using this technology, targeted gene deletions were first demonstrated at an endogenous locus in *Drosophila melanogaster*[13] and later in several other organisms[10,14]. Targeted insertions were also achieved by providing exogenous DNA repair templates[15-18]. Although ZFNs have proven to be effective in generating DSBs, design and optimization of efficient ZFNs for every new target sequence can be time consuming, labor-intensive and challenging.

Transcription activator-like effector nucleases (TALENs) are functionally similar to ZFNs. TALEs were discovered in pathogenic bacteria Xanthomonas that colonize plants and alter their gene expression[19]. Each TALE consists of a DNA binding domain that is 33-35 amino acids long and recognizes a single nucleotide in the target DNA. A series of such domains can be custom arranged to target the TALE to a specific genomic locus. TALEs lack nuclease activity, and the arrays are flanked by terminal domains[20,21]. Therefore, for genome editing purposes, TALEs have been attached to a FOK1 nuclease domain similar to ZFNs[22]. These TALENs have been shown to be efficient at inducing targeted

DSBs to edit the genome of several model organisms[23-26]. Because of their simple design and effectiveness TALENs quickly emerged as a powerful tool for genome editing.

1.2 Prokaryotic anti-phage defense systems- CRISPR

The CRISPR-Cas system is an anti-viral adaptive immune mechanism that is highly conserved among prokaryotes. CRISPR loci of the prokaryotic genomes contain repeats of 25-50 bp long DNA sequences interspaced by non-repetitive sequences[27,28]. It has been shown that the interspaced unique sequences (also known as spacers) were acquired from extrachromosomal elements and genomes of bacteriophages[29]. A cluster of coding regions known as Cas genes (CRISPR associated genes) generally flank the CRISPR loci and are also well conserved among prokaryotes[30,31]. These findings lead to the inference that the CRISPR loci and associated proteins may have functions in host defense mechanisms and interspaced unique sequences may be remnants of genetic material from previous phage invasions or extrachromosomal elements[29,31]. Building up on the observations of previous *in-silico* analyses Barrangou et al. demonstrated that presence of CRISPR spacers directly corelates with the bacterial resistance to phage invasions. In addition, they also showed that the proteins Cas5 (now known as Cas9) and Cas7 participate in acquiring and maintaining the phage resistance[32]. The mechanism of phage inactivation became clearer with the finding that spacer sequences code for small RNAs[33] and that the bacteria use these small RNAs to cleave the double stranded DNA

of the invading phages[34]. This finding was surprising because until then it was thought that the small RNAs target RNA molecules, similar to RNA interference (RNAi). In their groundbreaking work, Marraffini and Sontheimer proposed that CRISPR-Cas9 could be used as a tool for genome editing[34]. Later studies demonstrated that Cas9 is the only protein needed for DNA cleavage[35] and a common trans-activating (tracr) RNA binds the spacers to form a complex with Cas9 protein[36]. Based on the genes involved and the genetic architecture of the loci, CRISPR systems are classified into three major types (Type I, II and III)[37]. Whereas Type I and III systems use multiprotein complexes to generate guide RNAs and to cleave the target DNA, Type II systems use only Cas9 protein for maturation of crRNA and DNA cleavage[35,37]. Because of the simplicity of the type II system, efforts were focused on using Cas9 protein for genome editing purposes. Purification and biochemical characterization of Cas9 and guide-RNA (tracrRNA and crisprRNA) complexes showed that this system can cleave dsDNA *in vitro* and remarkably only needs Cas9, crRNA and tracrRNA[38,39]. Interestingly, Cas9/guide RNA complexes cleave their target DNA only in the presence of NGG proto spacer adjacent motif (PAM) downstream of guide binding sequence. CRISPR loci in the bacterial genomes lack PAM motif and therefore PAM aids in differentiating self-DNA from foreign-DNA. In summary, CRISPR-Cas systems are prokaryotic anti-phage defense mechanisms that capture and genetically register the nucleic acid fingerprints (spacers) of the

viruses during infections and utilize these fingerprints to mount immune responses during future infections.

1.3 CRISPR/ Cas9 Technology

1.3.1 Mechanism of DNA cleavage by Cas9 guideRNA complex

Cas9 is an endonuclease that contains two nuclease domains, HNH and RuvC[38,39]. Each of these domains cleave one strand in the double stranded target DNA. Cas9 protein is directed to the target DNA by its guide RNA. Each guide RNA consists of a ~36bp crispr RNA (spacer + repeat, crRNA) that provides target specificity and a trans-activating CRISPR RNA (tracrRNA) that basepairs with the crRNA. Each Cas9-gRNA ribonucleoprotein (RNP) complex first searches for PAM (NGG) sequence in the target genome, then through its helicase activity Cas9 unwinds the double stranded DNA in a directional manner that starts at the 3′ end of the target and proceeds towards the 5′ end[40]. crRNA then base pairs to the target sequence that is immediately upstream of the PAM site. Once the complex is stabilized, the HNH and the RuvC domains of Cas9 cleave the target strand (strand base paired with the crRNA) and the non-target strand (complementary strand) between 17th and 18th basepairs of the duplex DNA respectively. These cleavage events generate a precise blunt ended double strand break that will be fixed by the cell[38-40]. The affinity and the stability of the Cas9 RNP complex on the target DNA primarily depend on the sequence composition of the crRNA. For example, Cas9 complex is destabilized if the crRNA sequence doesn’t base pair with the target DNA. Similarly, mismatches

between the target DNA and the crRNA, specifically from 14th nucleotide to 20th nucleotide (nt), are not well tolerated. These mismatches may significantly reduce the efficiencies or completely abrogate the cleavage events[39,41]. In addition, mutations in the nuclease domains of Cas9 also eliminate the cleavage activity of the Cas9 RNP complexes both *in vitro* and *in vivo*[39]. Understanding the factors that influence the dynamics between Cas9 and the target DNA helps to achieve better on-target efficiencies and to suppress off-target cleavage activity. How modulating these factors influences Cas9 activity are discussed in the sections below.

1.3.2 CRISPR/ Cas9 Technology for genome editing

CRISPR-Cas9 genome editing technology essentially involves two main events: 1. generation of DSBs in the host genome and 2. repair of the induced breaks in the DNA. Double strands breaks are one of the most toxic lesions in cells[42]. Therefore, cells employ sophisticated machinery to quickly recognize and repair these DSBs. DSBs are repaired either by non-homologous end joining (NHEJ) or homology directed repair (HDR)[42,43] (Figure 1.1). Generally, NHEJ is more efficient, and it is also the dominant pathway compared to the HDR pathway. Therefore, for genome editing purposes end-joining is the preferred route to obtain loss of function alleles by generating simple indels (insertions and deletions). HDR on the other hand is more complex and less efficient, but it is also more precise compared to the NHEJ pathways. Therefore, HDR is the

preferred pathway when precision editing is required. However, the type of pathway used by the cell depends on several factors such as end-resection, homology search, cell type, cell cycle, type of the lesion etc. Improving the efficiencies of these two pathways has been a primary goal in all genome editing applications. Therefore, understanding the functions of the factors involved and the overall mechanisms of these major repair pathways helps in improving the editing efficiencies.

1.4 Mechanisms of double strand break repair

1.4.1 Non-homologous end-joining

DNA double strand breaks are routinely induced by cell intrinsic factors such as reactive oxygen species, replication stress and transposon jumping. Similarly, exposure to environmental factors such as ionization radiation, ultraviolet light or chemical agents can cause DNA damage. If these lesions are not repaired, they can lead to genotoxicity and cell death[42,44]. To quickly repair such pathological lesions, cells have evolved a repair mechanism called non-homologous end-joining (NHEJ). Fundamentally, this pathway involves direct ligation of the broken ends, and it is not dependent on the nucleotide composition at the break point. In NHEJ, DSBs are quickly fixed through a series of steps involving Ku70-Ku80, nuclease complex containing artemis and DNA dependent protein kinase (DNA-PK), DNA polymerases (pol lambda and pol mu) and DNA ligase complex of ligase IV (Lig4) and XRCC4[44,45]. Deficiencies in this pathway have been

implicated in chromosomal aberrations, malignant transformation, deficiencies in V(D)J recombination and several other diseases[46-48].

When DSBs arise, free ends at the break site are first recognized by evolutionarily conserved Ku70 and Ku80 heterodimers[49]. Since end resection triggers homology directed repair, end-protection by Ku proteins is crucial in directing the cells towards NHEJ[50,45]. The Ku70/Ku80 heterodimeric complexes on each end of the DSB then recruit nuclease complexes of Artemis and DNA-PK. Artemis:DNA-PK nuclease complexes have 5′ and 3′ exonuclease activity that can resolve incompatible DNA ends (such as overhangs and hairpins) into blunt ends. Artemis:DNA-PK complxes are also required to remove damaged nucleotides[51,44]. Once the ends are resolved, the ligase complex comprising ligase IV, XLF and XRCC4 joins the ends of one or both the strands. *In vitro* biochemical experiments showed that XRCC4::Lig4 complex alone is sufficient to complete the ligation of sticky or blunt ends[52]. Finally, DNA polymerases Pol λ (lambda) and Pol μ (mu) directly interact with Ku:DNA complex and are recruited to fill in any gaps at the ligated junctions[53,54].

Because of its relatively simple mechanism with few proteins and rapid repair process, NHEJ is the most efficient and predominant pathway for fixing a DSB. However, since the ends are directly ligated, insertions, deletions, or replacements of sequences at the breakpoint are frequently detected (Figure 1.1). Variation in the nuclease activity of Artemis:DNA-PK complexes and the addition of nucleotides by the polymerase leads to heterogenous repair

outcomes. Even though NHEJ is not precise it helps the cells by quickly ligating the broken ends of the DNA and by reducing the toxicity associated with DSBs. Therefore, to generate loss of function alleles utilizing cell's NHEJ mechanism is the most efficient method. However, it is not useful if precision editing is required. But in cells (e.g., non-dividing cells) where HDR is not active one has to primarily rely on end-joining pathways for genome editing (See HDR section for details).

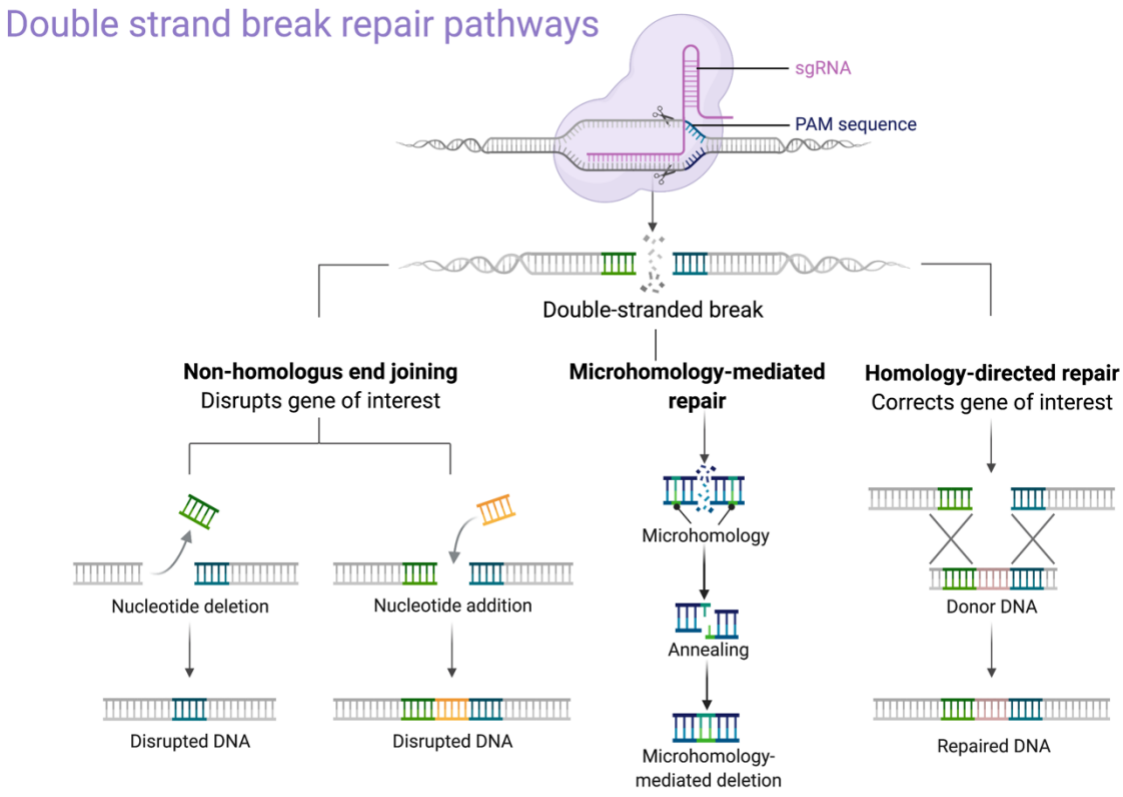


Figure 1.1 Double strand break repair pathways.

Double strand breaks are fixed by one of the three dominant forms of repair pathways. In NHEJ ends are directly ligated in a sequence independent manner but MMEJ may become the predominant form of repair if the sequences at the ends are homologous to each other. In the absence of end-joining pathways and in the presence of a donor template (e.g., homologous chromosome or exogenously provided donors) cells use HDR to fix the break in a precise manner.

1.4.2 Micro homology mediated end joining (MMEJ)

Microhomology mediated end joining can be classified as a special case of NHEJ pathway. Like NHEJ, MMEJ does not require a homologous allele to repair the DSB, instead this pathway uses the homology between the sequences flanking the DSB, deletes the sequence that is internal to the homologous sections and ligates the ends (Figure 1.1). But unlike classical NHEJ, MMEJ does not require Ku proteins and Lig4. Studies in yeast have demonstrated that even strains lacking functional Ku are capable of repairing the DSBs albeit at much lower (20-fold) efficiencies. Interestingly, in Ku mutants, repair results in the deletion of few nucleotides near the DSB[49-55]. Using yeast strains containing artificially inserted Ho endonuclease sites, Ma et al. found that deletions of same type were present in several independent colonies and these deletions were caused by micro-homology of about 8-bp between the sequences on either side of the endonuclease cleavage sites[56]. These observations suggested that small stretches of homology between the ends on opposite strands triggers MMEJ pathway and the length of the homology is thought to be relatively short (5 to 25 nt).

MRN complex (Mre11/Rad50/NBS1) initiates MMEJ by binding to the internal region at the DSB[56]. CtIP binds to the MRN complex and activates the nuclease activity of Mre11[57,58]. Mre11 then resects the strands internally from 3′ to 5′ to expose the homologous 3′ overhangs on both sides of the DSB[59,60]. Mre11 mutants fail to initiate MMEJ as overhang formation to expose microhomologous

regions is a requirement for this pathway[61-63]. Once the homologous sequences are annealed, ERCC1-XPF endonuclease removes the non-homologous 3′ flaps that are internal to the microhomologous regions. Finally, gaps are filled by polymerase Q (PolQ) and the ends are sealed by Lig1 and Lig3[63]. Flap removal in MMEJ results in deletions of the internal sequences. Therefore, MMEJ can be mutagenic. But because MMEJ produces the same type of deletion in majority of the cells, it can be a used to delete pathological repeat sequences just by inducing a DSB[64] without using donor templates.

1.4.3 Homology directed repair

Homology directed repair is the most precise and probably the most complex form of DSB repair. Unlike NHEJ or MMEJ, homology-based repair uses a homologous sequence as template to repair the DSBs. Homologous Recombination (HR) can be initiated either by programmed double strand breaks or by unscheduled DNA breaks. Defects in HR pathway lead to accumulation of damaged DNA or hyper recombination resulting in cancers[43].

HR is a tightly regulated process and the factors involved in this pathway are evolutionarily conserved across all domains of life. Defects in this pathway result in abnormal cell division, developmental arrests, and cancers[43,65,66]. HR also plays a critical role in meiotic recombination, meiotic chromosome segregation and DNA replication. Meiotic homologous recombination is initiated by topoisomerase like enzyme, Spo11[67-70]. Spo11 induces programmed DNA breaks and covalently attaches to the 5′ termini of the free DNA ends[67,71,72].

Endonuclease protein complex containing Mre11, Rad50 and Xrs2 (MRX complex) and Ctp1 cleave the DNA strands on either side of the Spo11 complex[73,74]. This endonuclease activity releases Spo11 and the oligonucleotides bound to it (called Spolligos)[74] exposing the 3′ ends of the DSB. The 5′ ends are then resected to produce long single stranded overhangs. The extent of resection and the length of the 3′ overhangs dictate, to a large extent, whether the breaks will be repaired by HR, Single strand annealing (SSA) or MMEJ. If the resection is long enough, RecA family proteins Rad51 and Dmc coat the 3′ overhangs and initiate the homology search. Once the template has been identified on the homologous chromosome, the 3′ overhangs invade the unwound duplex leading to the formation of D-loops. DNA polymerase δ (Polδ) then extends the 3′ end using the homologous DNA as template. After synthesis, the repair process can be completed either by crossover or by non-crossover. In case of crossover, the other 3′ overhang is captured (also known as second end capture) and extended by Polδ to form a double holiday junction (HJ). The dHJ is then resolved by Resolvase resulting in crossover recombinants[43,75]. In the second mechanism, the first 3′ extended over the homologous template is released from the D-loop to anneal to the other 3′ overhang at the DSB. Finally, through a series of steps both the 3′ ends are extended and ligated to the 5′ ends to fill the gaps. This form of the repair is known as synthesis dependent strand annealing (SDSA) and doesn’t result in homolog crossovers (non-crossover).

Non programmed DSBs are induced by a variety of factors such as replication stress, chemical agents, radiation, or endonucleases. Even though these DSBs are independent of Spo11 mediated cleavage the downstream homology directed repair steps are essentially identical to HR. MRN nuclease initiates 5′ to 3′ resection at the break and Exonuclease (EXO1) extends the resection[76]. Finally, RPA coats the single stranded overhangs, initiates homology search, and completes repair.

For the purposes of genome editing, DSBs induced by Cas9 nuclease can be fixed using exogenously supplied template DNA. The precision nature of the HDR pathway allows insertion of long edits into the host genomes. However, it is important to note that HDR pathways are tightly regulated according to the cell cycle. HDR has been shown to be active primarily during S and G2 phases. Furthermore, some cell types such as post-mitotic (neurons) are thought to lack HDR capabilities. Therefore, the repair outcomes in these cells mostly contain imprecise edits that result from NHEJ or MMEJ pathways. Understanding the factors that influence the choice of the repair pathway and the repair outcome could help in increasing the editing efficiencies as discussed in later sections.

1.5 Factors affecting efficiency of genome editing

Breaks in the host genome are necessary to incorporate user defined precise changes. Therefore, the success of genome editing depends on Cas9 cleavage

efficiencies. It is important to understand the main factors that contribute to the success or failure of Cas9 cutting efficiencies. Furthermore, understanding the dynamics between the sgRNA and its target sequence also helps in better prediction of the off-target sites and off-target donor integration.

In this section, I discuss the factors that significantly influence efficiencies of Cas9 cleavage and HDR *in vivo*. I review recent technological advancements that improve editing efficiencies.

1.5.1 Factors influencing the cleavage efficiencies of Cas9

Target site sequence

The indel inducing efficiencies of Cas9/guide RNA complexes can vary widely between targets. Based on the cutting efficiencies of a guide RNA it can be classified as a good guide or a bad guide. Nucleotide composition of the target sequence determines the indel efficiencies to a great extent. It has been shown that the nucleotides on the PAM proximal end (14 to 20) of a target sequence determine the performance of the guide RNA. Mismatches in this region strongly inhibit cutting whereas mismatches on the PAM distal end of the target sequence (1-10 nt) are generally tolerated. These findings suggest that the stability of R-loop between the guide RNA and the target strand is critical to achieve optimal Cas9 efficiencies. How can the stability be increased?

Serval groups have performed large scale multiplexed gRNA screens in human cell lines to identify signatures of efficient and inefficient guides. Using deep

sequencing and machine learning techniques guide RNA prediction tools have been developed[77]. First, it is has been shown that gRNAs with high GC content are more efficient at inducing indels at the target site. Interestingly, however, gRNA with more than 80% (16 to 20 nt) GC content performed poorly[78,79]. Furthermore, gRNAs that contain guanine at position 20 performed better and cytosine at this position was disfavored. In contrast, cytosine was preferred to guanine at position 16[78]. Taken together these studies show that 16 to 20 nt of the gRNA are critical for Cas9 activity, therefore this region is referred as 'seed region'[79]. Using zebrafish embryos as the model system, Moreno-Mateos et al. showed that the stability of the gRNA directly contributes to the higher cleavage activity. They suggest that gRNAs containing more guanines on the 5′ end could form G-quadruplex structures that can minimize exonuclease activity and improve half-life of the molecules in the cells[80]. Finally, sequence downstream of PAM site has also shown to influence the cutting efficiencies. In zebrafish experiments guanine depletion in this region corelated with better cutting efficiencies. In contrast, guides that end with GG motif upstream of NGG performed 10-fold better than guides without GG motif in *C. elegans*[81]. Similarly, G at position 20 is strongly preferred in mammalian cells[82,83]. Future improvements in machine learning algorithms with large scale multiplexing assays would further improve the prediction quality.

The structure of local chromatin and transcriptional state of the target sequence is another determining factor of guide RNA efficiencies. Even if the sequence

rules of the sgRNA design are followed cleavage efficiencies are lower in heterochromatin regions[84,85]. In contrast some studies showed that heterochromatin regions do not significantly affect Cas9 performance[86]. Interestingly, at the same loci other nucleases such as Cpf1 performed much worse than spCas9[86,87]. At loci where heterochromatin strongly hinders the nuclease activity it would be difficult to achieve optimal editing. Therefore, further improvements in guide design are needed to address this challenge.

Modality of Cas9 and Guide RNA delivery

For genome editing applications, particularly *in vivo*, proper delivery of Cas9 and guide RNAs into the target cells and their stability in the cells are crucial factors that determine the editing efficiencies. Using viral vectors (e.g., AAV) encoding Cas9 and guideRNAs is an efficient way to achieve *in vivo* editing. However, toxicity and the possibility of integration of viral vectors into chromosomal breaks pose challenges[88]. It has also been shown that delivering Cas9/gRNA either as mRNAs or as preassembled RNP complexes yields more efficiency compared to Cas9/sgRNA in DNA form[88,89]. Furthermore, RNPs or mRNA of Cas9 are transient and reduce the side effects (such as off-target indels) associated with long term expression of Cas9[90-92]. Although exogenous Cas9/ sgRNA DNA constructs can be more toxic they are also more stable in cells. On the other hand, Cas9/ sgRNA RNPs or mRNAs are less toxic yet less stable. Therefore, achieving a good balance between toxicity, stability and efficacy is necessary to successfully implement genome editing protocols.

Chemically modified guide RNAs

RNAs are generally not stable in cells and are prone to degradation by nucleases[93]. Exogenous RNA triggers Toll like receptor (TLR) mediated innate immune pathways[94] and *in vitro* generated guide RNAs elicited strong immune responses in human cells[95]. Therefore, to develop potent guide RNAs that are amenable to non-viral delivery (e.g. lipid nano particles (LNPs)), nuclease resistant, immune neutral and stable in cells for longer periods researchers have resorted to chemical modifications. Chemically modified RNAs have proven to be beneficial and effective in other therapeutic applications such as RNAi and ASOs[96-98].

A variety of modifications were incorporated into synthetic guide RNAs to improve editing efficiencies[99-104]. In mammalian cell lines, guide RNAs modified at the 5′ and 3′ termini with 2′-O-methyl (2′OMe), 2′-O-methyl 3′phosphorothioate (PS), or 2′-O-methyl 3′thioPACE modifications improved indel frequencies over unmodified guides. Particularly, guides with phosphorothioate and thioPACE modifications increased indel frequencies more than 25-fold in tissue culture settings[103]. Guided by the crystal structure of Cas9, indel efficiencies were further improved by using fully modified guide RNAs[99,101]. Replacing most or all the 2′ OH groups in crRNA and tracrRNA with 2′-deoxy-2′-fluoro-ribonucleotide (2′F RNA) or 2′OMe modifications significantly enhanced editing efficiencies. The importance of chemical modifications is further underscored by efficient editing in primary cells and in mouse liver[101,105]. Incorporation 2′-O-methyl-3′-

phosphonoacetate modifications reduced off-target cutting without compromising on target efficiencies[106]. Interestingly, however, other terminal modifications such as amine, DBCO or azide are tolerated but did not improve editing efficiencies over unmodified guide RNAs[100] suggesting that internal modifications are critical to increase stability. In summary, chemical modifications to guide RNAs allow Cas9 protein and Cas9 mRNA to be delivered in LNPs to obtain high editing efficiencies for *in vivo* and *ex vivo* applications. Importantly, these chemically modified sgRNA did not increase off-target indels[99,103,104,106].

1.5.2 Cas9 and Guide RNA engineering to improve specificity

Wildtype spCas9 tolerates a few mismatches in the guide sequence that could contribute to high off-target activity for some guideRNAs[107]. Therefore, in addition to improved on-target cutting efficiencies genome editing applications require improved specificity i.e., elimination of off-target indels. To achieve this goal both the nuclease and its guide should be optimized.

High-Fidelity Cas9 Variants

Search for Cas9 variants that show reduced off-target cutting has led to the development of a quadruple substitution variant (N497A/R661A/Q695A/Q926A). Guided by structural studies Kleinstiver et al engineered the quadruple variant and demonstrated that off-target activity can reduced[108]. Later studies showed that a single mutation at position 691 (arginine to alanine) in the Cas9 protein has been shown to remarkably[109] reduce off-target cutting while retaining high on-target efficiencies[105]. Data from the single mutant strongly indicates that other

multiple mutations are unnecessary to achieve high fidelity. Cas9 orthologs from other bacterial species also have shown higher specificity compared to wildtype spCas9[110,111].

RNA-DNA chimeric guides

To reduce guide RNA off-target binding chimeric RNA-DNA guides have been engineered. To this end, replacing the 5′ end (1-6 nt) of guide RNA sequences with DNA nucleotides has significantly reduced off-target activity while retaining on-target efficiency. However, 3′ end of the guide is more sensitive to nucleotide changes and abrogated on-target activity in human cell cultures[109,112]. Similarly, DNA substitutions at the PAM distal region of Cas12a guide RNAs improved specificity but compromised on-target activity as well. These chimeric guide RNAs would be particularly beneficial in applications where high on-target activity is not necessary to rescue the phenotype of the mutant allele. For example, in disease models where cells with wildtype copies gain proliferative advantage over the cells with mutant copies of the gene.

Truncated guide RNAs

As discussed in the above sections, guide RNAs are particularly sensitive to mismatches or chemical modifications in the PAM proximal regions (14-20nt). However, modifications to the 5′ end are generally well tolerated. Using this finding as the foundation, efficiencies of truncated guideRNAs were evaluated for on-target and off-target activity. Guides RNAs lacking one, two or three 5′ terminal nucleotides performed as efficiently as full-length counterparts.

Additionally, these short guide RNAs were inefficient at inducing off-target indels[113,114]. Using shorter guides (19 nt) in combination with Hi-Fi Cas9 (R691A) variant virtually eliminated all the off-target indels[114]. It is imperative that future genome editing therapeutic platform switch to this combination of Cas9 and guide RNAs.

Extended guide RNAs and tracrRNA

Endogenously, in bacteria, tracrRNA is necessary to produce matured crRNAs. crRNA and tracrRNA form a duplex at the complementary repeat sequence. For efficient use in genome editing applications, the two RNAs have been joined through GAAA tetra loop to form a chimeric single guide RNA[39]. The rest of the tracrRNA forms three more stem loop structures (stem loop 1, 2 and 3). To optimize cutting efficiencies, some recent studies have focused on engineering secondary structure of guide RNA by modifying the 5′ end of the spacer sequences or by extending the stem loops. To this end, Kocak et al, designed a robust hairpin structure on the 5′ of the sgRNA that destabilizes the interaction between the sgRNA and off-target sites due to higher energy requirements for R-loop formation. Unlike previous 5′ sgRNAs that were shown to be processed into 20nt spacers, hairpin-sgRNAs were stable in cells and elicited high cutting efficiencies with reduced off-target activity[115]. Interestingly, hairpins did not reduce binding between spacer and the off-target DNA. Incorporation of aptazyme into sgRNA sequence is an attractive strategy to control the dynamics and kinetics of Cas9 guideRNA complexes. Blocking the 5′ end of the guide

sequence and controlling the binding through small molecule dependent aptazyme has resulted in improved editing efficiencies. This method offers control over the cleavage dynamics and limits Cas9 RNPs activity to a defined time window[116]. Similarly, extending repeat-anti repeat duplex at the 3′ end of crRNAs and stem loop 2 has significantly increased Cas9 cutting efficiencies[117,118].

Variables that are hard to control

Factors that influence Cas9 cutting efficiencies discussed so far can be optimized or controlled to achieve higher editing efficiencies. However, factors such as cell type, delivery considerations, tissue specificity for *in vivo* editing, lack of PAM sites at the locus of interest are difficult to manipulate. For example, immune and hematopoietic cell lines are very difficult to transfect. Similarly, neurons are hard to edit. Preexisting immunity to Cas9 is another factor that could compromise *in vivo* editing efficiencies. Streptococcus pyogenes, the species of origin of most widely used Cas9 protein, frequently infects humans[119]. Prior infections of these bacteria generate antibodies against Cas9 through adaptive immune responses[120-123]. The potency of therapeutic Cas9 could be significantly dampened by the pre-existing immunity that may lead to failure in genome editing. Safety and specificity of Cas9 RNPs is of utmost importance for therapeutic applications. By combining the most effective strategies, future engineering of Cas enzymes and sgRNAs could yield a more robust system with high specificity and efficiency.

1.6 Factors influencing HDR and methods to improve efficiencies

Pathogenic genetic variants such as single nucleotide mutations cause more than four thousand unique monogenic disorders in humans[124]. Such mutations lead to diseases and need to be corrected to alleviate the disease symptoms. Corrections of such genetic defects at the DNA level needs precision genome editing. Using cellular HDR pathways, precise edits can be introduced into the genome by providing exogenous DNA donor templates. Precise edits are also needed to study gene functions in animal models. Using CRISPR/cas9 technology and donor DNA, introducing epitope tags, small inserts or SNPs into the locus of interest facilitates studies of biological pathways. Unfortunately, most cell types and model systems have been more resistant to HDR mediated changes as compared to simple indels by NHEJ. Therefore, there is a great need to understand the crucial factors that influence HDR and to improve the technologies to achieve better efficiencies. Also, it is critical to achieve the best editing with little amounts of donor DNA and without toxicity to the cells. In this section, I discuss the cell extrinsic and cell intrinsic factors that influence HDR efficiencies either positively or negatively. I also discuss the future perspectives and recent technologies that could increase the precision.

Cell type and cell cycle dependence

The cell type in which genome editing must be performed plays a significant role in determining HDR efficiencies. Terminally differentiated cells (e.g., neurons) are

more resistant to genome editing compared to transformed cell lines such as HEK293T[125]. Two factors stand out as reasons for low HDR efficiencies in non-dividing cells. First, HDR is typically active only during S and G2 phases of the cell cycle[126]. BRAC1 binds to DNA at the DSBs and promotes end-resection. BRAC1 also recruits PALB2, BRAC2 and RAD51 to the DSB to initiate homology search and repair. However, BRCA1 protein complex formation is inhibited during the G1 phase which indirectly suppresses HDR. During the S phase, cells promote HDR by allowing recruitment of BRAC1 and its cofactors to the DSBs. During the M-phase, mitotic kinases (CDK11 and PLK1) phosphorylate RNF8 and 53Bp1 and inhibit their recruitment to the DSB thereby suppressing HDR[127,128]. Therefore, poor efficiencies in non-dividing cells could be due to restriction of HDR to G2-S phases. Second, absence of nuclear envelop breakdown to facilitate nuclear entry of the donor DNA could contribute to low HDR efficiencies. Nuclear pore complex limits the size of the DNA (to less than 500bp) that can passively diffuse from cytoplasm to nucleus[129]. Therefore, large donors (> ~1kb) may gain nuclear access only when the nuclear envelope breaks down during cell divisions.

NHEJ based strategies circumvent inefficiencies associated with HDR in non-dividing cells. Homology-independent targeted integration (HITI)[130] has shown to be successful in editing non-dividing cells[131]. However, in this approach, cells directly ligate the donor molecules into genomic DSBs instead of using them as templates for repair. This direct ligation could introduce undesired indels at the ends of the integration. Therefore, although this approach can be used to introduced

whole gene or exons (with ends in introns) it would not work for correction of single nucleotide mutations.

In dividing cells, HDR can be enhanced by pharmacological cell cycle synchronizing[132]. Cells arrested with aphidicolin or nocodazole and released immediately prior to transfecting (or nucleofecting) with Cas9 RNPs and donor DNA yield high HDR efficiencies[132]. Similarly, the activity of Cas9:geminin fusion protein is restricted to S and G2 phases and improved HDR efficiencies[133]. These studies demonstrate that HDR is tightly regulated according to cell-cycle phase.

Suppression of NHEJ to favor HDR

Every DSB can be fixed by one of the several repair pathways (NHEJ, MMEJ, HDR etc). Unfortunately, in somatic cells end-joining (EJ) pathways precede HDR. In addition, repair through EJ pathways proceeds faster compared to HDR which is needed to protect the cells from accumulation of deleterious chromosomal breaks. However, in precision genome editing, this dominance of EJ pathways reduces HDR efficacy. In order to increases the frequencies of HDR mediated repair, methods to suppress EJ pathways were explored[134-139] (Figure 1.2).

KU70 and KU80 heterodimer binds to DSBs and recruits the downstream factors in the NHEJ pathway. Inhibiting KU proteins by siRNAs suppresses NHEJ and promotes HDR in human and mouse cell lines[134]. Similarly, pharmacological inhibition of DNA-PK (by NU7441)[138], 53BP1[136] or ligase IV (by Scr7) suppressed NHEJ events and improved HDR efficiencies[135]. Co-injecting CRISPR Cas9

mixtures and Scr7 has also showed promise in improving HDR efficiencies in mouse zygotes[135]. Similar NHEJ inhibition approaches used in *C. elegans*[140] and drosophila genome editing have improved HDR efficiencies[141]. However, it is important to note that the effects of NHEJ inhibitors vary between cell types warranting careful testing in the recipient cells[137]. Also, suppression of NHEJ could lead to accumulation of DNA breaks, inability to perform V(D)J recombination and activation of apoptotic pathways[48,142-144]. Therefore, further studies are needed to characterize the side effects and to use this strategy for *in vivo* applications. Interestingly, a screen for HDR enhancers identified chemicals that may not suppress NHEJ components. Blocking ab3-adrenergic receptor by L755507 or inhibition of intracellular protein transport from the ER to the Golgi apparatus by Bre-feldin A increased HDR by 2-3 fold[145].

Ectopic expression or stimulation of proteins in HDR pathway

HDR is mediated by factors that participate in end-resection (eg., BRCA1 complex) and homology search and strand invasion (eg., RAD51). Instead of suppressing end-joining pathways, strategies have been explored to stimulate proteins involved in the HDR pathway. To this end, small molecule screens were performed in human cells to identify stimulants of HDR activity[146,147]. RAD51-stimulatory compound (RS-1) acts as a RAD51 agonist and stabilizes its active form on single stranded overhangs at DSBs. *In vitro* experiments demonstrated that RAD51 stabilization

increases the lengths of RAD51 ssDNA filaments[147]. Using RS-1increased HDR in human cells at Cas9 induced DSBs[146] and in zebrafish embryos[148].

Ectopic co-expression of RAD52 and dominant negative form of 53BP1 promoted HDR when ssODN donors were used as repair templates. Interestingly, however, this strategy did not improve HDR with dsDNA donors suggesting that the factors may promote SSA pathway[149-151]. CtIP is required for strand resection at DSBs. Specifically, the N- terminal fragment is required for localization of Ctip to DNA break sites and to promote recombination[152]. Cas9 fused to the N-terminal fragment of CtIP promoted HDR[153,154]. In contrast, Over-expressing hyperrecombination variants of BRCA1 produced mixed results. While $BRCA1^{M1775R}$ modestly increased HDR in HEK293A cells $BRCA1^{K1702M}$ did not show any difference as compared to cells with over-expression of wildtype BRCA1 or no over-expression[146]. These studies clearly demonstrate that by modulating recombination pathway proteins HDR can be improved. Essentially, all the strategies discussed here increase concentration of recombination factors near DSBs and promote templated repair.

Donor Types and optimization

Exogenous donor templates are required to introduce precise changes into the locus targeted by Cas9 nuclease. Design and optimization of donor templates are crucial to achieve best HDR efficiencies. The type of donor employed largely depends on the size of the desired insert. For longer insertions (e.g., > 3 KB) plasmids are preferred whereas linear dsDNA donors are used for inserts 200bp to

3kb long. However, for inserts shorten than 200bp or single nucleotide changes ssODN donors are preferred.

ssODNs perform better than other donors presumably because they engage SSTR pathway instead of homologous recombination[155-157]. By optimizing the design of ssODNs several groups have achieved increased knock-in efficiencies. Designing ssODNs that are complementary to the non-target strand showed improved editing efficiencies[156,158]. It was also shown that by using asymmetric lengths of homology arms around the DSB, efficiencies could be further boosted compared to symmetric donors[156]. Using nuclease resistant phosphorothioate-modified ssODNs efficient knock-ins were obtained in rats and mice[157]. Single stranded donors proved to be efficient in all most systems due to minimal toxicity[155-157,159-161]. Although less popular, long single stranded donors have been also shown to be efficient[162,163]. But generating long single stranded donors can be cumbersome and time consuming. single stranded (long and short) donors also introduce more mutations at the junctions and may not be ideal for longer inserts.

Linear or circular dsDNA donors are preferred for longer inserts (>200bp). dsDNA has unique advantages over other types of donors: 1. Long insertions can be achieved, 2. dsDNA engages HR pathway instead of SSTR which is more precise 3. dsDNA can be easily generated by PCRs or by simply cloning the plasmids in bacteria. However, dsDNA donors also pose challenges such as toxicity and integration into off-target locations in the genome[88]. The general design considerations for dsDNA have largely remained simple without much innovation—

the insert is flanked by homology arms that are at least 90bp[132,164]. One of the few design changes made to dsDNA donors was to use biotinylated termini and Cas9:streptavidin fusion protein to improve HDR (see later section).

RNA templated DNA repair is an intriguing phenomenon that has been shown to be active in yeast[165,166]. Similarly, RNA donor templates facilitated precise gene editing in rice plants[167]. However, this pathway is not well explored in mammalian systems. A recent study proposes that DNA polymerase (Polθ) can use RNA templates for DNA repair in human cells[168]. It is possible that most organisms can use RNA as repair templates, but RNA may not be available as mature transcripts are actively transported out of the nucleus. RNA may be less toxic and less immunogenic compared to DNA donors. Therefore, Further studies exploring this concept will be of great value.

Single stranded circular DNA templates used as repair templates have shown improvements over dsDNA donors[169]. Finally, studies in mouse and human embryos have shown that heterozygous mutant alleles can be fixed without using exogenous donor DNA. A DSB on the mutant allele triggers repair and uses the wildtype homolog as the donor (interhomolog repair)[170,171]. The frequency of interhomolog repair and therefore homozygosity can be improved by over expression of RAD51. Although this phenomenon is very efficient in embryos it may not be suitable for editing in somatic cells[170]. Therefore, exogenous donors may be more efficient to edit somatic cells.

Controlling Donor localization

The availability of donor molecules at DSB is critical to facilitate precise templated repair. In the absence of donors, DSB can be either repair by end-joining mechanisms or by using the homologous allele as the repair template. Increasing the local concentration of donor molecules by controlling their spatial and temporal presence increases HDR efficiencies. Co-delivering Cas9-strepatavidin fusion protein and biotinylated donors has been shown to improve HDR efficiencies in human cells and mouse zygote injections[172-175]. Covalently linking O6-benzylguanine (BG)-labeled donor molecules to Cas9-SNAP tag fusion protein showed similar positive effects on HDR efficiencies[176]. Instead of using chemical modifications, domains that recognize specific sequences on donor DNA have been fused to Cas9 to physically link Cas9 to donor molecules[177-179]. Some of the domain include transcription factor domain or Porcine Circovirus 2 (PCV) Rep domain[177,178]. Similarly, using dead Cas9 (dCas9) RNPs to complex with donor may promote HDR[164]. HDR improvement is dependent on the presence of Nuclear Localization Signal (NLS) on dCas9 protein. On similar lines, engaging two or more pathways to import Cas9 and donor into the nucleus would be more effective than relying on just one pathway. Lui et al optimized the NLS sequences on Cas9 and using a combination of two different sequences they demonstrate that both Cas9 cutting and HDR efficiencies have increased significantly in mammalian cells and zebra fish[180].

In summary, Increasing the nuclear concentration of donor molecules increases the chances of finding repair templates and promotes HDR.

1.7 CRISPR in Clinic

The fundamental advancements made in CRISPR/Cas9 technology paved way to bring it into clinical settings in less than a decade. Several non-profit organizations and biotechnology companies are actively working to test the safety and efficacy of Cas9 mediated DNA editing in patients. Currently, most of the clinical trials are focused on fixing genetic mutations in the hepatic system, ocular system or defects in hematopoietic cells (*ex vivo*) due to the ease of delivery. Lack of tissue specificity with non-viral delivery has limited the treatments to these cell types.

Advances in lipid nano particle (LNP) technology have demonstrated high levels of editing by Cas9 in liver[181]. Recently, clinical trials have shown that CRISPR technology could potentially cure Transthyretin amyloidosis by *in vivo* editing the mutant TTR gene in liver[181]. A Single injection of Cas9 (mRNA) and guideRNA targeting mutant transthyretin (TTR) gene in liver has resulted in 90% reduction of blood TTR levels[181]. Many other clinical trials are also underway[182]. Other conditions that could be corrected with CRISPR technology include Sickle Cell Disease and Beta-Thalassemia[183], cholesterol reduction to control heart diseases by suppressing PCSK9 expression in liver[184] etc. With improvements in HDR efficiencies, engineering allogenic and autologous chimeric antigen receptor (CAR) T-cells would become much easier and faster. Although these treatments

have been safe and effective so far, it is also important to carefully analyze the long term off-target effects of CRISPR/Cas9. Advances in non-viral, tissue specific delivery methods will enable targeting solid tumors and other genetic defects that remain incurable today.

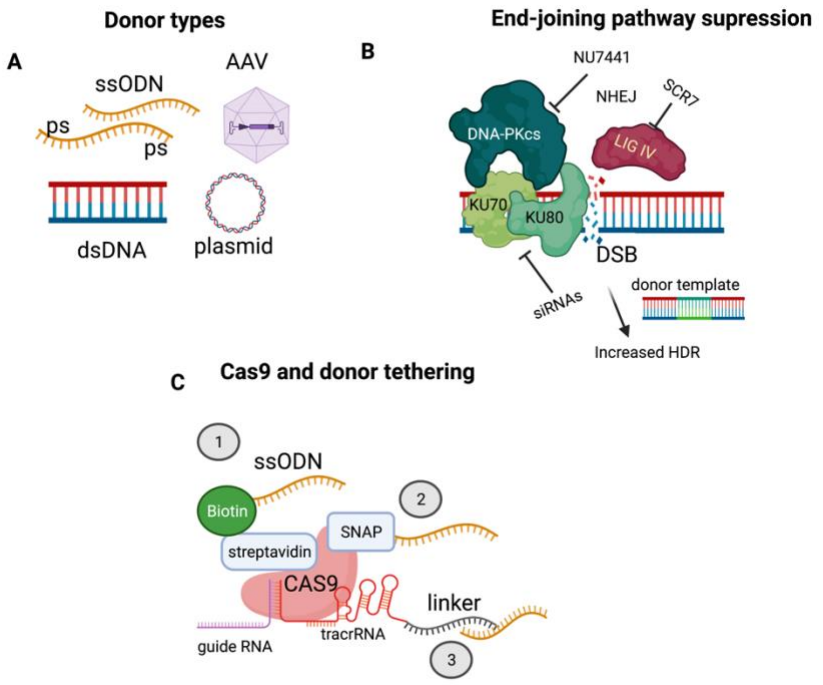


Figure 1.2 Strategies to improve HDR efficiencies. A) Various types of donor modalities used for templated repair are shown (ps: phosphorothioate). B) Suppression of competing end-joining pathways either genetically or chemically improves HDR efficiencies. However, NHEJ suppression could have unforeseen side effects. C) Increasing local concentration of donor molecules at DSBs helps in increasing HDR. Streptavidin:Cas9 fusion protein attached to biotinylated donors (1), Cas9 and donor conjugation through SNAP-tag (2) or through a bridging linker RNA (3) make the donors available at the site of DNA break. These strategies improve HDR efficiencies without suppressing endogenous pathways. Tethering approach can be used either with ssODNs or with dsDNA donors.

1.8 Genome editing in *C. elegans*

In the nematode worm *Caenorhabditis elegans* MosSCI has been the go-to technique for introducing transgenes[185,186]. This method has helped in studying the spatiotemporal regulation and expression patterns of proteins throughout the developmental stages. More recently, targeted gene editing techniques such as Zinc Finger Nucleases[8], TALENs and CRISPR/ Cas[39] systems have started to replace the traditional methods owing to their ability to produce precise insertions or deletions of DNA sequences in a targeted manner. Among these techniques, CRISPR genome editing has taken the center stage due to its remarkable simplicity and robustness.

Here I review various techniques and current advancements in *C. elegans* germline genome-editing. Compared to other complex eukaryotes, *C. elegans* is unique because genome editing is performed directly in the germ cells of an adult animal instead of zygotes. This provides a unique advantage of targeting hundreds of oocytes in a single adult. Therefore, in theory every injected animal is capable of producing more than 100 genome edited progeny (average brood size of a WT animals is ~300). At the same time editing *C. elegans* germline also poses several challenges. Maturing germ cells in the adult hermaphrodite gonads contain four sets of each chromosome. However, by the completion of meiosis only one set is contributed to the zygote while three sets end up in polar bodies. Therefore, to achieve high proportion of edited F1 animals 1. The locus of interest on all the four chromosomes of the oocytes should be edited and 2. all

the oocytes that get fertilized should be edited. Hence, it is critical to achieve high cutting efficiencies among germ cells of the injected worms. In this section I will discuss the advancements and the best practices to achieve best editing efficiencies in *C. elegans* germline.

C. elegans germ cells repair Cas9 induced double strand breaks by direct end joining reactions that are mediated by mammalian ligase IV homolog *lig-1* (for HDR see later sections). Initial protocols delivered Cas9 and single-guide RNA (sgRNA) as plasmids[187-190] or as RNA[26,191]. Although these protocols proved that Cas9 is effective at inducing DSBs in *C. elegans* germline, cutting efficiencies were rather low. For example: virtually in all the initial reports less than 5% of the F1 progeny carried indels at the target loci. To circumvent these low cutting efficiencies, positive selection strategies were developed to identify the F1 animals with Cas9 activity[140,192-194]. These strategies involved using a locus that exhibits phenotype with the loss of function alleles as a proxy to screen for the locus of interest (Co-CRISPR) or insertion of a gain of function marker to disrupt the gene function and generate a selectable marker simultaneously. Although these selection strategies reduced laborious downstream genotyping, they did not directly address the poor cutting efficiencies. Low cutting efficiencies could also be due to robust transgene silencing mechanisms in the worm germline[195-197]. High transcription from Cas9 and guideRNA extrachromosomal arrays may induce the small RNA mediated silencing that limits the editing efficiencies.

Findings in other models have shown that purified Cas9 protein and *in vitro* transcribed sgRNA assembled into RNP complexes are more efficient compared to plasmids delivery[88]. Cho et al directly injected Cas9 RNPs into *C. elegans* gonads and demonstrated that this method can be a viable alternative to the nucleic acid based Cas9 delivery. Paix et al adapted this method and demonstrated that high frequency of cutting can be achieved at the *dpy-10* locus. Cas9 RNP usage quickly become the standard to generate DSBs in the worm germline. However, most protocols suffered with lack of reproducibility and consistency. Furthermore, HDR efficiencies remained poor.

In this thesis I aim to address the challenges associated with CRISPR/Cas9 genome editing. Building upon existing protocols we have further improved editing efficiencies and demonstrate the generality of our methodologies in multiple model organisms. The protocols we developed here show remarkable consistency and efficiency at multiple loci in *C. elegans*, human cells and mouse zygotes.

Preface to Chapter II

Contents of this chapter appeared in the following publication

Ghanta et al., Microinjection for Precision Genome Editing in *Caenorhabditis elegans*. STAR Protocols (2021).

Takao Ishidate contributed to Figure 2.2B

Chapter II: Microinjection for Precision Genome Editing in *Caenorhabditis elegans*.

Abstract

In *Caenorhabditis elegans,* targeted genome editing techniques are now routinely used to generate germline edits. The remarkable ease of *C. elegans* germline editing is owed to the syncytial nature of the pachytene ovary which is easily accessed by microinjection. With proper guidance any researcher can learn to genome edit this organism. This protocol is meant to help navigate and troubleshoot the entire genome editing procedure, lowering the barrier to accessing this powerful genetic animal system.

Graphical Abstract

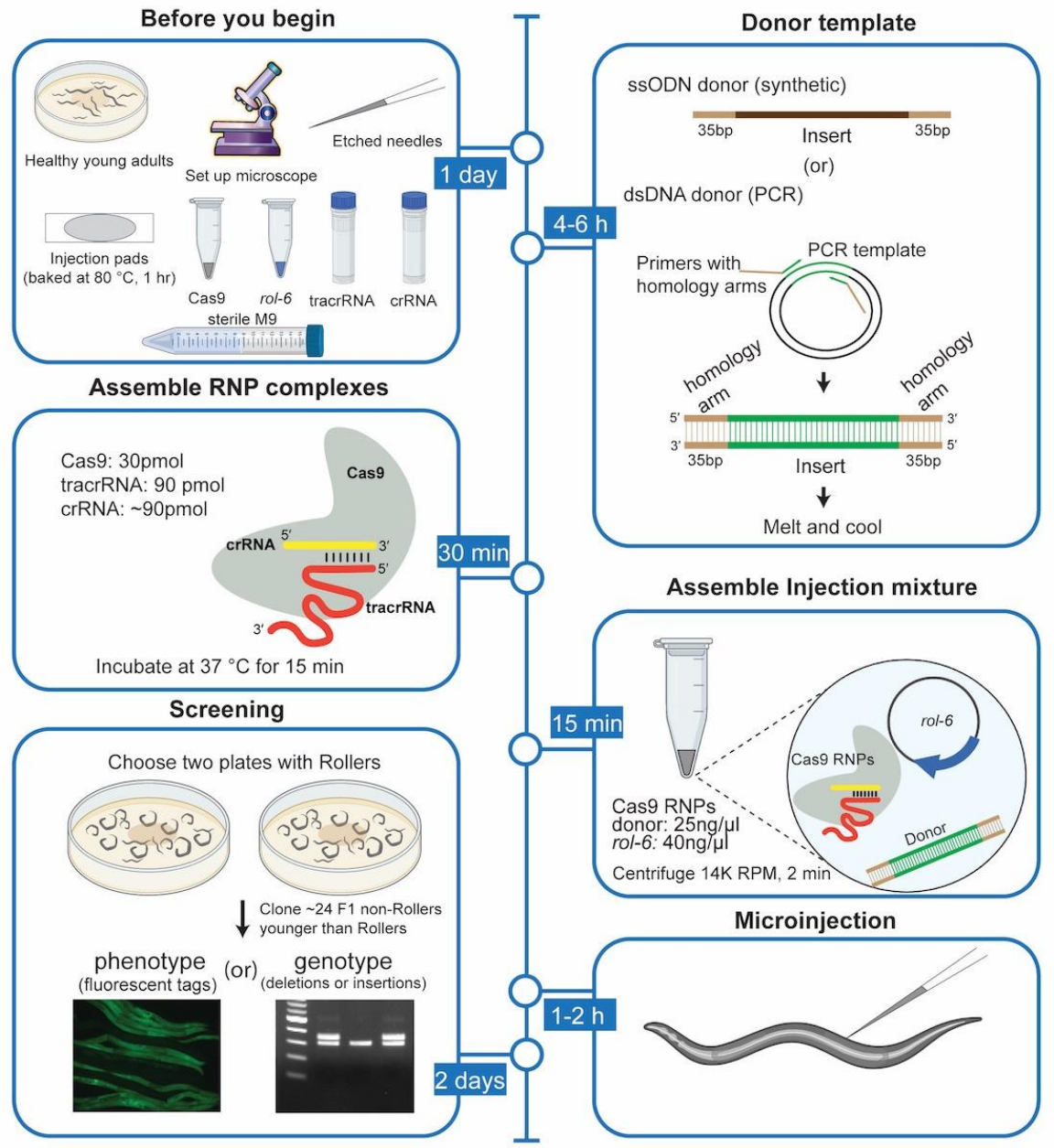


Figure 2.1 Graphical abstract of CRISPR/Cas genome editing protocol

Table 2.1 Key Resources

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Hydrofluoric acid (HF)	Sigma-Aldrich	339261
Halocarbon Oil- 700 CAS#9002-83-9	Sigma-Aldrich	H8898
Agarose	Genesee Scientific	20-102GP
tracrRNA	IDT	Cat# 1072532
TE 7.5 (10 mM Tris, 0.1 mM EDTA)	IDT	Cat#11-01-02-02
Nuclease-Free Duplex Buffer (30 mM HEPES, pH 7.5; 100 mM potassium acetate)	IDT	Cat#11-01-03-01
Polyethylene Glycol 8000 (PEG)	MP	#195445
Q5 polymerase	NEB	Cat# M0491S
S. pyogenes Cas9 3NLS protein	IDT	Cat#1081058
Cas12a protein	IDT	Cat#10001272
Critical Commercial Assays		
Gel extraction kit	Qiagen-Qiaquick	#28706
Experimental Models: Organisms/Strains		
C. elegans N2 strain	Caenorhabditis Genetics Center	N/A
Oligonucleotides		
CRISPR-Cas9 crRNA, 2 nmol	IDT	N/A
A.s. Cas12a crRNA, 2 nmol	IDT	N/A
ssODN donors (ultramer)	IDT	N/A
GFP forward primer	N/A	AGTAAAGGAGAAGAACTT

GFP reverse primer	N/A	TTTGTATAGTTCATCCATGC
Universal linker forward	N/A	TCCGGAGGGAGTGGA
Universal linker reverse	N/A	AGAACCTCCGCCACC
Recombinant DNA		
GFP-linker plasmid	Addgene	N/A
FLAG-TEV-degron-linker plasmid	Addgene	N/A
mCherry-linker plasmid	Addgene	N/A
PRF4::rol-6 (su1006)	N/A	N/A
Other		
Glass capillaries	World Precision Instruments	#1B120F-4
Cover slips 24 x 60 mm No.1	Globe Scientific Inc	#1419-10
Cover slips 22 x 22 mm No.1	Globe Scientific Inc	#1404-10
Mouth pipette -15 In Drummond aspirator tube assembly	Thermo Fisher Scientific	#2118010
PCR purification columns	Qiagen-Minelute	#28604
Ampure XP beads	Beckman Coulter	Ref# A63880
Tygon tubing E-3603 (ID:1/32 in; OD:3/32 in; Wall: 1/32 in)	Saint-Gobain	#00444
Microloader Tips (Femtotips)	Eppendorf	930001007
Dissecting scope	Nikon	SMZ745
Inverted microscope (DIAPHOT 200)	Nikon	Current successor: Eclipse Ti2
Fluorescence dissecting microscope	Zeiss	Axio Zoom. V16

Needle puller Narishige PN-30	Narishige	Current successor: PC-100
Microinjector	Tritech Research	Analog MINJ-1
Micromanipulator	Narishige	MN-151

Before You Begin

Culturing Worms

Timing: 3 days

Worms grown for microinjection should be healthy — fed well, contamination free and under controlled temperature. An approach that ensures ideal worms for injection is to culture broods from individual animals. Animals cultured by chunking the normal growth media (NGM) tend to vary significantly in their quality for the purpose of injections. If chunked worms must be used it is helpful to pick several L3 and L4 hermaphrodites from the chunked plate onto a fresh plate one day prior to injection. Synchronized worms obtained by bleaching gravid adults can also be used, but synchrony limits the variety of stages available from the culture on a given day to a much narrower developmental window than is achieved by brood culture.

The choice of animals may also depend on the needle. A blunt or broken needle can still work for injections into older larger adults but will invariably kill young

adults. When attempting the first injections for a new batch of animals, pay close attention to the size of the gonad cytoplasm. Recently starved adults may appear healthy otherwise but will have a shrunken core gonadal cytoplasm that provides a very poor target for microinjection. Unhealthy animals are a very common reason for unsuccessful injections. Refer to the WormBook chapter for detailed protocols on maintenance of *C. elegans*[198].

Preparation of Injection Pads

Timing: 2 h

Dried 2% agarose pads (also known as injection pads) are used to immobilize the worms during the microinjection procedure[199,200]. If not prepared properly the pads can either dehydrate the worms too quickly and/or worms may not stick to them. High quality injection pads are critical to performing good injections.

1. Prepare 2% (w/v) agarose solution in distilled water by briefly bringing to a boil in a microwave oven. Keep the solution as liquid and well mixed on a stir plate with medium heat (with a stir bar).
2. Use a Pasteur pipette to place two drops (~50 µl each) of hot agarose solution on a 24 x 60 mm glass coverslip and immediately drop another coverslip on top, forming an agarose sandwich, such that the coverslips are parallel to each other.
3. After the flattened agarose solidifies (~5 minutes) remove one coverslip by gently sliding.

4. Place the coverslips with wet agarose facing up on Whatman paper cut to fit on an oven tray for baking. Array the coverslips singly so they do not stack or overlap.
5. Repeat the agarose sandwich process until about 100 pads are arrayed on the Whatman paper for baking.
6. Bake the agarose pads at 80 °C for 1 hour.
7. Once dried, the agarose pads will not stick to each other can be transferred to a 10 cm petri dish and stored together indefinitely. In areas with high humidity, it may be necessary to use a sealed desiccated container, or to re-bake the pads periodically as needed. For convenience, correct side (agarose side) of the slide can be marked with an asymmetric letter before storing in the container.

Preparation of Needles

Timing: 2 h

Here we describe needle preparation method that we routinely use in the lab, other labs may use slightly different techniques. Injection needles are prepared by pulling and etching borosilicate glass capillaries (1.2mm) with inner filaments. The inner glass filament provides capillary action and aids in easier back filling of the needles. Although pre-pulled needles (e.g: Eppendorf Femtotips or Tritech Research: MINJ-PP) can be purchased, they are unnecessary and cost-prohibitive

in our opinion. Needles can be stored and used indefinitely (at least in the order of months) without losing their shape due to glass flow.

8. Pulling the capillaries: Needles suitable for penetrating the outer cuticle of adult worms must be both rigid and sharp. Such needles are easily prepared using an economical needle puller. There are three key adjustment parameters: the “heat” setting, a “gentle pull” setting, and a “strong pull” setting. During the gentle pull the needle gradually lengthens as the glass closest to the filament warms, upon reaching a predetermined length, this slow pull triggers a strong pull that draws the needle out and breaks the filament into two halves. If your device pulls from one end only, the end closer to the filament will have a longer taper, however both needles should be nearly identical at the tip and thus both halves should be saved for use.

a. Adjusting the needle puller is an empirical process that can be greatly aided by keeping a reference needle to use for comparison. Needles should taper to a sharp stiff point. These parameters are attained by adjusting the “heat” and “pull-strength” settings on the puller. In general, lower heat and greater pull strength give stiffer needles. Occasionally, the heating filament is bumped by the blank needle and becomes bent. Use extra care when loading or unloading the needles to avoid this. Examine the first few needles under a dissecting scope comparing their shapes to a reference needle kept for this purpose. The settings we use may not translate to other systems and are only approximate starting values (Heat: 97°C, magnetic sub:22 and

magnetic main: 90 on Narishige PN-30 needle puller). These settings should be optimized for your system.

b. Once the settings are optimized, pull needles and monitor needle shape by eye and by periodically re-checking the needles under the dissecting scope. If the needles begin to deviate in shape or if the puller begins to delay during the slow pull phase or stops pulling altogether, the filament was likely bumped out of alignment. Carefully, re-align the filament, repeat the adjustments to the heat and pull strength, and then continue pulling until you have made enough needles to last for several weeks.

c. Load the pulled needles into a tray or box (with a lid) by propping them up in rows on Time Tape (rolled lengthwise with sticky side out) or along a thin line of rolled clay.

9. Etching Needles: Etching the needles takes less than 1 minute per needle and is worth the added time as it saves hours that will otherwise be lost attempting to brake open needles on the inverted microscope. To etch the needles, pressurize each needle one at a time by loading them onto the microinjection system set to 80 psi. This is easily done by sliding the Tygon tubing from the system's compressed air source directly over the blunt end of the needle. While holding the needle in your hand, apply pressure (using the foot pedal from the injector system). Keep the needle pressurized throughout the test. If the needle tubing interface leaks (emits a hissing sound) it is necessary to more firmly seat the needle.

a. Pipette 50 µl of distilled water near the center of a 35mm plastic petri dish (do not use a glass dish). Pipette a second drop of 15 ul water nearby about a half inch from the larger drop. Carefully, using a plastic pipette tip add 15 ul of a commercially sourced 48% HF solution to the 15µl drop of water to obtain a drop containing 24% HF.

b. Bubble test: Place the petri dish containing the water drop and 24% HF droplets under the dissection scope and carefully, at low magnification, bring the pressurized needle close to the (larger) water drop, pause before touching. If the needle is badly broken, then before the needle touches the water drop you will see the water drop distort from air blown from the needle. Immediately discard such needles. Otherwise, while keeping the needle pressurized, insert the tip of the needle into the water drop. If the needle emits a slow stream of tiny air bubbles, immediately remove it from the water while keeping it pressurized (to prevent water from entering the needle). The needle can then be transferred directly to a final storage tray, with lid (marked as “bubbles in water”). These needles have course openings and though useful are not ideal, unless for example an injection solution contains particulates that are clogging finer tipped needles.

c. If the needle fails to bubble in water, transfer the needle (while keeping it pressurized) into the drop of HF solution. In about one or two seconds it should start to bubble in HF. Immediately transfer back to water, (again maintaining pressure at all times, including during the transfer). Usually, the

needle will not bubble in water. If so, transfer it to a tray labeled “bubbles in HF.” These are ideal needles. If the needle produces a slow stream of bubbles in the second water test, it can be saved in the “bubbles in water” tray. A fast bubble stream in water, indicates a faulty needle that is over etched, and should be discarded. If this happens repeatedly add additional water to further dilute the HF solution.

d. Store the etched needles in a closed container on rolled tape or clay so that the needles are firmly secured.

Critical: Keep the needle pressurized throughout the entire process. If HF enters the needle, it will quickly over-etch and destroy the needle. If water is allowed to enter the needle it will leave residues when it evaporates and could also dilute your injection mixes.

Alternatives: Validating the needles in the bubble tests is a major time saver. When using pre-tested needles, a single needle can be loaded with each injection mix and will almost always yield excellent flow. If untested needles are used, or if a needle becomes clogged while in use, the needle can be opened by physically breaking the tip against the agarose pad or by touching it to the edge of a coverslip.

Note: After several needles are etched it may be necessary to trim a few mm off of the Tygon tubing to remove cracks that prevent a good seal. Etching can also be done by loading each needle into a needle holder that creates a seal with a screw lock and gaskets. However, we find this takes more time than using the simpler direct tubing interface.

Critical: HF is highly corrosive and in large quantities is very dangerous, handle the solution with great care and follow all product safety guidelines. Appropriate personal protection equipment such as eye googles, lab coat and gloves should be used. Purchase the 48% HF solution in a small quantity. If possible, purchase HF in an anti-spill bottle that permits micropipetting from a small reservoir in the lid.

Preparation of stock solutions

Note: proceed to either step 10 or 11 based on the Cas protein used in this protocol.

Timing: 30 min

10. **If using Cas9, prepare Cas9 RNPs**

    a) Cas9 protein- aliquot 0.5 µl (5 µg or 30 pmol) either in PCR tubes or 1.5 ml tubes and store at -80 °C (avoid freeze/thaw cycles)

    b) tracrRNA – 0.4 µg/µl (18 µM) in IDT nuclease free duplex buffer, store at -20 °C (store 30 µl aliquots at -80 °C for long term storage)

    c) crRNA – 0.4 µg/µl (34 µM) in TE pH 7.5, store at -20°C (store 10 µl aliquots at -80 °C for long term storage)

11. **If using Cas12a, prepare Cas12a RNPs**

    a) Cas12a protein- aliquot 0.5 µl (5 µg or 32 pmol) and store at -80 °C (avoid freeze/thaw cycles)

b) Cas12a-crRNA – 40 µM in TE pH 7.5, store at -20 °C (store 10 µl aliquots at -80 °C for long term storage)

12. ssDNA oligo donor – 1 µg/µl in nuclease free water, store at -20 °C
13. PRF4::*rol-6 (su1006)* plasmid (midi prep)*:* 500 ng/µl working solution in nuclease free water, store at -20°C.

Notes: Commercial Cas9 protein can be diluted 1:1 with 1x PBS and stored as 2.5 µg (0.5 µl) aliquots to reduce the total injection volume to 10 µl. We chose to reconstitute tracrRNA in Duplex buffer to provide salts to the Cas9 RNP complexes. We have not explored the long-term stability and cutting efficiencies of the injection mixtures prepared with tracrRNA reconstituted in other solvents such as TE or water. We have also not explored replacing TE or water with nuclease free buffer in the injection mixture. Aliquots of Cas9 protein (-80 °C) can be stored in -80 °C for at least for a year without compromising efficiencies. crRNAs and tracrRNA can be stored at -80 °C for at least two years and at -20 °C for at least a year without compromising efficiencies. Refer to previously published protocols if Cas9 and guide RNA plasmid-based approach is used[201,202].

Guide RNA Design

Depending on the GC content and the availability of Protospacer Adjacent Motif (PAM) sites either Cas9 (PAM: NGG) or Cas12a (Cpf1) (PAM: TTTV) can be used as nuclease to introduce double strand breaks. In our experience Cas9 guide

RNAs (gRNA) with low GC content perform poorly. It may be more efficient to use Cas12a[203,204] at those AT-rich target sites.

To generate loss of function deletions, choose two gRNA target sites such that majority of the gene is deleted. Bioinformatics tools such as CRISPOR (http://crispor.tefor.net/) or CHOPCHOP (https://chopchop.cbu.uib.no)[205] help in identifying gRNAs that have no off-target cleavage activity[206]. When the choice of guide RNA is more restricted the CRISPOR tool can be used to identify potential off-target sites that can be screened for later and/or crossed out of the strains. If using *C. elegans* strains other than N2, polymorphisms specific to your strains need to be considered for gRNA design and off-target assessment.

Templated precise editing- Donor Design

Although non-homologous end joining (NHEJ) has been shown to be the dominant type of repair mechanism in many cell types, this is not the case in the pachytene germline of *C. elegans*. When using properly prepared donor molecules, up to two thirds of the post injection progeny will exhibit homology-directed repair[203,207]. Depending on the length of the desired edit, one can choose to use short single stranded oligodeoxynuleotides (ssODN)[140,160,194,208-211] or linear double stranded DNA (dsDNA)[203,208,210-213] as donor templates. Refer to[203,207] for HDR efficiencies achieved using the protocols described here.

14. **ssODN Donor Design**

ssODNs up to 200 nt can be obtained commercially at reasonable prices. Therefore, any edit that is shorter than 130bp can easily be obtained using a synthetic ssODN donor (35 nt homology arms)[208,211].

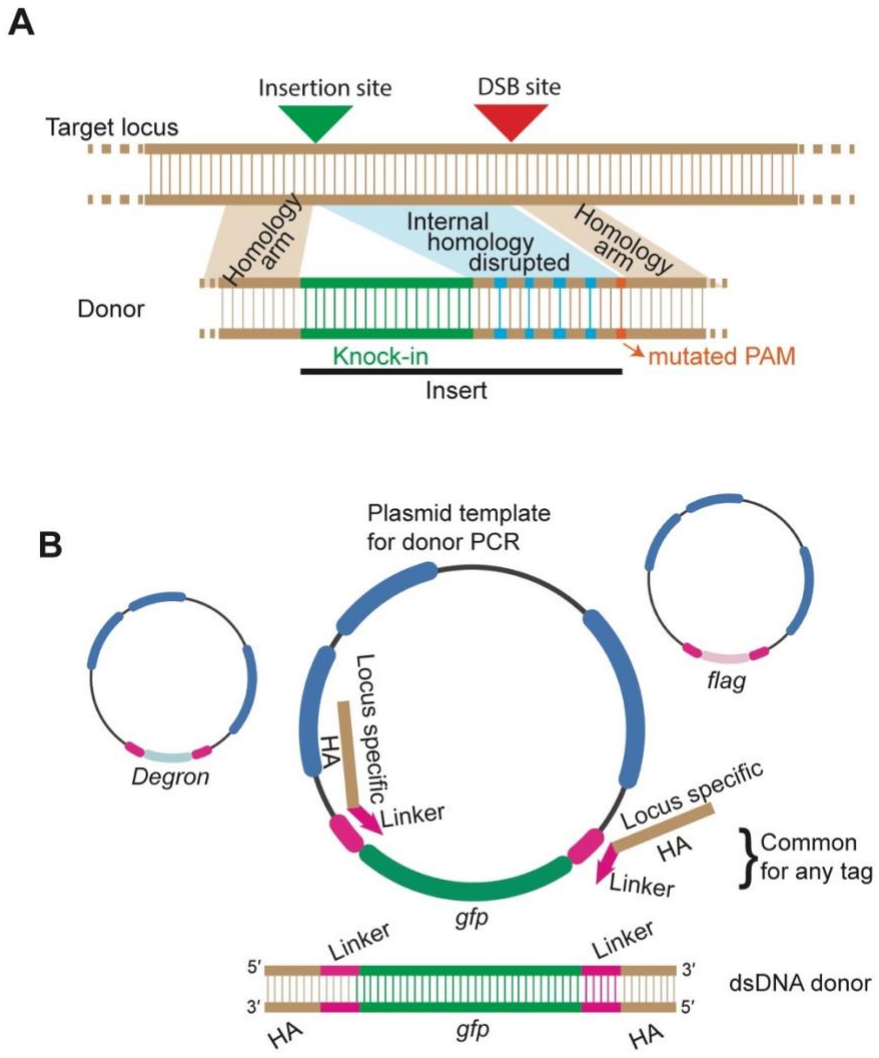


Figure 2.2 Donor design strategy.

(A) If the site of insertion of a tag (knock-in) and the site of double strand break are not close to each other, disrupting the internal homology between these two sites

increases the efficiency of precise repair. Sequence with disrupted homology should be considered as part of the insert and the homology arm should begin after the last silent mutation. (B) Schematic design for a set of locus-specific homology arms (HA) with PCR primers which also serve as linkers between the tag and the protein of interest. For a given locus, the same set of oligos can be used to generate dsDNA donors to knock-in any tag. Plasmid containing the tag flanked by linkers is used as PCR template. Diagrams are not drawn to scale.

15. **To generate short inserts (<130bp)**

Add 5′ homology sequence (~35nt) in front of the tag (or mutations) and 3′ homology sequence (~35nt) at the end. To prevent recutting of the edited allele, introduce silent mutations into the PAM site or into the guide binding sequence in the donor molecule if it is not already split by the insert. Three or four silent mutations in the PAM proximal region should suffice and 35nt after the final silent mutation should be used as homology arm (Figure 2.2 A).

16**. To generate point mutations:** Choose 35nt sequence upstream and 35nt sequence downstream of the DSB site as homology arms for the donor. Introduce the desired nucleotide changes in the donor and the PAM or guide binding sequences. Silent mutations can be introduced to create restriction sites to easily genotype precisely edited progeny of the injected animals.

17**. To generate large deletions:** Use two guides that remove majority of the target locus (for example: to precisely remove a structural domain or an entire coding region). Order an ssODN that contains 35nt of sequence upstream of the left-guide DSB and 35nt downstream of the right DSB to act as a donor. You may wish to incorporate a new gRNA binding site inside the break points on this donor so that you can re-cut the locus for any downstream purposes. Deletions of 1kb are routinely achieved with this strategy but it can also work for larger deletions. Donor can be omitted if precise deletion is not necessary.

Critical: For both ssODNs and dsDNA donors, if the cut site is not ideally situated, for example is offset by several nucleotides from the desired point of donor insertion, it is critical to mutate the intervening nucleotides, for example by recoding amino acids. Minimizing sequence identity in the donor internal to the desired insertion prevents partial insertions due to template switching (Figure 2.2 A)[208,211].

18. **dsDNA Donor design and generation**

Timing: 4–6 h

Generate dsDNA donor templates using unmodified or 5′ Triethylene Glycol (TEG) modified primers. TEG modification is also known as spacer 9 (SP9). Melting the dsDNA donors is critical to achieve high HDR efficiencies[203]. Using end-modified donors (and melting) further increases HDR efficiencies[207].

a. Synthesize unmodified (or 5′ SP9 modified) PCR primers with standard desalting (IDT); 35nt as homology arms and 20nt complementary to the desired insert (e.g., *gfp*). SP9 modifications are available at 100nmol scale from IDT.
b. Perform PCRs to amplify the insert (e.g: *gfp*) from a plasmid using High-Fidelity polymerase.
c. Perform agarose gel electrophoresis to ensure that a single specific band is obtained. If non-specific amplification is observed, set up a temperature gradient (50 °C to 64 °C) and find the optimal annealing temperature.

d. Purification of donor DNA: Use one of the following methods depending on your experimental conditions.

i. Perform PCR purification using spin-columns (e.g., Qiaquick Minelute) and elute DNA in 15 to 20 µl of nuclease free water. Generally, column purification is sufficient, and you may proceed to step (18.e) However, some primer (55nt long) pairs produce long dimers that may contain the homology arms. Spin-columns may not be able to remove long (>80bp) dimers completely. These short "dimer donors" out compete the full-length donors (e.g., *gfp*) and integrate heavily into the DSB site.

Note: Dimers may or may not be visible on the agarose gel. It is helpful to use a primer analysis tool (e.g., OligoAnalyzer-IDT) to predict homo- or hetero-dimer formation.

ii. If primer-dimer formation is a concern, perform PCR clean-up using 0.6X SPRI paramagnetic beads (AMPure XP) to size-exclude shorter (<200bp) products. For example: incubate 100 µl of PCR product with 60 µl of beads, wash the beads with 70% ethanol twice and elute with nuclease free water (refer to the manufacturer's protocol for detailed procedure).

iii. Gel-extract the donor DNA if primer dimers are clearly visible on agarose gel. However, gel extracted DNA can be toxic due to the presence of guanidine hydrochloride (component of binding buffer)

in the final elute. Strong absorbance at 230nm on Nanodrop suggests presence of GuHCl in the elute. During the wash steps, incubating the spin-columns with wash buffer for 10 min before centrifugation helps in reducing salt contamination; repeat the washes 2 to 3 times. For best results, gel-extracted DNA should be further purified with 1X to 1.5X AMPure-XP beads (strongly recommended).

e. Dilute a portion of the purified donor to 100 ng/µl and transfer about 5.5 µl (for 20 µl injection mixture) to a PCR strip tube and proceed to the melting step.

f. "Melt" the donor

Temperature	**Time**
95 °C	2 min
85 °C	10 sec
75 °C	10 sec
65 °C	10 sec
55 °C	1 min
45 °C	30 sec
35 °C	10 sec
25 °C	10 sec
4 °C	hold

Ramp down at 1 °C/sec at every step

Notes: we store purified donors at -20 °C and melt them right before adding to the injection mix. We have not explored storage and re-use of melted donors. We have only tested the editing efficiencies with donors melted at 50-150 ng/µl concentrations. Melting donors at concentrations that are too high or too low may negatively influence editing efficiencies.

Alternatives: If the goal is to insert a common tag such as FLAG, GFP, degron etc. then it is a good idea to order a set of homology arms with "universal" PCR primers which also serve as linkers between the tag and the protein of interest. A single set of gene-specific homology arm primers can then be used to insert any tag into your locus of interest (Figure 2.2 B). It is almost always more cost effective to order universal primers, even if inserting a short tag such as FLAG could be achieved with ssODNs. Any gene you decide to study will be tagged in multiple different ways during the course of your work, and universal primers will save both time and money. Plasmids containing a set of commonly used tags flanked by these linkers are available from Addgene (see Key Resource Table 2.1).

Materials and Equipment

Reaction setup:

Reagent	**Amount**
5× Q5 buffer	10 µL
10mM dNTPs	1 µL
10 µM Forward Primer	2.5 µL

10 µM Reverse Primer	2.5 µL
Q5 polymerase	0.5 µL
Template plasmid (5 ng)	1.0 µL
Nuclease-free water	33.5 µL
Total	**50 µL**

Thermocycling conditions:

Step	Temperature	Time
Initial denaturation	98°C	1 min
Amplification (30 cycles)	98°C	10 sec
	64°C	20 sec
	72°C	40 sec
Final extension	72°C	3 min
	4°C	Hold

If necessary, find optimal annealing temperature by performing a thermal gradient (50°C to 64°C).

While this protocol has focused on GFP knock-in with the microinjection equipment detailed above, the general principles of the procedure described here can be used for inserts of comparable size using other equipment.

Step-by-Step method details

Note: if using Cas9, follow steps 1-8; if using Cas12a, follow steps 9-16.

Genome editing using Cas9 nuclease

Add components of the injection mixture to the tube containing Cas9 in the following order:

1. Prepare Cas9 – 0.5 µl of 10 µg/µl stock (30 pmol)
2. Add tracrRNA – 5 µl of 0.4 µg/µl stock (90 pmol)
3. Add crRNA – 2.8 µl of 0.4 µg/µl stock (95 pmol) (if two guides are needed add 1.4 µl of each)
4. Pipette the mixture gently several times and incubate at 37 °C for 15 minutes.
5. Add ssODN donor – 2.2 µl of 1 µg/µl stock (or)

   Add melted dsDNA – 500 ng (final concentration: 25 ng/µl for ~1kb donors or 45 fmol/µl)
6. Add PRF4::*rol-6 (su1006)* plasmid – 1.6 µl of 500 ng/µl solution
7. Add nuclease free water to bring the final volume to 20 µl and pipette gently several times.
8. To avoid needle clogging, centrifuge the mixture 13000 x g for 2 min, transfer about 17 µl of the mixture to a fresh tube and keep the tube on ice; proceed to loading the needles.

Genome Editing using Cas12a nuclease

Add components of the injection mixture to the tube containing Cas12a in the following sequence:

9. Cas12a – 0.5 µl of 10 µg/µl stock (32 pmol)
10. Add Cas12a-crRNA – 2.5 µl of 40 µM stock (100pmol)
11. Add TE pH 7.5 – 3.0 µl
12. Pipette the mixture gently several times and incubate at 37 °C for 15 minutes
13. Add ssODN donor – 2.2 µl of 1 µg/µl stock (or)

    Add melted dsDNA – 500 ng (final concentration: 25 ng/µl for ~1kb donors or 45 fmol/µl)
14. Add PRF4::*rol-6 (su1006)* plasmid – 1.6 µl of 500 ng/µl stock
15. Add nuclease free water to bring the final volume to 20 µl and pipette gently several times.
16. To avoid needle clogging, centrifuge the mixture at 13000 x g for 2 min, transfer about 17 µl of the mixture to a fresh tube and keep the tube on ice; proceed to loading the needles.

Notes:

Injection mixtures can be prepared at room temperature

1:3 molar ratio of nuclease to guide RNA is used to saturate the nuclease. We have not explored the cutting efficiencies with other molar ratios.

Protein aggregation or precipitation is not an issue at these Cas9 RNP concentrations.

TE is added in step 11 of Cas12a mixtures for easier pipetting. This step can be omitted by further reducing the concentration of crRNA stock.

Although we haven't explored the optimal dose range for ssODNs, given the efficiencies obtained with dsDNA at 25 ng/µl, much lower doses of ssODN could be used[211] *.*

Final injection mixture can be stored at 4°C and re-used for more than a year without compromising efficiencies. We have injected 16 months old mixes and confirmed that HDR efficiencies are high.

In our experience adding any double stranded DNA before RNP complex formation reduces HDR efficiency.

Loading needles and preparing the microscope

Timing: 15 min

17. Using microloader tips, load 1.5 µl of the injection mixture into an etched needle.
18. Leave the needle pointing tip downward for several minutes to allow air bubbles to work their way out of the tip. This can be done easily by sticking the needle to double sided tape or clay on a vertical surface.
19. Before mounting your needle onto the micro-manipulator make sure to align and adjust the microscope. To do this place a worm on an injection pad under halocarbon oil. Bring the worm into focus under the 10X objective,

then close down the condenser iris and make sure the shadow of the iris is centered in the field of view and sharply in focus. Also adjust the Nomarski optics and ensure that you have the proper filters and polarizers in the light path. Consult your manual for appropriate steps to focus the condenser and adjust the microscope. These adjustments are essential and must be done properly or the injections will be difficult if not impossible.

20. Insert the needle into the tube connected to the pressure regulator and lock it in place on the needle holder, then position the holder downward (~15°) relative to the stage. Position the needle using the Z-axis adjustment such that it doesn't touch the injection pad or the condenser when the microscope head is lowered.
21. With the light on and the condenser iris closed down partially, position the needle using the coarse lateral adjustments so that the needle tip is illuminated in the center of the light. If your manipulator is mounted to the condenser arm, these adjustments can be made while the arm is tilted back away from the stage. Do not look through the eyepiece when making these coarse adjustments. Instead, look directly at the needle tip as you move the lateral and longitudinal adjustments on the manipulator. First, with the needle tip pushed past the center of the light path, move it forward or backward with the coarse adjustments until it is illuminated. You will see a spark of light when the needle is in the light path (close down the condenser iris to further pin-point the needle, via this spark of light, in the center of the

field of view). Next, move the needle laterally (to the right if using a right-handed system) until the spark of light disappears, and then move it back toward the left until it lights again. Now raise the needle up close to the condenser to ensure ample working distance when you lower the condenser arm. Assuming you successfully aligned your microscope light source as described in step 19 above, your needle tip is now in the center of the field of view and will only need to be lowered in order to bring it into focus when needed.

22. Using the worm loaded earlier as a reference, focus on the worm using the 10X objective and then while watching the needle directly (not through the eyepiece) use the Z-axis adjustment to lower the needle until it touches the oil above the worm.
23. Look through the eyepiece and slowly center and bring the tip of needle into focus using X-, Y- and Z-axis adjustments on the micromanipulator. Once the tip is in focus, do not change the positions of X- or Y- axis.
24. Using 40X objective, examine the needle tip and apply pressure continuously for one or two seconds. One can judge the size of the opening (hence the quality of the needle) based on the diameter of the bubble produced in this interval. The bubble should be about the diameter of the gonad after one second. Adjust the pressure to the needle as needed. Pressures in the range of 40 to 80 psi are typically used during injection.

Mounting Worms

Timing: 1 min

Worms move by swimming in a thin aqueous layer that surrounds their body surface. Agarose pads immobilize the animals by drying this layer of moisture, sticking the animals to the pad surface. If the pad is too thick it will quickly desiccate the animal, if too thin, the animal will crawl away. Even high-quality injection pads will dry the worms out after a few minutes, so do not mount more worms than you can easily inject in a 10-to-15-minute interval. With experience mounting each worm should take less than one minute.

25. Determine the upper side of the injection pad by scratching an edge of the dried agarose with a razor blade or other metal object. This scratch mark can also be used for breaking needles later if needed. Mark on the agarose side of the pad can also be used if the slide is marked with an asymmetric letter.
26. If desired, you can mark the boundaries of the agarose with a thin marker on either surface of the coverslip. Since the dried agarose is virtually transparent marking the boundaries helps in placing the worms on the agarose.
27. Use a Pasteur Pipette to transfer a small drop of halocarbon oil onto the injection pad away from the agarose (Figure 2.3). Use surface tension (not suction) on the pipette to move the oil to the injection pad. For working under the dissecting microscope place the injection pad on an inverted lid of petri dish for easier handling.
28. Touch your worm pick to the surface of the halocarbon drop, and then use the viscous oil on the bottom of your pick to collect one or a few worms from your

culture plate using the dissecting microscope. It is usually possible to choose worms that are off the bacterial lawn, but it may be necessary to move the worms briefly to an unseeded plate prior to picking them up with the halocarbon oil.

Alternatives: Excess bacteria can also be removed from the worms after transferring them to the large halocarbon oil droplet on the injection pad. After depositing the worms into the large oil droplet, flame the pick and return to the oil drop. First quench the pick in the oil away from the worms and then move the pick underneath the worms and then quickly pull the worms through the oil to remove bacteria. Once the desired number of animals are floating in the oil drop and free of bacteria, select the worms for injection by placing the pick under them in the oil and moving them by lifting them on the pick to a dry area of the pad. Depending on your abilities and your experimental needs you can transfer either one (recommended for beginners), or up to several animals onto your pad for the first round of injections.

29. Touch the pick down to the surface and move it laterally to deposit the worms in a row one after another (Figure 2.3). They should float off the pick in the oil. Carefully orient the worms with vulva and row of eggs facing to the left (if a right-handed system is being used). Gently push the worm down with the worm-pick by rubbing the pick along the body of the animal. If the animal is mostly stuck, then move onto the next animal. Make sure all the animals remain stuck and properly oriented. If possible, keep the worms within about 2-3 mm of each

other so that you can move from one to the next without difficulty while working on the inverted microscope (Figure 2.3). Finally, after the worms are stuck down dip your pick back into the large oil drop on the edge of the pad, then use your pick to deposit additional oil to fully cover the immobilized worms. This additional oil will prevent the worms from desiccating during the procedure.

Alternatives: If you accidently orient some worms incorrectly you can either reorient them by gently pushing the worm off the pad with your pick and re-sticking them, or you can inject the worms oriented one way and then flip the slide to inject the ones facing the other way.

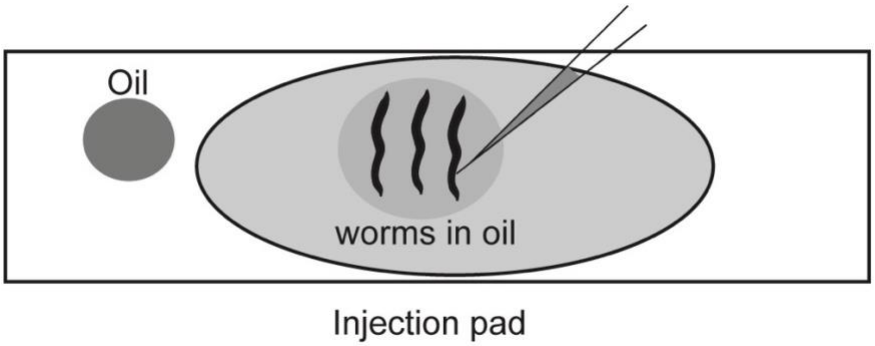


Figure 2.3 Mounting worms.

Use halocarbon oil to mount worms onto the injection pad. Align all the pachytene zones (sites of injection) of all the gonad arms towards the needle. Diagrams are not drawn to scale.

Microinjection and Worm Recovery

Timing: 30 to 60 min

Microinjection into syncytium of hermaphrodite gonad[214,215] delivers Cas9-guideRNA complexes into dozens of germ nuclei simultaneously. Aim to inject the animals in the mid pachytene region. Injecting either in the distal or in the proximal end of the gonad is not ideal. Since the number of embryos produced by wildtype animals is limited by the availability of sperm, germ cells present in the distal region of the gonad arms do not get fertilized. Therefore, you should target the germ cells in the mid to late pachytene region. It should take approximately one minute to inject both gonad arms of each worm, and another minute or two to remove the injected animals to culture plates.

30. Focus the objective (40X) onto the gonad (pachytene region) so that you can see two rows of nuclei (as shown in Figure 2.4 A). This focal plane ensures that when inserted, the tip of the needle enters the core of the gonad and not above or below the somatic gonad. Do not focus on the honeycomb structure of the germ nuclei.
31. Without changing the focus and while looking through the eyepiece, slowly lower the needle into the oil drop. By adjusting the position of the stage bring the cuticle (at the pachytene region) very close to the needle tip (Figure 2.4 A).
32. Gently tap the micromanipulator so that the needle penetrates the cuticle and the somatic gonad, and enters the core germplasm. Gently move the stage

towards the needle to adjust the position of the tip in the gonad. Do not insert the needle too far into the gonad as it ruptures the cuticle.

33. Press the foot pedal to inject the mixture into the germplasm and make sure that the injection mixture flows smoothly onto either side of the needle tip over a period of two seconds. You should see the fluid entering the proximal region of the gonad and pushing the oocytes towards the spermatheca (Figure 2.4 B). If the flow is excessive the worm or gonad can explode, reduce the pressure to achieve a lower flow rate. Inject into both arms of the gonad.
34. Raise the needle to remove the injection pad. Move the pad onto the inverted lid under the dissecting scope.
35. Recover the injected worms using mouth pipette (pulled glass capillary attached to Drummond aspirator tube) and sterile M9 buffer or MPEG (0.5% PEG in M9). Occasionally, worms stick to the glass pipette during the recovery process, and they are less likely to stick when MPEG is used. Other buffers are not necessary. Avoid using plastic pipette tips to recover the injected worms because they will stick to the tips.
36. If the injections differed in quality, for example if only one arm was injected, these differences should be noted, and the worms separated NGM plates labeled according to injection quality. When all the worms are injected place each injected animal singly onto a fresh NGM plate and culture for about 3.5 days at room temperature (22 °C to 23 °C) unless genetic or experimental conditions demand other temperature conditions.

Screening and Genotyping

Timing: 2 days

If both arms of the hermaphrodite gonad are injected, a good injection should yield 20 to 40 F1 Rollers. Screening strategy depends on the type of edit generated.

37. About 72 hours post injection, score the number of F1 Rollers and choose two plates with the highest number. The Rollers provide an important trouble-shooting metric (see below) and should always be included whenever possible.

    a. For indels (without donor template) or ssODN donors, choose two P0 plates that segregate the highest number of F1 Rollers; pick 12-24 young adult F1 Rollers onto separate plates and allow them to have progeny prior to genotyping by PCR.

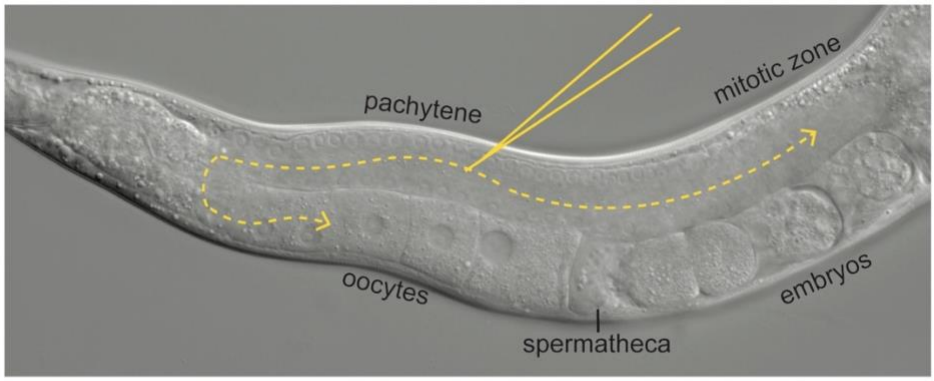
    b. For long dsDNA donors screening should be performed differently depending on the quality of the injections and the type of edit desired. If excellent bilateral injections were performed resulting in greater than 20 Roller progeny per injected animal, then choose two plates that segregate the highest number of F1 Rollers. If a fluorescent protein tag was inserted a quick look under the compound fluorescence microscope (or a dissecting Fluorescence microscope if the fluorescence is expected to be bright) will allow you to gauge how many F1 progeny to pick. To check under the compound microscope (preferred for genes whose expression levels are low or unknown) place ~50 gravid adult F1

progeny younger than the rollers onto 2% wet agarose pads on a microscope slide (do not use dry injection pads) and immobilize with levamisole, then push the worms into proximity with the pick, gently place a cover slip on the top and check for fluorescent signal. If you wish you can memorize the location of positive animals and while working on a dissecting scope push the cover slip aside (do not flick it open) to directly recover the fluorescent positive animals. If PCR genotyping is the desired approach for detection, choose about 24 non-Rollers that are younger than the Rollers and place them onto separate NGM plates. Younger animals among Rollers can also be picked. For inexperienced injectors, we recommend using 5′ end-modified dsDNA donors and picking F1 Rollers.

38. Genotype the F1 adults directly after allowing them to lay progeny. Alternatively, to avoid PCR false positives due to mosaicism, lyse several F2s from each plate and genotype. At least one of the genotyping primers should lie outside of the homology arms to avoid amplification from transiently retained donor molecules.

    Note: Strains obtained by genome editing should be back crossed to eliminate off-target indels. However, off-target indels on the same chromosome and close to the target site may be linked to the target site. Therefore, generating at least two independent alleles using different gRNAs is recommended.

A



B

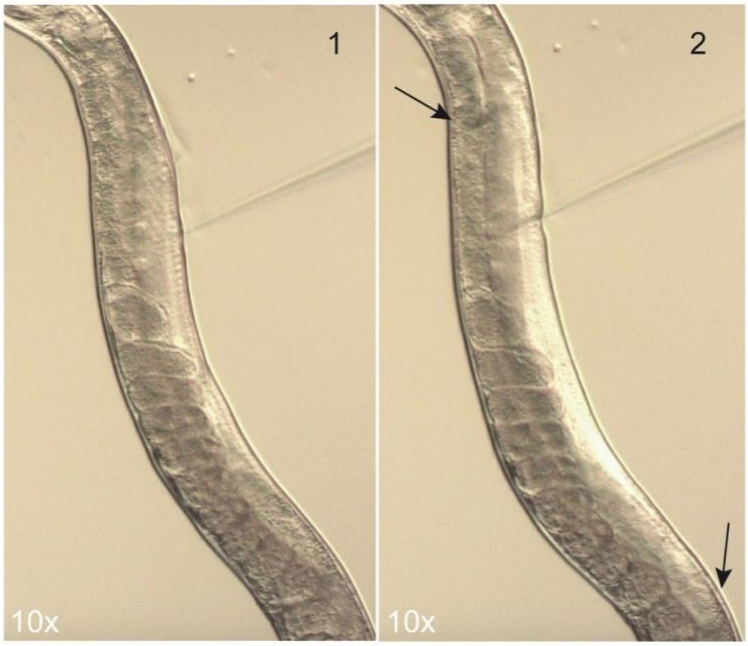


Figure 2.4 Microinjection.

(A) Using the 40X objective, focus on the two rows of germline nuclei as shown and insert the needle into the pachytene zone. Flow of the injection mixtures should reach mature oocytes in the proximal end of the germline as shown with yellow dotted line.

(B) Illustration of the flow of injection mixture in germline is shown (1) before and (2) after injection. Arrows marks the region in the gonad until where the mixture has reached. 10X magnification is shown for illustration purposes, use 40X objective for the microinjection procedure.

Expected outcomes

This protocol is used to generate precise genome edits in *C. eleg*ans germline. Loss of function alleles, point mutants and fluorescent tags can be generated with high efficiencies. A single injected animal can produce precision edits in among as many as two thirds of progeny from the post-injection cohort[203].These advances increase the utility of *C. elegans* for understanding fundamental questions of animal biology and for investigating the physiology of human alleles implicated in disease.

Limitations

Obtaining high *knock-in* efficiencies of large inserts (for example, 3kb or more) is still a challenge. Future improvements in protocols and development of new strategies will address this limitation.

Troubleshooting

Problem 1. Worms mounted onto injection pads dry immediately (step 29)

Potential solution:

- Ensure that the agarose solution is made with water. Beginners may mistakenly use DNA gel running buffers instead of water.
- Ensure that the pads are thoroughly dried by baking and that they are not too thick
- Animals that are too young tend to dry much faster
- Ensure that the animals have never starved and are well fed

Problem 2. Worms don’t stick to the injection pads (step 29)

Potential solution:

- Injection pads may not be dry: Re-bake at 80 °C for an hour
- Too much bacteria: make sure that bacteria is not transferred onto the injection pad

Problem 3. Needles clog during injections (step 33)

Potential solution:

- Dust particles in the injection mixture clog the needles. Make sure that the mixtures are spun down at high speed for at least 2 minutes prior to loading into the needles
- High concentrations of protein could clog the needles. Ensure that the correct concentrations are used in the injection mixture
- Needle opening may be too small: gently touch the tip of the needle to the agarose to break it open

Problem 4. Needle opening is either too wide or too narrow, or tip is too flexible (bends when pushed against cuticle (step 32)).

Potential solution:

- Adjust the settings of the needle puller to obtain tapered stiff needles
- Adjust the contact time of the needles in HF

Problem 5. Low survival rate of the injected animals (step 35)

Potential solution:

- Choose young adults for injection
- Blunt needles with wide openings could severely damage the cuticle
- Injected worms should be recovered with care using M9 buffer

Problem 6. Small brood sizes after injection (step 37)

Potential solution:

- Small brood sizes are observed if the target locus of the guide RNA is essential for viability
- Make sure that the donor DNA is free of contaminants such as high salts and the concentrations of all the components are optimal
- Use young adults for injections. Old animals do not yield many progeny
- Do not inject excess amounts of the mixture. If flow is excessive, reduce the pressure to lower the flow rate

Problem 7. No Rollers (step 37)

Potential solution:

- Check the integrity of the *rol-6* marker plasmid on agarose gel. Purify the plasmid DNA to remove contaminants.
- Use appropriate concentrations of the Cas9 RNPs, donor DNA and *Rol-6* plasmid as suggested in the protocol. Excess amounts of Cas9 RNPs or DNA is toxic.

- For beginners, practicing microinjection procedure by injecting Prf-4: Rol-6 plasmid alone can be helpful. More than 40 F1 Rollers can be obtained from each injected animal.
- If loss of function mutations of the target locus impair gamete or embryo viability the number of F1 Rollers produced will be dramatically low.

Problem 8. No indels (step 38)

Potential solution:

- AT rich Cas9 guide RNAs can be inefficient. If the target sequence is AT rich, switch to Cas12a based editing.
- Use appropriate concentrations of Cas9 and crRNAs as suggested in the protocol. We recommend using a final concentration of 1.5 µM Cas9 RNPs in the mixture. Increasing the concentration to 3 µM does not increase cutting efficiency[208].

Problem 9. Indels at homologous locus (step 38)

Potential solution:

- If an identical target site exists in another locus, majority of the F1 progeny may have indels at both the loci. Use a different guide to avoid indels at the homologous locus.
- If the guide choice is limited, reduce Cas9 RNP concentration by five-fold. Overall indel efficiencies may drop but single mutants can be easily isolated.

Problem 10. No GFP insertions (steps 37 and 38)

Potential solution:

- Make sure that the dsDNA donor is melted before adding to the Cas9 RNP mixture
- Donor DNA should be free of any primer dimer or smaller DNA species that can potentially inhibit HDR efficiencies
- It is difficult to obtain HDR insertions with inefficient guide RNA. Genotype and sanger sequence the target locus in about 12-24 Rollers and ensure that cutting efficiency of the guide is high. Alternatively, TIDE analysis on pooled F1 genomic DNA can be performed[208,216]. Also, rule out partial insertion of the donor (e.g., due to primer dimer insertion).
- If the expression pattern of the protein is unknown it may be helpful to perform fluorescence-based screens with animals of mixed stages including older embryos on the plates.
- If the initial screening was performed by fluorescence microscopy, genotype the locus by PCR and sequencing the insertion sites. Low expression levels may be the reason for “No GFP”.

Acknowledgments

We thank the members of the Mello lab and the worm community for sharing their observations and difficulties which helped in making the protocols better. This work was funded by NIH grant R37GM058800 to C.C.M. C.C.M is a Howard Hughes

Medical Investigator. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440).

Author Contributions

K.S.G and C.C.M wrote the manuscript. K.S.G, T.I and C.C.M performed the experiments. C.C.M supervised the study.

Declaration of Interests

Some of the findings described here are part of the patent applications filed by the University of Massachusetts Medical School with K.S.G and C.C.M as co-inventors. C.C.M is a scientific founder and advisory board member of CRISPR therapeutics.

Preface to Chapter III

Contents of this chapter appeared in the following publication

Robust genome editing with short single-stranded and long, partially single-stranded DNA donors in *C. elegans.* Gregoriy A. Dokshin*, **Krishna S. Ghanta***, Katherine M. Piscopo, and Craig C. Mello. (*co-first author)

I contributed to Figures 3.1- 3.4

Gregoriy Dokshin contributed to Figures 3.1- 3.4

Katherine M. Piscopo contributed to figure 3.2C

Chapter III: Robust Genome Editing with Short Single-Stranded and Long, Partially Single-Stranded DNA Donors in *Caenorhabditis elegans*

Abstract

CRISPR-based genome editing using ribonucleoprotein (RNP) complexes and synthetic single stranded oligodeoxynucleotide (ssODN) donors can be highly effective. However, reproducibility can vary, and precise, targeted integration of longer constructs – such as green fluorescent protein (GFP) tags remains challenging in many systems. Here we describe a streamlined and optimized editing protocol for the nematode *C. elegans*. We demonstrate its efficacy, flexibility, and cost-effectiveness by affinity-tagging all twelve of the Worm-specific Argonaute (WAGO) proteins in *C. elegans* using ssODN donors. In addition, we describe a novel PCR-based partially single-stranded "hybrid" donor design that yields high efficiency editing with large (kilobase-scale) constructs. We use these hybrid donors to introduce fluorescent protein tags into multiple loci achieving editing efficiencies that approach those previously obtained only with much shorter ssODN donors. The principals and strategies described here are likely to translate to other systems and should allow researchers to reproducibly and efficiently obtain both long and short precision genome edits.

Introduction

In theory, CRISPR/Cas9-based genome editing enables researchers to rapidly generate designer alleles of any locus for genetic, cytological, or biochemical analyses. In practice, however, we have found that the technology remains far from routine for many users, especially in applications where long templated insertions are desired. Here we explore the basic principles behind a robust editing pipeline. We demonstrate pronounced toxicity of RNP at high concentrations and provide a strategy for optimizing RNP levels using a co-injected, easily scored reporter. Finally, we show that generating hybrid, partially single stranded long DNA donor molecules dramatically promotes templated repair for the insertion of longer edits such as GFP. Although, we have only tested these strategies in *C. elegans*, it seems likely that the principals revealed here will be relevant in other systems. The key features include:

Utilization of a DNA expression vector as a co-injection marker that controls for injection quality, permits optimization of Cas9 RNP concentration, and monitors toxicity among a cohort of progeny inheriting long DNA required for templated repair.

Employment of hybrid PCR-based donors with single-stranded homology arms for consistent, high-efficiency insertion of large constructs.

Material and Methods

Strains and genetics: All the *C. elegans* strains were derived from Bristol N2 background and cultured on Normal Growth Media (NGM) plates seeded with OP50 bacteria[2].

Table 3.1: List of *C. elegans* Strains

Genotype	Strain Name
glh-1(ne4595[gfp::glh-1])I	WM606
wago-4(ne4479[gfp::wago-4])II	WM610
sago-2(ne4480[gfp::sago-2])I	WM611
ppw-1(ne4481[gfp::ppw-1])I	WM609
gcna-1(ne4598[gfp::gcna-1])III	WM613
dvc-1(ne4599[mcherry::dvc-1])V	WM614
top-2(ne4600[top-2::mcherry])II	WM615
wago-1(ne4435[3xflag::wago-1])I	WM616
wago-2(ne4416[3xflag::wago-2])I	WM617
ppw-2(ne4419[3xflag::ppw-2])I	WM618
wago-4(ne4422[3xflag::wago-4])II	WM619
wago-5(ne4427[3xflag::wago-5])II	WM620
sago-2(ne4478[3xflag::sago-2])I	WM621
ppw-1(ne4428[3xflag::ppw-1])I	WM622
wago-8(ne4432[3xflag::wago-8])V	WM623

hrde-1(ne4437[3xflag::hrde-1])III	WM624
wago-10(ne4430[3xflag::wago-10])V	WM625
wago-11(ne4614[3xflag::wago-11])II	WM626
nrde-3(ne4615[3xflag::nrde-3])X	WM627
rde-1(ne4616[3xflag::rde-1])V	WM628
ergo-1(ne4617[3xflag::ergo-1])V	WM629

Protocol for genome editing:

Materials:

1. S. pyogenes Cas9 3NLS (10 µg/µl, IDT)
2. tracrRNA (IDT)
3. crRNA (2nmol or 10nmol, IDT)
4. ssDNA donor (standard desalting; 4 nmol Ultramer, IDT)
5. PRF4::rol-6(su1006) plasmid

Re-suspension:

1. Aliquot 0.5 µl (5µg) of Cas9 protein and store at -80°C to avoid freeze/thaw cycles. Use 1 aliquot per injection and add all the other reagents sequentially to the Cas9 tube.
2. tracrRNA – 0.4 µg/µl in IDT duplex buffer, store at -20°C
3. crRNA – 0.4 µg/µl in IDTE P.H 7.5, store at -20°C
4. ssDNA oligo donor – 1 µg/µl in ddH2O, store at -20°C

5. PRF4::rol-6 (su1006): 500 ng/µl

Donor Design and Generation

ssODN donors:

For generating short inserts (<130bp):

To generate a ssODN donor, add 35 bases of 5′ homology sequence in front of the tag and 35 bases of the 3′ homology sequence at the end. Remember to mutate the PAM site or the guide binding sequence if it is not already disrupted by the insert. If the guide binding sequence is mutated, length of homology sequence should be 35bp from the last mutation[212].

For generating point mutations:

Pick 35bp homology upstream and 35bp homology downstream of your guide cut site, which should ideally be within 20bp of the desired mutation site. Introduce the desired mutations in the donor and the PAM/ guide binding sequence.

For large deletions with 2 guides:

Pick 35bp homology upstream of the left-guide cut site and 35bp homology downstream of the right-guide cut site and put them together. Everything in between will be removed. Deletions up to 1kb can be easily achieved through this strategy. In principle, this should work for larger deletions as well.

dsDNA asymmetric-hybrid donors:

1. Order 140bp oligos from IDT as Ultramers; 120bp as homology arms and 20bp complementary to GFP (or any other desired insert). Also, order standard oligos just complementary to your insert (no homology arms).
2. Generate two PCR products; one with 120bp homology arms and the other without any homology arms (only insert sequence) using an insert containing plasmid as the template for PCR; perform 4 to 8 50 µl reactions for each product.
3. Run 5 µl on an agarose gel to check if a single bright band at ~1050bp (gfp +120bp+120bp) is obtained (in some cases the template plasmid band might be detected. It can be ignored as it does not interfere with HDR.). If non-specific amplification is observed, set up a temperature gradient and find the optimal temperature. (OPTIMIZATION NOTE: we find that amplification of challenging templates with long ultramer primers is aided by increasing the template plasmid concentration and decreasing primer concentrations.)
4. Combine all the PCR reactions of each product and column purify (we use Qiagen minElute kit), elute in 10-20 µl of water depending on brightness of the band, aiming to get >300ng/ul concentration.
5. Mix 1:1 of the purified PCRs (2µg:2µg for 20µl injection mix), heat to 95ºC and cool to 4ºC to re-anneal (95ºC-2:00 min; 85ºC-10 sec, 75ºC-10 sec, 65ºC-10 sec, 55ºC-1:00 min, 45º-30 sec, 35º-10 sec, 25º- 10 sec, 4º-forever.)
6. Add this donor cocktail to the rest of the injection mixture ONLY after pre-incubating Cas9, crRNA and tracrRNA (see below).

(STORAGE NOTE: we store individual purified donors at -20C and pre-assemble the cocktail directly before injection. We have not explored storage and re-use of pre-assembled cocktails)

Preparing injection mixtures:

Add components of the injection mixture in the following sequence:

1. Cas9- 0.5 µl of 10 µg/µl stock
2. Add tracrRNA – 5µl of 0.4 µg/µl stock
3. Add crRNA – 2.8µl of 0.4 µg/µl stock (if you are using two guides add 1.4 µl of each)
4. Incubate this mixture @37°C for 10 minutes before adding any DNA. Adding any double stranded DNA before RNP complex formation reduces HDR efficiency dramatically.
5. Add ssODN donor – 2.2 µl of 1µg/µl stock or dsDNA donor cocktail – 200 ng/µl (total 4 µg) in the final injection mixture
6. Add PRF4::rol-6 (su1006) plasmid – 1.6 µl of 500 ng/µl stock
7. If needed, add nuclease free water to bring the volume to 20µl
8. To avoid needle clogging, centrifuge the mixture @14000rpm for 2 min, transfer about 17 µl of the mixture to a fresh tube and proceed to loading the needles.

Note: some protocols provide additional KCl in the injection mixture to reduce Cas9 aggregation. Under the reduced Cas9 concentration conditions we have

not found aggregation to be an issue and find that KCl from IDT's buffer added to tracrRNA is sufficient.

Micro-injection and screening:

1. Inject 10 to 20 animals and transfer them onto individual plates. After about 3 days, score for F1 rollers and place each roller onto a separate NGM plate.
2. In general, injections with ssODNs yield a greater number of F1 rollers per injected animal compared to the injections with dsDNA.
    a. For ssODN-based editing: Pick 2 plates that segregate the greatest number of F1 rollers and from these 2 plates, pick about 24 F1 rollers and place them onto separate plates.
    b. For dsDNA-based editing: Pick at least 24 F1 rollers from several plates and place them onto to separate plates.
3. To avoid false positives due to mosaicism in F1 animals, pick several F2s from each plate, perform lysis and genotyping. Genotyping primers should lie outside the homology arms to avoid false positives from transiently retained donor molecules. In some circumstances large inserts do not readily amplify in heterozygotes (because the small wild type band amplifies preferentially) In those situations it might be necessary to employ one primer inside the insert for each junction.
4. Alternatively, if the expression levels are detectable, insertions of fluorescent tags can be screened under a microscope either by using high magnification

fluorescence microscope (mount several F2 animals onto 2% agarose pads) or by using a fluorescence dissecting scope.

Results

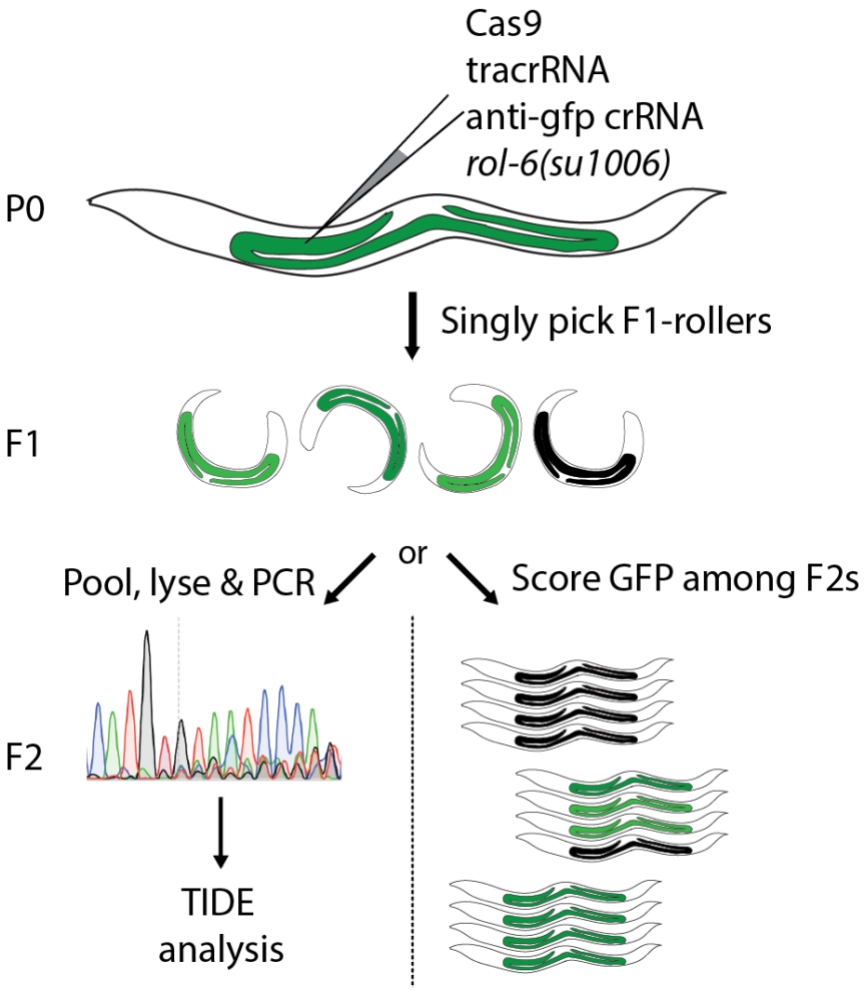
Cas9 RNP mixtures can be toxic at high concentrations

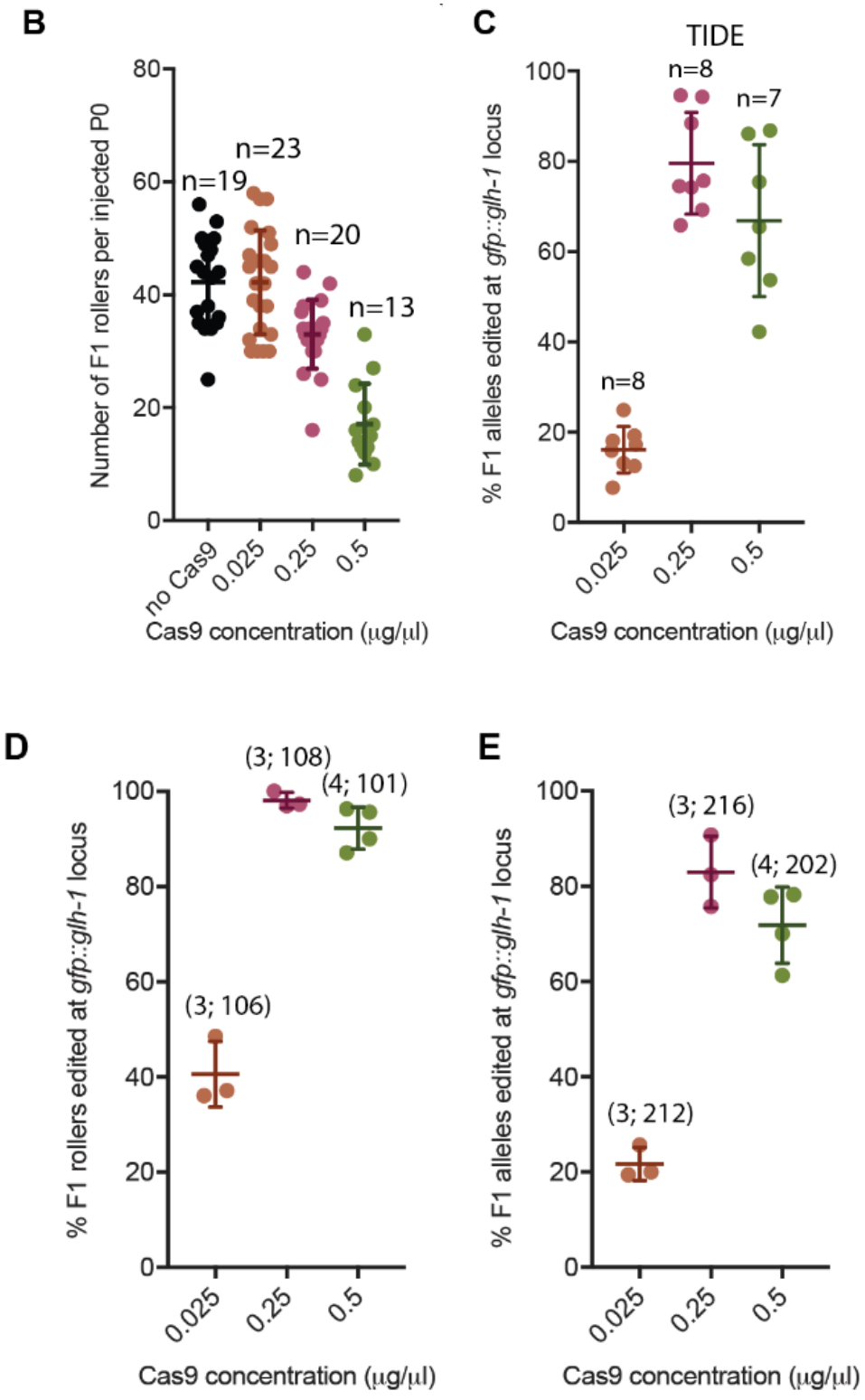
In the course of adopting Cas9 RNP editing methodologies[212] we decided to monitor injection quality by adding the pRF4::*rol-6(su1006)* plasmid to the injection cocktail[214]. We were very surprised to find that, despite giving high numbers of edited progeny, the numbers of transgenic Roller (*rol-6*) animals were greatly reduced. For example, in the course of two independent attempts to target the *vet-2* locus (a non-essential gene) we only recovered total of 32 rollers from 51 injected P0 worms, an average of less than one roller per injected P0. Moreover, we noted that the few surviving Roller animals obtained were often sick and sterile (data not shown), suggesting that toxicity, or off-target genome editing might cause the lack of Roller transgenics.

To address these possibilities, we performed a titration of RNP concentrations while holding the Roller DNA concentration constant. We then examined both the genome editing efficiency and the frequency of Roller transgenics among F1 progeny of the injected animals. Worms expressing the bright fluorescence marker GFP::GLH-1 were co-injected with 40ng/µl pRF4::*rol-6(su1006)* plasmid and dilutions of Cas9 RNPs loaded with an anti-gfp guide (Figure 3.1 A). In our pilot studies we recovered very few Rollers at 2.5µg/µl of Cas9 used in initial *C. elegans* Cas9 RNP protocols[212,217], we therefore decided to begin with a 5 fold dilution, 0.5 µg/µl as a starting RNP concentration. This concentration was recently proposed by Prior and colleagues[209]. Injections using 0.5 µg/µl of Cas9

resulted in an average of 17 F1 roller progeny per injected P0 animal. Reducing the concentration by 2-fold, down to 0.25µg/µl doubled the frequency of F1 rollers to 33, while a ten-fold dilution to 0.025µg/µl resulted in 43 F1 roller progeny per P0 (Figure 3.1 B). These latter two F1 roller frequencies are comparable to the rate of 42 F1 rollers per P0 obtained when pRF4::*rol-6(su1006)* is injected alone (Mello et al[214] and Figure 3.1 B). Taken together these findings suggest that RNP concentrations below 0.25µg/µl do not interfere with expression of the co-injected Roller marker gene.

We next asked how Cas9 RNP concentrations affected the in-del frequency at the *gfp::glh-1* locus. To measure in-del rates in a high throughput fashion we used the TIDE analysis pipeline, which estimates the in-del rates in a mixture of PCR products (Figure 1A, left)[216]. To do this we PCR amplified the *gfp::glh-1* locus from pools of at least 10 F1 rollers segregated by an injected P0 worm and subjected the mixture to sanger sequencing and TIDE analysis. Using this approach, we found that at 0.025µg/µl ~16% of alleles carried an in-del. The number of edited alleles increased to ~80% at 0.25µg/µl (Figure 1C) but did not increase further when the Cas9 concentration was doubled to 0.5µg/µl, and in fact appeared to decline slightly to ~67% (Figure 3.1 C). Because GFP::GLH-1 is easily detectable in adult animals under the fluorescence-dissecting microscope, we were able to validate the TIDE results directly using microscopy (Figure 3.1 A, right). For example, we determined that at 0.25µg/µl of Cas9 ~98% of all F1 rollers segregated GFP-negative (successfully edited)


Cas9
tracrRNA
anti-gfp crRNA
rol-6(su1006)
P0
Singly pick F1-rollers
F1
Pool, lyse & PCR
or
Score GFP among F2s
F2
TIDE
analysis


B
Number of F1 rollers per injected P0
n=19
n=23
n=20
n=13
no Cas9
0.025
0.25
0.5
Cas9 concentration (µg/µl)
C
TIDE
% F1 alleles edited at gfp::glh-1 locus
n=8
n=7
n=8
0.025
0.25
0.5
Cas9 concentration (µg/µl)
D
% F1 rollers edited at gfp::glh-1 locus
(3; 108)
(4; 101)
(3; 106)
0.025
0.25
0.5
Cas9 concentration (µg/µl)
E
% F1 alleles edited at gfp::glh-1 locus
(3; 216)
(4; 202)
(3; 212)
0.025
0.25
0.5
Cas9 concentration (µg/µl)

F

Brood	(-/-)	(-/*gfp*)	unedited	edited (%)
1 (n=33)	18	14	1	96.96
2 (n=38)	31	7	0	100
3 (n=37)	25	11	1	97.29

F1 rollers edited at the *gfp::glh-1* locus
with 0.25 µg/µl (1.5 µM) of Cas9 protein

Figure 3.1: **Determining optimal Cas9 RNP concentrations.**

(A) Schematic representation of the optimization workflow. Cas9 protein loaded with anti-gfp guide is co-injected at several concentrations with 40ng/µl of pRF4::*rol-6(su1006)* plasmid into *gfp::glh-1* animals. Number of F1 Rollers segregated by each injected P0 is scored. F1 Rollers are then subjected to genotyping as a pool by TIDE analysis (left) or their F2 progeny are scored by microscopy (right).

(B) Number of F1 Rollers recovered from a P0 animal injected with *rol-6(su1006)* plasmid alone or with *rol-6(su1006)* and Cas9 RNP at three different concentrations. Each dot represents an individual animal and (n) refers to the number of broods scored for each condition. Only broods containing at least one roller were scored. (**C**) Percent of alleles carrying an in-del at the *gfp::glh-1* locus at three different Cas9 concentrations as determined by TIDE analysis. Each dot represents a pool of at least 10 F1 Rollers from one injected P0 and (n) refers to the number of broods scored in each condition.

(D) Percentage of F1 rollers segregating GFP-negative F2 progeny plotted versus the concentration of Cas9 protein used in the injection mixture. Numbers in parentheses indicate: (number of injected P0s; number of F1 rollers). (E) Percentage of edited gfp::glh-1 alleles calculated based on numbers of homozygous and heterozygous F1 rollers (in Figure 1D) plotted versus the concentration of Cas9 protein used in the injection mixture (compare with TIDE data in Figure 1C). Numbers in parentheses indicate: (number of injected P0s; number of alleles). (**F**) A detailed breakdown of the F1 rollers among the three broods from the 0.25µg/µl Cas9 injection. (n) refers to total number of F1 rollers. All error bars represent standard deviation from the mean.

progeny (Figure 3.1 D). Furthermore, ~68% were homozygous, producing only GFP negative progeny, while another ~30% were heterozygous. Based on these numbers we can calculate that 83% of all *gfp::glh-1* alleles were successfully edited at 0.25µg/µl of Cas9 (Figure 1 E and F). These numbers correlate well with TIDE data (Figure 3.1 C), and thus lend confidence to the calculations of the percentage of *gfp::glh-1* alleles cleaved at each Cas9 concentration (Figures 3.1 C and E). Finally, to determine the reproducibility of these findings we repeated the injections with a previously characterized moderately efficient guide targeting the *unc-22* locus[192] and observed similar results (data not shown).

Efficient editing with ssODN donors using a Roller plasmid co-injection marker

The above findings demonstrate that Roller plasmid co-injection identifies animals that are highly likely to undergo CRISPR-induced DNA double strand breaks. We next wished to test this methodology for achieving homology-directed repair (HDR). To do this we decided to introduce a (3X)FLAG-affinity tag into each of the twelve worm-specific Argonautes (WAGOs) as well as two additional Argonautes ERGO-1 and RDE-1. For each gene we designed guides targeting the PAM site closest to the ATG start codon (without any further optimization or guide testing) (Figure 3.2 A)[212]. *wago-1* and *wago-2* are highly similar near the ATG and no specific guide could be designed; thus, one guide targeting both loci was used.) Each mixture was then injected into adult N2 worms using standard worm gonadal injection methodology (Figure 3.2 B)[200]. For each experiment we injected ~10 P0 animals and singled ~24 F1 Rollers from plates segregating the

most Rollers (indicative of the best injections). After producing broods, Rollers were genotyped for 3XFLAG insertions (Figure 3.2 B).

We were able to recover 13 out of 14 tagged strains among the first 24 F1 progeny screened by PCR from each set of injections, and in every case we recovered multiple independent alleles (Table 3.2). Although genotyping suggested that we recovered a significant number of putative F1 homozygotes (Figure 3.2 B [asterisks]), these animals were not used to establish lines. These F1 homozygotes are expected to carry two different alleles of the modified gene. For example, without additional analysis homozygous F1s could not be conclusively distinguished from trans-heterozygotes carrying a correct edit, over a partial or imprecise edit or an edit that deleted one of the genotyping primer binding sites. Thus, for simplicity of the genetic analysis, independent lines were established by selecting homozygous F2s segregated by three different heterozygous F1 animals (Table 3.1). Accuracy of each insertion was validated by sequencing. The average success rate for precise insertion of the 3XFLAG tag ranged from 10-73% and averaged ~34% (Table 3.2 and Figure 3.2 C).

Plotting the insertion efficiency versus the distance between the Cas9-induced cut and the desired insertion site (directly after the ATG start codon [Figure 2A]) we found no strong correlation up to 20 base pairs away (Figure 3.2 C). The *wago-6 (sago-2)* locus was the only outlier, likely because the nearest available cut site suitable for use with the original donor design was 27 bases away from the site of insertion. Although a number of insertions were recovered at this locus

they were either out of frame or contained random DNA sequences (data not shown). The *wago-6* gene contains a second PAM site located right at the ATG start codon. This site was not used originally because the 3XFLAG donor sequence (which starts with a “G”) would not disrupt the PAM. Moreover, the alternative approach to prevent re-cutting of the repaired locus, mutating the guide binding site, would require introducing potentially undesirable mutations into the 5’ UTR. To solve this problem, we added an extra CCC, proline codon, to the 3XFLAG donor sequence, immediately downstream of ATG (Figure 3.3). Using this donor and guide we recovered *flag::wago-6* alleles in 52% of the F1 Roller animals analyzed.

In all of the edited strains the Roller phenotype was expressed only transiently during the F1, indicating extra-chromosomal expression[214]. These findings demonstrate the general utility of the Roller marker for identifying edited animals without introducing additional edits or undesired phenotypes into the resulting strains. In addition, these findings indicate that as long as the desired insertion site resides within 20 bp of the cut site, ssODN donors provide for highly efficient editing.

Hybrid dsDNA donors promote the integration of large constructs

High rates of HDR have been reported using PCR-generated double stranded DNA (dsDNA) ~1kb-sized donors with ~35bp homology arms[212]. However, we have struggled to reproduce these successes using the original or optimized

protocols (data not shown). Extending the homology arms from 35bp to 120bp resulted in low, ~2%, but reproducible levels of GFP or mCherry integration at 6 different loci (Table 3.3 and Table 3.4). Thus, in our hands, there was a large gap between the efficiency of templated repair using ssODNs versus longer dsDNA donors.

A recent study proposed that ssODN donors are integrated by a highly efficient single stranded template repair (SSTR) pathway, while dsDNA donors rely on a less efficient homology-directed repair (HDR) pathways[218]. We therefore wondered whether we could achieve the improved efficiency of ssDNA by employing large PCR-based donors with single stranded overhangs. To test this idea, we generated two PCR donors to target the same locus: one with a 120bp left homology arm and a 35bp right homology arm, and the other with 35bp on the left and 120bp on the right. By mixing these donors at equimolar quantities, then melting and re-annealing the mixture we should get a mixture of four different molecules (Figure 3.4 A) two of which have either 3′ or 5′ single stranded overhangs. Alternatively, hybrid asymmetric PCR donors were prepared by annealing molecules with 120bp homology arms to a PCR product containing just the insert, with no homology arms (Figure 3.4 B). 200ng/µl of blunt donor or hybrid cocktail was used in the optimized editing protocol (Figure 3.2 B), and integration was scored by PCR and multiple positives were validated with sequencing across the junction as well as by microscopy. Strikingly, both types of hybrid dsDNA donor cocktails consistently yielded higher rates of accurate

integration at three different loci, compared to melted and re-annealed traditional blunt donors (Figure 3.4 C). We were successful at generating N- and C-terminal fusions with GFP and mCherry tags at rates comparable to ssODNs', ~20% of F1 Rollers. Hybrids between the PCR product with 120bp homology arms and a PCR product containing just the insert, lacking arms, (Figure 3.4 B) yielded the best precise editing rates, indicating that homology arms on the shorter product are not required to stimulate recombination. Hybrids with shorter overhangs (60bp of homology) provided some precise insertion, but were not as effective as 120bp arms (Table 3.4).

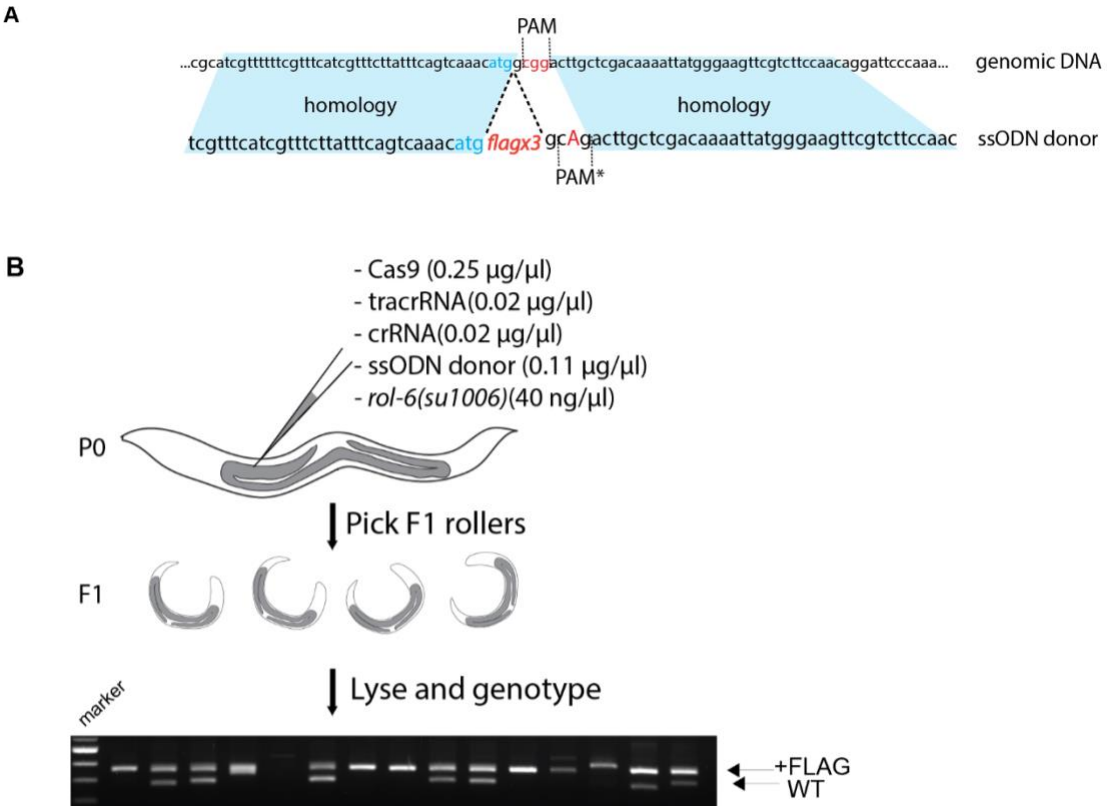

A
PAM
...cgcatcgtttttttcgtttcatcgtttcttatttcagtcaaacatggcggacttgctcgacaaaattatgggaagttcgtcttccaacaggattcccaaa...
genomic DNA
homology
homology
tcgtttcatcgtttcttatttcagtcaaacatg flagx3 gcAgacttgctcgacaaaattatgggaagttcgtcttccaac
ssODN donor
PAM*
B
- Cas9 (0.25 µg/µl)
- tracrRNA(0.02 µg/µl)
- crRNA(0.02 µg/µl)
- ssODN donor (0.11 µg/µl)
- rol-6(su1006)(40 ng/µl)
P0
Pick F1 rollers
F1
Lyse and genotype
marker
+FLAG
WT
C
% HDR
100
80
60
40
20
wago-4
rde-1
wago-1
wago-6^
wago-3
wago-10
wago-8
wago-2
wago-7
wago-5
wago-12
ergo-1
wago-11
wago-9
wago-6
ATG
10
20
30
no.of bp between cut and insertion sites

Figure 3.2. Efficient integration of 3XFLAG ssODN donor at fourteen of the *C. elegans* argonaute genes using pRF4::*rol-6(su1006)* co-injection marker. (A) Schematic of donor design for 3XFLAG insertion directly downstream of the ATG (based on[212]. Blue shading highlights homology arms, red letters indicate the PAM site, blue letters represent the START codon, capital A is the mutation introduced to disrupt the PAM site in the donor. (B) Schematic of the CRISPR protocol. Simplified injection mixture contains just the RNP components, the ssODN donor, and *rol-6* plasmid. Approximately 24 F1 rollers from two best injection plates were single cloned and genotyped. Lower band is the wild-type PCR product; upper band is upshifted due to 3xflag insertion. * mark putative homozygotes. (C) Efficiencies of 3XFLAG insertion plotted versus distance of the guide cut site from the START codon. Detailed underlying data supplied in Table 3.2. Each dot represents targeting of one gene. ^ indicates the repeated attempt at targeting *sago-2* using the donor described in Figure 3.3.

Figure 3.3. Alternative targeting strategy for *wago-6 (sago-2)*. A schematic representation of the donor design to introduce *flagx3* into the *wago-6* locus using the ATG-proximal PAM site that would not be normally disrupted by *flag* insertion. Briefly, introducing an extra proline codon (dark blue capital CCC) upstream of the flagx3 sequence in the donor for *wago-6* locus disrupts the PAM site in the donor.

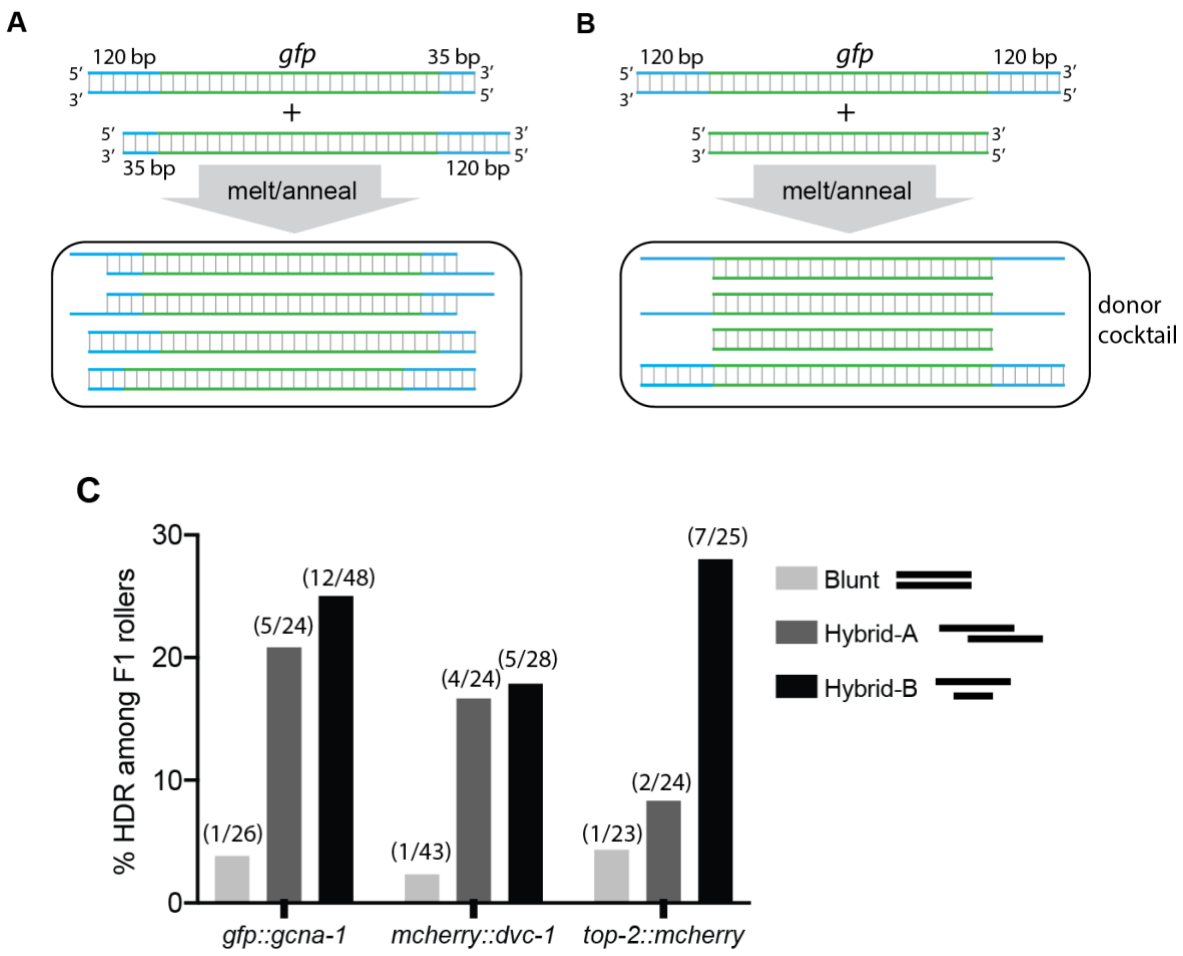


Figure 3.4 Efficient editing with long, partially single-stranded dsDNA donors. (A) and (B) Schematics of the strategy for generating hybrid dsDNA donor cocktail featuring molecules with ssDNA overhangs. (C) Integration efficiencies of GFP or mCherry fluorescent tags using blunt donors or hybrid dsDNA donor cocktail at diverse loci, plotted as a fraction of F1 rollers positive for appropriate insert as detected by PCR. Numbers above each bar indicate number of insert-positive rollers over total number of rollers.

Discussion

CRISPR/Cas9 genome editing is highly efficient in *C. elegans* and should be accessible to investigators of all levels of experience. The protocols described here establish clear benchmarks for implementation and troubleshooting. We demonstrate efficient editing at diverse genomic loci provided that editing targets are reasonably proximal (<20bp) to a PAM site. For short inserts (<140bp), we find the best efficiency with ssODN donors, as was previously reported[209,212]. For longer inserts, we recommend using donors that are hybrids of two asymmetric PCR products or a hybrid of a traditional symmetric donor and the insert (Figure 3.4 B). The detailed version of our protocol is included in the supplemental materials.

An important feature in any microinjection protocol is the inclusion of metrics that enable troubleshooting. For example, during development of basic DNA transformation protocols for *C. elegans* it was found that poorly purified DNA, too much DNA, or even specific DNA sequences can be toxic. Thus, the inclusion of a DNA marker such as the pRF4::*rol-6(su1006)* plasmid, that reports on the viability of progeny inheriting DNA, enables quick identification of toxic injection mixtures[214]. The current findings suggest that, like DNA preparations, RNP mixtures can interfere with inheritance of co-injected DNA. Importantly, RNPs distribute so widely, to even hundreds of progeny, and induce oligo mediated templated repair so efficiently[212], that the absence of the 20 or 30 progeny that typically inherit large co-injected DNA molecules could easily escape detection.

Thus, we propose that using only indel frequency or the oligo-driven repair efficiency to monitor the activity of the editing mixture, is not sufficient. Instead we recommend the inclusion of a plasmid-DNA-driven visible marker such as *rol-6*. Expression of this marker reports on a segment of the brood that inherits long double stranded DNA, and thus identifies animals that likely also inherit long dsDNA donor templates and may thus incorporate longer edits such as GFP insertions. We are not arguing that this procedure yields higher rates of editing than other (properly optimized) protocols, but rather that the current methods provide important metrics for troubleshooting, particularly when longer DNA insertions are desired.

Early *C. elegans* genome editing protocols employed DNA vectors to express all the editing components. To trouble shoot these protocols we and others advocated using known and validated guide vectors such as *unc-22* or *dpy-10* as "co-CRISPR" markers[192] Cutting at the previously validated target locus reported on DNA-driven Cas9 activity, and because co-injected DNA vectors are generally inherited together[214] also reported on the viability among animals inheriting co-injected DNA repair templates[192]. In the current protocol we utilize RNP driven Cas9 activity and show that RNP activity distributes much more broadly and does not report on viability of animals receiving large DNA templates required for longer edits. Moreover, whereas vector driven guides were often non-functional – with synthetic guide RNA preparations we have yet to encounter guides that do not cut the target locus. Thus, in practice it has not been necessary to monitor

the activity of each new synthetic RNA guide. Instead we recommend a sequential test, to first monitor Cas9 RNP activity and toxicity by, using a known and well-validated guide RNA (such as the GFP guide described here) and pRF4::*rol-6(su1006)* plasmid DNA. Every new batch of editing enzyme should first be tested to ensure, both that editing occurs and that viable rollers are obtained at reasonable levels (typically a few from each injected animal). These same validated conditions are then used with new guides, until or unless a problem occurs. For example, in the event that the desired edits are not obtained among the rollers, or if rollers are absent, then further tests will be needed to ensure that the new guide preparation is functional and not toxic. If rollers are absent the particular guide RNP might be toxic, perhaps cleaving an essential locus at very high efficiency. Further dilution of the guide/RNP mix until rollers are once again observed would likely solve this problem. If editing is still not observed, one might also wish, at that point to perform a co-CRISPR test with two RNPs mixed together to monitor guide RNA toxicity. In practice we just have not, as yet, needed to undertake these additional trouble-shooting step when using the very robust RNP methodology.

There are several additional advantages to using the pRF4 roller expression marker for RNP-based editing. The Roller phenotype is dominant and easily scored under the light dissecting microscope. Plasmid DNA preparation is inexpensive, and injection of plasmid DNA at these concentrations results, primarily, in transient F1 expression without further inheritance in subsequent

generations. Indeed a recent study employed an *mCherry::myo-2* plasmid to identify ssODN templated editing events[209], demonstrating the feasibility of using other plasmid-based co-injection markers for genome editing. However, we find Roller more convenient as a fluorescence dissecting scope is not needed for scoring.

Our findings suggest that Cas9 RNP mixtures can be toxic and can eliminate F1 progeny that receive the largest amounts of co-injected long dsDNA. In this study we tested only one source of commercially available Cas9 protein. Since RNP activity and toxicity will likely vary depending on the specific target or guide sequence, or due to variations in protein preparation or impurities, we recommend that Cas9 RNP preparations be tested routinely for optimal concentration using the simple and inexpensive *rol-6*/TIDE approach (Figure 3.1 A).

We do not yet know how hybrid dsDNA PCR donors stimulate HDR, and it will be important to fully test the limits of this approach in terms of maximal donor length and minimal single-stranded overhangs and optimal donor concentrations. It seems likely that other modifications, such as chemical modifications to the ends of the donor molecule may drive even greater efficiencies. The procedure for generating hybrid donors is extremely easy to implement and we anticipate that these types of donors will also stimulate precise editing in other systems. In summary, it is now as easy to precisely edit the worm genome as it is to generate the iconic Roller transgenics first described by Mello et al.[214]. We strongly

encourage even the total novice worm breeder to begin editing the genome of this fascinating “yeast” of metazoa.

Locus	# F1 rollers positive for insertion by PCR	Total # of F1 rollers genotyped
wago-1	18	30
wago-2	10	30
wago-3	17	30
wago-4	22	30
wago-5	8	29
sago-2 (wago-6)	0	30
*sago-2 (wago-6)^*	33	64
ppw-1 (wago-7)	7	24
sago-1(wago-8)	9	24
hrde-1 (wago-9)	3	30
wago-10	10	24
wago-11	8	48
nrde-3 (wago-12)	5	24
ergo-1	4	24
rde-1	14	24

Table 3.2: 3XFLAG tag insertions in N-termini of fourteen argonaute genes using ssODN donors and *rol-6* co-injection marker. Breakdown of numbers used to derive the %HDR efficiencies plotted in Figure 2C. ^indicates the repeated attempt at targeting *sago-2* using the donor described in Figure 3.3.

Locus	no. of F1 rollers *(rol-6)*	
	Screened	GFP+
wago-4	54	1 (1.85%)
sago-2 (wago-6)	119	1 (0.84%)
ppw-1 (wago-7)	76	1 (1.32%)

Table 3.3: HDR efficiencies of GFP insertion with blunt-ended PCRs as donors. All the donors consist of 120bp long homology arms on both ends.

Locus	Donor type	F1 rollers	
		Screened	GFP/mCherry+
gcna-1	120bp blunt	26	1 (3.85%)
	60bp hybrid B	24	4 (16.67%)
	120bp hybrid A	24	5 (20.83%)
	120 hybrid B	48	12 (25.0%)
dvc-1	120bp blunt	43	1 (2.33%)
	60bp hybrid B	19	0 (0%)
	120bp hybrid A	24	4 (16.67%)
	120 hybrid B	28	5 (17.86%)
top-2	120bp blunt	23	1 (4.35%)
	60bp hybrid B	32	1 (3.13%)
	120bp hybrid A	24	2 (8.33%)
	120 hybrid B	25	7 (28.0%)

Table 3.4: Efficiencies of long donor integration. Lengths (in bp) refer to length of homology arms.

Acknowledgements: We thank Geraldine Seydoux at Johns Hopkins University and Scott Wolfe at the University of Massachusetts Medical School, Worcester for sharing Cas9 protein for pilot experiments. This work was funded by: American Cancer Society Fellowship (G.A.D.), Howard Hughes Medical Institute (C.C.M.), NIH P01# HD078253 (C.C.M).

Preface to Chapter IV

Contents of this chapter appeared in the following publication

Ghanta, K.S., and Mello, C.C. (2020). Melting dsDNA Donor Molecules Greatly Improves Precision Genome Editing in *Caenorhabditis elegans*. Genetics 216, 643-650.

My contributions to this chapter are Figures 4.1- 4.10

Craig Mello contributed to Figures 4.1B, 4.2, 4.5D, 4.7 and 4.10C-D

Chapter IV: Melting the dsDNA donors potentiates precision genome editing in *Caenorhabditis elegans*

Abstract

CRISPR genome editing has revolutionized genetics in many organisms. In the nematode *Caenorhabditis elegans* one injection into each of the two gonad arms of an adult hermaphrodite exposes hundreds of meiotic germ cells to editing mixtures, permitting the recovery of multiple indels or small precision edits from each successfully injected animal. Unfortunately, particularly for long insertions, editing efficiencies can vary widely, necessitating multiple injections, and often requiring co-selection strategies. Here we show that melting double stranded DNA (dsDNA) donor molecules prior to injection increases the frequency of precise homology-directed repair (HDR) by several fold for longer edits. We describe troubleshooting strategies that enable consistently high editing efficiencies resulting, for example, in up to 100 independent GFP knock-ins from a single injected animal. These efficiencies make *C. elegans* by far the easiest metazoan to genome edit, removing barriers to the use and adoption of this facile system as a model for understanding animal biology.

Introduction

In the nematode worm *C. elegans*, genome editing can be achieved by direct injection of Cas9 guide-RNA ribonucleoprotein (RNP) complexes into the syncytial ovary[208,212,217]. In the worm germline, such injections afford the editing machinery simultaneous access to hundreds of meiotic germ nuclei that share a common cytoplasm. Under optimal conditions the frequency of F1 progeny with indels caused by non-homologous end joining (NHEJ) can be greater than 90% of those progeny expressing a co-injection plasmid marker gene[208]. Leveraging these high cutting efficiencies, precise genome editing is readily achieved using short (under ~200 nucleotide [nt]), single-stranded oligodeoxynucleotide (ssODN) donors, permitting insertions of up to ~150 nt in length[160,194,208,209,212]. However, with longer dsDNA donors (~1kb), HDR events are recovered at lower frequencies, require more complex protocols, high concentrations of the donor DNA, and typically require screening the broods of multiple injected animals[140,192,194,202,208,210,211,213,219-221].

There are multiple reasons why longer repair templates may be less efficient as donors for HDR compared to ssODNs. First, empirical studies suggest that long dsDNA is more toxic than short single-stranded DNA[214], limiting safe donor concentrations to less than 200 ng/µl for ~1kb donors. Second, upon injection into germline cytoplasm, dsDNA molecules quickly form large extra-chromosomal arrays via both end-joining and homologous recombination pathways, and appear to do so while sequestered away from genomic DNA[214,215].

Concatenation of donor molecules into large arrays would have the effect of lowering the number of individual molecules available to access and to template repair at the target site double strand break (DSB). Moreover, if injected DNA assembles concatenates while sequestered from the nuclear DNA—perhaps within de novo nucleus-like organelles[222]—this process could preclude templated repair of a genomic target site until after the sequestered concatenates gain nuclear access after nuclear envelope breakdown occurs post-fertilization.

In the previous chapter, we showed that CRISPR-mediated HDR could be increased ~4-fold by mixing, melting, and re-annealing overlapping donor molecules to create donor templates with single-stranded overhangs[208]. In those previous studies, we limited our analysis to a cohort of F1 'Roller' progeny that express the co-injection marker gene *rol-6 (su1006)*. Here, to explore editing efficiency outside the Roller cohort, we scored the entire brood of each injected animal for precisely edited progeny that incorporate and express fluorescent protein markers (GFP or mCherry). We show that the vast majority of insertions occurred later in the brood, after the cohort of progeny that express the Roller phenotype. Whereas overhangs improved the frequency of editing among the F1 Rollers[208], they had no benefit within this latter segment of the brood. Instead, melting the donor molecules, alone, sufficed to increase the HDR frequency to as high as 50% of the post-injection progeny. We provide a protocol and troubleshooting strategies that enable even a novice injector to achieve their editing goals and to optimize editing efficiencies.

Material and Methods

Strains and genetics

All the strains were generated in the Bristol N2 background unless specified otherwise and cultured on normal growth media (NGM) plates seeded with OP50 bacteria[2]. Strains used in this study are listed in Table 4.1.

At CSR-1 locus, GFP was introduced between FLAG::linker(9bp) and TEV in FLAG::linker::TEV::CSR-1 strain.

Scoring methodology

Injected P0 animals were individually cultured on NGM plates at room temperature (22°C- 23°C) unless specified otherwise. P0 animals with more than 100 post-injection progeny and at least 20 Rollers were selected— except at 100 ng/µl and 200 ng/µl of dsDNA donor where number of Rollers can be lower than 20 due to toxicity— and their F1 progeny were scored between 72 and 90 hours post-injection. All the F1 progeny from each brood were mounted onto 2% agarose pads and screened under fluorescence microscope for GFP or mCherry expression. GraphPad Prism (Version 8.4) was used to perform statistical tests and calculate P-values.

Oligos and donors

End-modified donors were generated by PCR using oligos with 5′ SP9 modifications (IDT). Oligos used for to generate *hrde-1* and *F53H1.1 gfp* donors also contain 15bp linkers on either end of *gfp* which also serve as PCR primers.

Sequences of all the crRNAs and oligos are provided in Supplemental Material, Table 4.2 and Table 4.3 respectively.

Table 4.1: List of *C. elegans* strains

Genotype	Strain Name
F53H1.1(ne4815[gfp::F53H1.1]) IV	WM703
glh-1(ne4816[gfp::glh-1]) I	WM704
csr-1(ne4483[flagx3::linker::tev] IV	WM705
znfx-1(ne4817[mcherry::znfx-1]) II	WM706
hrde-1(ne4818[gfp::hrde-1]) III	WM707
csr-1(ne4819[flagx3::linker::gfp::tev::csr-1]) IV	WM708

Table 4.2: Sequences of crRNAs

ID#	Target locus	Sequence (5′→3′)	Notes
CMG-18	*hrde-1*	CATAATTTTGTCGAGCAAGT	To insert N-terminal tag
CMG-25	*csr-1*	AAGATGTTCAGGGCAAGTCT	To insert N-terminal tag
CMG-33	*Flagx3::linker::tev*	TATAAAGACGATGACGATAA	To insert gfp at csr-1 locus
CMG-34	*glh-1*	TTTTCTGCGAAAATGTCTGA	To insert N-terminal tag
CMG-35	*glh-1*	TGCGAAAATGTCTGATGGTTG	To insert N-terminal tag; (A.s. Cpf1)
CMG-77	*F53H1.1*	TTCCAGTTTTCGATGGGTCG	To insert N-terminal tag
CMG-79	*F53H1.1(A.s. Cpf1)*	CAGTTTTCGATGGGTCGCGGC	To insert N-terminal tag
CMG-88	*znfx-1*	AGGTTTCTGACCATTGAATA	To insert N-terminal tag

Table 4.3: Sequences of oligos

ID#	Sequence (5′→3′)	Notes
cmo-KG475F	ATTTTCTGGAAAAATCTTAA	*gfp::glh-1* donor
cmo-KG476R	TTAGCAGCACTTTCGCTATC	
cmo-KG830F	/SP9/TCGTTTCATCGTTTCTTATTTCAGTCAAACATG TCCGGAGGGAGTGGA	*gfp::hrde-1* donor
cmo-KG831R	/SP9/GTTGGAAGACGAACTTCCCATAATTTTGTCGA GCAAGTCTGCAGAACCTCCGCCACC	
cmo-KG832F	/SP9/ATCCAAAAATCCCCAATTTTTTCCAGTTTTCGA TGTCCGGAGGGAGTGGA	*gfp::F53H1.1* donor
cmo-KG833R	/SP9/ACGCTTTCGTTTGTGCTCTTTGTGCTCGCCGC GACCAGAACCTCCGCCACC	
cmo-T1193F	CTTGTTTCAGACCAATTCGCCAACCGTATTCAATGG TCTCAAAGGGTGAAGAAGA	*mCherry::znfx-1* donor
cmo-T1194R	GGCGGCGGGAGCCCTGGGGGGGCGAGGTTTCTGA CCTTATACAATTCATCCATGC	
cmo17648	AATCTCAATCAGGACGGTAAAG	*hrde-1* indel detection
cmo17649	GAACTCCTAGGCATAATGTTGA	
cmo-JG55F	ACATAAAACGATAAATCGGC	*F53H1.1* indel detection
cmo-JG56R	TTCCGTGACTCTTCCATTTC	
cmo-KG825R	CGCCGTTTTACTCTCTTT	
cmo17659	CGATTGGAAGTAGAGGTTCT	*gfp::csr-1 donor*
cmo17660	ATCATGATATTGACTATAAA	
cmd-25	gaactatacttttttcaggacttaactctgacatgGATTACAAAGACC ATGATGGTGACTATAAGGATCATGATATTGACTATA AAGACGATGACGATAAGGGTGGCGGAGAGAACCTC TACTTCCAATCGaaCcaAaaAcaGaaCccTagGctAgcAct Aaacatcttcgggcttgagctctctgagcgaacgat	ssODN donor to knock-in flagx3::linker::tev at *csr-1* locus

Cas9 Based Genome Editing

I. Materials:

1. *S. pyogenes* Cas9 3NLS (10 µg/µl, IDT)
2. tracrRNA (IDT)
3. crRNA (2 nmol or 10 nmol, IDT)
4. ssODN 4 nmol Ultramer (standard desalting, IDT)
5. PRF4::*rol-6(su1006)* plasmid (high quality Midi or Maxiprep)
6. SPRI paramagnetic beads (AMPure XP, Beckman Coulter)

Re-suspension (Stock Solutions):

1. Aliquot 0.5 µl (5 µg or 30 pmol) of Cas9 protein and store at -80°C (avoid freeze/thaw cycles)
2. tracrRNA – 0.4 µg/µl (18µM) in IDT nuclease free duplex buffer, store at -20°C (aliquots at -80°C)
3. crRNA – 0.4 µg/µl (34 µM) in TE PH 7.5 (IDT), store at -20°C (aliquots at -80°C)
4. ssDNA oligo donor – 1 µg/µl in ddH2O, store at -20°C
5. PRF4::*rol-6 (su1006):* 500 ng/µl, store at -20°C

II. Injection mixture preparation:

Add components of the injection mixture to the tube containing Cas9 in the following sequence:

17. Cas9 – 0.5 µl of 10 µg/µl stock (30 pmol)
18. Add tracrRNA – 5 µl of 0.4 µg/µl stock (90 pmol)
19. Add crRNA – 2.8 µl of 0.4 µg/µl stock (95 pmol) (if two guides are needed add 1.4 µl of each)

20. Pipette the mixture gently several times and incubate @37°C for 15 minutes. In our experience adding any double stranded DNA before RNP complex formation reduces HDR efficiency.
21. Add ssODN donor – 2.2 µl of 1 µg/µl stock (see note (3) below) (or)

    Add melted dsDNA – 500 ng (final concentration: 25 ng/µl for ~1kb donors or 45 fmol/µl)

22. Add PRF4::*rol-6 (su1006)* plasmid – 1.6 µl of 500 ng/µl stock
23. Add nuclease free water to bring the final volume to 20 µl and pipette gently several times.
24. To avoid needle clogging, centrifuge the mixture @14000rpm for 2 min, transfer about 17 µl of the mixture to a fresh tube and keep the tube on ice; proceed to loading the needles.

Notes:

1. *All the above steps in section II can be performed at room temperature*
2. *Aggregation is not an issue under these Cas9 concentrations.*
3. *Although we haven't explored the optimal dose range for ssODNs, given the efficiencies obtained with dsDNA at 25 ng/µl, much lower doses of ssODN could be used.*
4. *Final injection mixture can be stored at 4°C and re-used for several months (up to 6 months) without compromising efficiency; we have not yet tested mixes that are older than 6 months.*

Cas12a (Cpf1) Based Genome Editing

I. Materials:

1. A.s. Cas12a Ultra (10 µg/µl, IDT)
2. Cpf1-crRNA 21 bases long (2 nmol or 10 nmol, IDT)
3. ssODN 4 nmol Ultramer

4. PRF4::*rol-6(su1006)* plasmid (high quality Midi or Maxiprep)
7. SPRI paramagnetic beads (AMPure XP, Beckman Coulter)

Re-suspension (Stock Solutions):

1. Aliquot 0.5 µl (5 µg or 32 pmol) of Cas12a protein and store at -80°C (avoid freeze/thaw cycles)
2. Cas12a-crRNA – 40 µM in TE PH 7.5 (IDT), store at -20°C (aliquots at -80°C)
3. ssDNA oligo donor – 1 µg/µl in ddH2O, store at -20°C
4. PRF4::*rol-6 (su1006):* 500 ng/µl, store at -20°C

II. Injection mixture preparation:

Add components of the injection mixture to the tube containing Cas9 in the following sequence:

1. Cas12a – 0.5 µl of 10 µg/µl stock (32 pmol)
2. Add cas12a-crRNA – 2.5 µl of 40 µM stock (100pmol)
3. Add TE PH 7.5 – 3.0 µl
4. Pipette the mixture gently several times and incubate @37°C for 15 minutes
5. Add ssODN donor – 2.2 µl of 1 µg/µl stock (see note (3) below) (or)

   Add melted dsDNA – 500 ng (final concentration: 25 ng/µl for ~1kb donors or 45 fmol/µl)

6. Add PRF4::*rol-6 (su1006)* plasmid – 1.6 µl of 500 ng/µl stock
7. Add nuclease free water to bring the final volume to 20 µl and pipette gently several times.

8. To avoid needle clogging, centrifuge the mixture @14000rpm for 2 min, transfer about 17 µl of the mixture to a fresh tube and keep the tube on ice; proceed to loading the needles.

Notes:

1. *All the above steps in section II can be performed at room temperature*
2. *TE is added in step 3 for easier pipetting; by further diluting the crRNA stock this step can be omitted.*
3. *Although we haven't explored the optimal dose range for ssODNs, given the efficiencies obtained with dsDNA at 25 ng/µl, much lower doses of ssODN could be used.*

III. Donor Design and Generation

ssODN donors:

To generate ssODN donor, add 35 bases of 5′ homology sequence in front of the tag (or mutations) and 35 bases of the 3′ homology sequence at the end. Mutate the PAM site or the guide binding sequence if it is not already disrupted by the insert. If the guide binding sequence is mutated or if silent mutations are introduced between the guide cleavage site and the desired insertion site, length of homology sequence should be 35bp from the last mutation.

dsDNA donors:

Generate dsDNA donors by PCR either by using unmodified oligos or 5′ SP9-modified oligos.

1. Order unmodified (or 5′ SP9 modified) oligos with standard desalting (IDT); 35nt as homology arms and 20nt complementary to insert (eg: GFP). SP9 modifications are available at 100nmol scale from IDT.

2. Perform PCR with an insert-containing plasmid as the template for amplification; use High-Fidelity polymerase.
3. Run a few microliters of PCR on agarose gel to check if a single bright band is obtained. If non-specific amplification is observed, set up a temperature gradient and find the optimal temperature.
4. PCR clean-up: use one of the following three options depending on your experimental conditions.

    a. Purify the PCRs using spin-columns and elute DNA in 20 µl of nuclease free water. Generally, column purification is sufficient, and you may proceed to step 5. However, some primer pairs produce long (~80bp) primer dimers that may contain the entire homology arms. Spin-columns may not be able to remove dimers of this length completely. We found that these short “dimer donors” are preferentially used as templates over full-length donors with the desired insert (such as GFP). *Note*: Dimers may or may not be visible on the agarose gel.

    b. If dimer formation is a concern, use 0.6x SPRI beads (AMPure XP) to perform the clean-up instead of spin-columns. For example: add 60 µl of beads to 100µl of PCR, wash with 70% ethanol twice, elute in nuclease free water (refer to the bead manufacturer’s protocol for further details).

    c. If primer dimers are clearly visible on the gel, then it is best to gel-extract the DNA. However, gel extracted DNA can be toxic, presumably due to the presence of guanidine hydrochloride (component of binding buffer) in the final elute. To reduce salt contamination, incubate the column with wash buffer for 10 min before centrifugation; repeat washes 2-3 times. Strong absorbance at

230nm on Nanodrop suggests GuHCl contamination. For best results, gel-extracted DNA should be further purified with 1x to 1.5x AMPure XP beads (strongly recommended).

5. After purification, dilute a portion of dsDNA PCR donor to 100 ng/µl and transfer about 5.5 µl to a PCR strip tube and proceed to the heating step.

6. Heat to 95 ºC and cool to 4 ºC using thermal cycler (95 ºC-2:00 min; 85 ºC-10 sec, 75 ºC-10 sec, 65 ºC-10 sec, 55 ºC-1:00 min, 45 ºC-30 sec, 35 ºC-10 sec, 25 ºC-10 sec, 4 ºC-hold. Ramp down at 1 ºC/sec at every step).

7. Add melted donor DNA to rest of the injection mixture only after pre-incubating RNP complexes.

Note: we store purified donors at -20 ºC and melt them right before adding to the injection mix. We have not explored storage and re-use of melted donors.

IV. Micro-injection and Screening:

1. Inject 5 to 10 young adults and transfer them onto individual plates. If both arms of the hermaphrodite gonad are injected, a good injection should yield 20 to 40 F1 Rollers.
2. After about 72 hours post injection, score for number of F1 Rollers and choose 2 plates with the highest number of Rollers.

Note: We generally culture the injected animals at room temperature (~22°C-23°C).

3. a. For indels: choose 2 P0 plates that segregate the highest number of F1 Rollers; pick 12-24 F1 Rollers from these 2 plates and place them onto separate plates.

b. For ssODN-based editing: choose 2 P0 plates that segregate the highest number of F1 Rollers; pick about 24 F1 Rollers from these 2 plates and place them onto separate plates.

c. For dsDNA-based editing: Choose 2 plates that segregate the highest number of F1 Rollers and from these 2 plates, pick ~24 non-Rollers that are younger than Rollers and place them onto separate plates. Younger animals among the Roller cohorts can also be picked. For inexperienced injectors, we recommend using 5′ end-modified dsDNA donors and picking F1 Rollers.

4. To avoid false positives due to mosaicism in F1 animals, pick several F2s from each plate, perform pooled lysis and genotype. Genotyping primers should lie outside of the homology arms to avoid amplification from transiently retained donor molecules.

5. Alternatively, correct insertions of fluorescent tags can be screened under a fluorescence dissecting scope or by using high magnification fluorescence microscope. For high magnification screening, mount several F2 animals onto 2% agarose pads and immobilize with levamisole.

Results

Melting the donor dramatically stimulates HDR for longer edits

We recently showed that melting and reannealing donor molecules to create asymmetric donors with single-stranded homology arms can improve the frequency of CRISPR-mediated homology-directed insertions among transformants that were positive for a transformation marker[208]. Because transformation markers can cause confounding effects or toxicity, we decided to conduct an initial study in which markers were omitted altogether. For this purpose, we chose to target the insertion of *gfp* into the easily scored *glh-1* locus, which encodes a VASA-related DEAD-box protein that localizes robustly to germline perinuclear foci known as P granules or nuage.

We prepared the *gfp* donor by PCR using primers tailed with 35 nt of homology to the *glh-1* locus (Figure 4.1A). In order to separately analyze the consequences of melting and of generating single-stranded overhangs we prepared three types of donor, (i) PCR products that were never melted, “unmelted donors,” (ii) “melted donors” that were heated and allowed to reanneal, and (iii) “asymmetric melted donors” that were prepared by heating a mixture of two overlapping *gfp* PCR products (one with 35-bp homology to *glh-1* at each terminus and one without[208]. For simplicity, we refer to denaturing and quickly cooling the donor as "melting," (see Methods). We injected each type of donor along with Cas9-guide-RNPs targeting *glh-1* into the core cytoplasm of the pachytene syncytium just distal to the gonad turn. Ideal injections result when the flow of the

injection solution extends bilaterally from the injection site into the queue of oocytes at the proximal end and into mitotic region at the distal end[214]. Only animals with two such injections—one per arm—were analyzed.

As previously shown[208] the asymmetric melted donor outperformed the unmelted donor. The asymmetric *gfp::glh-1* donor yielded 381 GFP-positive transformants among 900 F1 progeny, or 42% of total post injection progeny. The unmelted symmetric donor in contrast yielded half as many edits, 161 GFP-positive transformants among 740 post-injection progeny, (22%). Surprisingly, the symmetric melted donor was just as effective as the asymmetric melted donor, yielding 331 GFP positives among 906 F1 progeny, (37%). Thus, when the entire brood is scored melted symmetric donor was as effective as its asymmetric counterpart. For melted donors, the number of GFP positive edits equaled approximately two-fifths of all post injection progeny exceeding the total number of Roller transgenics typically recovered per injected animal (Figure 4.2 and see below).

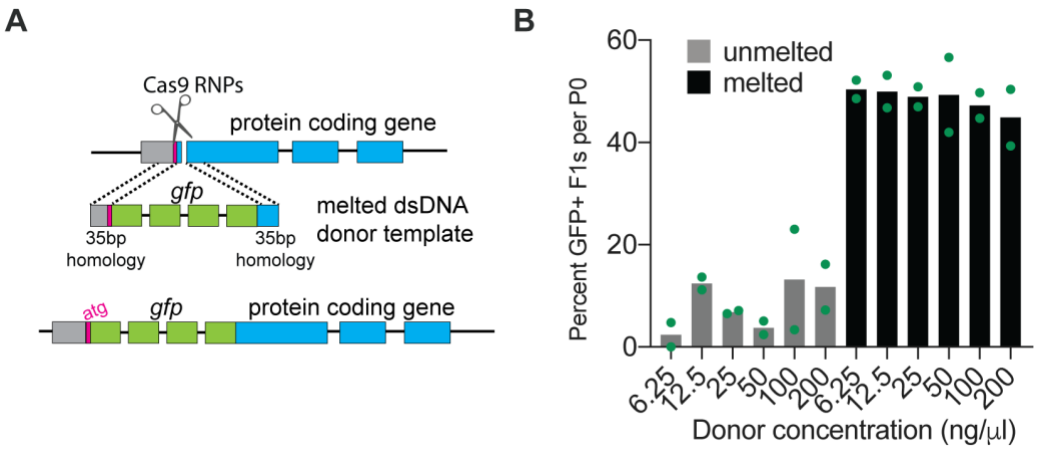


Figure 4.1. Melting dsDNA donors potentiates homology directed repair in *C. elegans*. (**A**) Schematic representation to insert *gfp* at the N-terminus of a protein coding gene immediately down stream of start codon (atg) using symmetric melted dsDNA donors and Cas9-guideRNA ribonucleoproteins (RNPs) is shown; grey segment represents sequence upstream of the start codon. Precise repair (HDR) enables fluorescent protein expression. (**B**) HDR efficiencies at the *glh-1* locus using symmetric unmelted (grey bars) or melted donors (black bars) with *rol-6* injection marker at indicated concentrations (n=2 broods) is plotted as percentage of F1s expressing GFP per injected animal (P0).

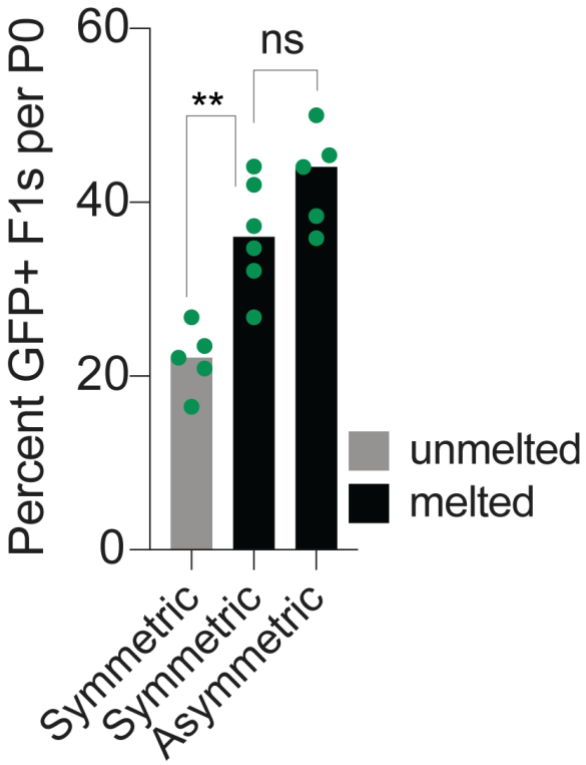


Figure 4.2: Melted dsDNA donors promote homology directed repair. HDR efficiencies at the glh-1 locus using symmetric (unmelted or melted) and asymmetric donors (n=5 or 6 broods) without rol-6 injection marker. Each data point (green) represents the percentage of animals expressing GFP among F1s scored per brood. Bars represent median. P-values (**, 0.0087 and ns, 0.1255) were determined by Mann Whitney test (unpaired, non-parametric, two-tailed)

Efficient HDR occurs over a broad range of donor concentrations

To explore how the frequency of *gfp* edits varied over a range of donor concentration, we injected unmelted or melted *gfp::glh-1* donor at concentrations of 6.25 ng/µl, 12.5 ng/µl, 25 ng/µl, 50 ng/µl, 100 ng/µl and 200 ng/µl (25 ng = 0.04 pmol). In order to control for injection quality, each injection mix included 40 ng/µl of the *rol-6(su1006)* co-transformation marker. For each donor mix, we injected 5 to 7 worms, singled those receiving optimal bilateral injections, and further analyzed two worms that made at least 100 post-injection progeny, including at least 20 Rollers. We then screened all the post-injection progeny—Roller and non-Roller—for germline GFP expression. We noted that the overall percentage of *gfp* insertions per injected animal (40%–50% for melted donors) (Figure 4.1B) was similar to levels achieved when the *rol-6* marker was omitted (Figure 4.2), suggesting that the *rol-6* marker does not interfere with the overall efficiency of editing. Surprisingly, the frequency of GFP-positive progeny per injected animal remained similar over a 32-fold range of donor concentrations. Melted donors consistently outperformed unmelted donors at every concentration (Figure 4.1B). These results suggest that, even at the donor concentration of 6.25 ng/µl, the HDR efficiency may be near saturation. At donor concentrations above 25 ng/µl, the frequency of Rollers per injected

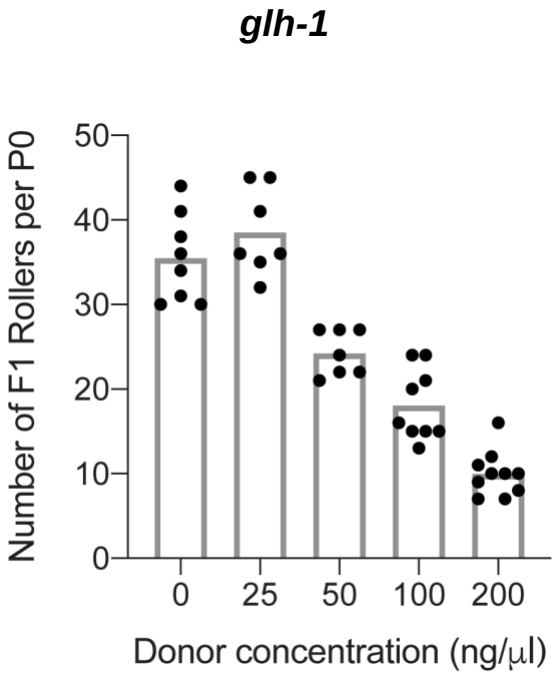


Figure 4.3. **High doses of donor DNA reduce the number of Rollers**. Injection mixes contain Prf4::rol-6 (su1006), Cas9 protein, crRNA targeting glh-1 locus and gfp::glh-1 dsDNA donor with ~35bp homology arms at indicated doses. Each dot represents the number of F1 rollers obtained per P0 animal and the bar represents mean; (n=7 to 10 broods per condition). dsDNA donors were not melted.

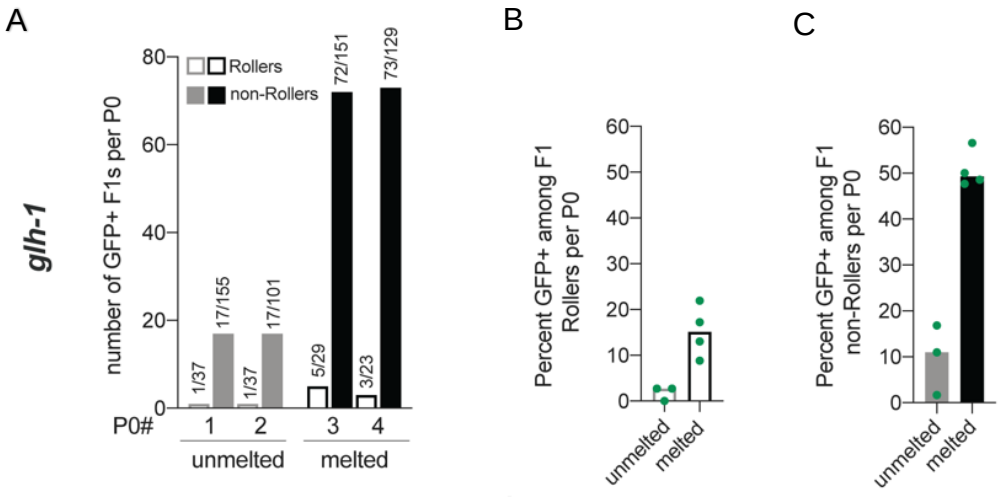


Figure 4.4. **HDR efficiencies are improved with melted donors**. Using unmelted and melted donors, HDR efficiencies at the *glh-1* locus is plotted as (**A**) number of fluorescence+ animals among Rollers and non-Rollers from two representative broods. Percentage of animals expressing fluorescence among, (B**)** Rollers and (**C**) non-Rollers, is plotted as percentage (n = 3 or 4 broods) for *glh-1* locus.

animal declined, suggesting that these higher concentrations cause toxicity (Figure 4.3). Taken together, these findings suggest that melted donors provide high rates of HDR with low toxicity over donor concentrations in the range of 6.25 ng/µl (0.01 pmol/µl) to 25 ng/µl (0.04 pmol/µl). Based on these findings we chose to use 25 ng/µl of donor in further investigations.

We next wished to examine how editing efficiencies vary among the Roller and non-Roller cohorts of post-injection progeny. We found that melted donors out-performed unmelted donors in both Roller and non-Roller cohorts (Figure 4.4), yielding several dozen *gfp* edited progeny per injected animal (as shown in two representative broods, Figure 4.4A). Strikingly, the fraction of GFP expressing progeny was much higher among non-Rollers (49%) (Figure 4.4C) compared to Rollers (15%) (Figure 4.4B).

To confirm the generality of these findings, we targeted two additional germline-expressed genes, *csr-1* and *znfx-1* (Figure 4.5, A-C). In both cases, melted donors consistently outperformed unmelted donors for *gfp* and *mCherry* insertions respectively (Figure 4.5, A and D). When melted donors were used, the fraction of animals with precision insertions was approximately ~10-fold higher than levels obtained with unmelted donors. This enhancement was observed in both the Roller (Figure 4.5, B and E) and non-Roller cohorts (Figure 4.5, C and F). We also explored whether melted donors were beneficial for editing with Cas12a (CPF1)[223] RNPs — which recognize an AT rich TTTV protospacer adjacent motif (PAM) sequence. Indeed, Cas12a editing yielded high HDR efficiencies

comparable to those achieved with Cas9 RNPs for *gfp* insertion at both the *glh-1* and *F53H1.1* loci (Figure 4.6) (See methods section for detailed protocol).

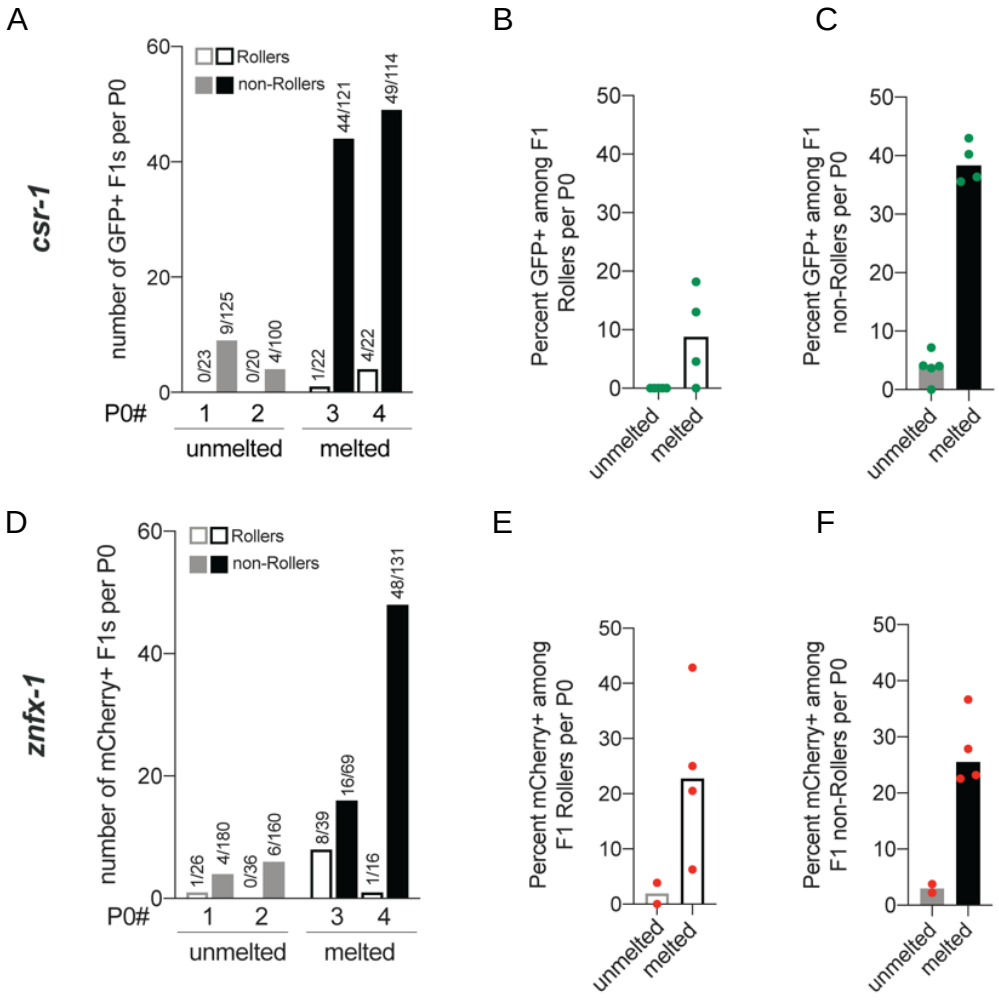


Figure 4.5. M**elted donors increase HDR efficiencies**. Using unmelted and melted donors, HDR efficiencies at the *csr-1 and znfx-1* loci are plotted as (A) number of fluorescence+ animals among Rollers and non-Rollers from two representative broods at *csr-1* locus. Percentage of animals expressing fluorescence among, (B**)** Rollers and (C) non-Rollers, is plotted as percentage (n = 3 or 4 broods) for *csr-1* locus. Similarly, improvement in fluorescent protein insertion efficiencies (HDR) with melted donors are shown for (D) representative broods (E) Rollers and (F) non-Rollers, at *znfx-1* loci. Each data point represents

the percentage of animals expressing fluorescent protein among F1s scored in each cohort per brood. Bars represent median. Number of fluorescence+ animals over number of animals scored is shown above the bars. Green dots represent GFP and red dots represent mCherry insertions.

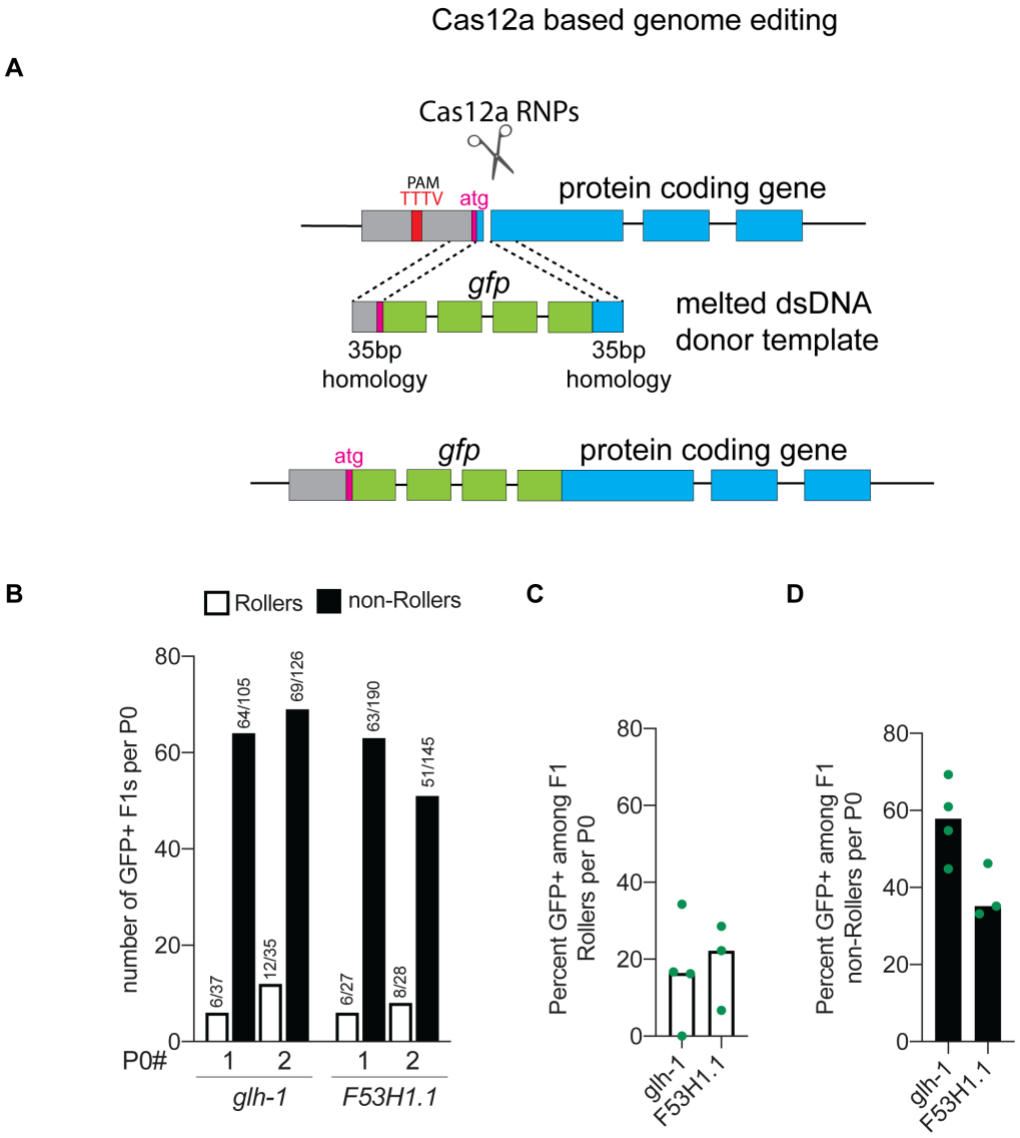


Figure 4.6. Precise genome editing using Cas12a nuclease and melted dsDNA donors. (A) Schematic representation of template dependent editing with Cas12a and melted dsDNA donor to insert gfp at the start codon (atg). Protospacer adjacent motif (PAM) for Cas12a system is TTTV, where V is A, C or G. HDR efficiencies at *glh-1* and *F53H1.1* loci are plotted as (B) number of GFP+ F1 animals among two representative broods, (C) percentage of GFP+

animals among F1 Rollers and (D) percentage of GFP+ animals among F1 non-Rollers; n= 3 or 4 broods. Number of GFP+ animals over number of animals scored are shown above the bars. Each data point represents the percentage of animals that are GFP+ among F1s scored in each cohort per brood and bars represent the median.

Editing efficiency peaks later in the brood after the Roller cohort of progeny are produced

The finding that HDR events are more prevalent among non-Roller progeny might reflect different developmental competencies of germ nuclei to form these distinct types of transgenics. For example, distal pachytene germ nuclei may be more receptive to recombination between the target chromosomal locus and the *gfp* donor, whereas more proximal germ nuclei may be more receptive to forming extra-chromosomal transgenes driven by recombination between co-injected DNA molecules (see Discussion)[214]. To examine these possibilities, we followed the production of Roller and GFP-positive progeny over the entire post-injection brood. Worms receiving 'ideal' bilateral injections of an editing mix prepared with melted *gfp*::*glh-1* donor (25 ng/µl) and *rol-6* co-injection marker (40 ng/µl) were cultured in two groups of 4 injected animals. Each group of animals was transferred every 4 hours to fresh plates to divide their broods into 12 segments over the next two days. Animals were transferred one more time on the third day (64 hours post-injection) thus dividing the progeny into 14 groups (Figure 4.6A). We then scored the frequency of Roller progeny and GFP-positive progeny in each segment.

Consistent with the idea that Roller extra-chromosomal transgenes assemble in more proximal germ cells, nearly 100% of the Roller progeny were produced within the first 28 hours post injection. The frequency of Rollers peaked between 8 and 12 hours post-injection where Rollers comprised 81% of the 47 progeny produced in the interval. The frequency of Roller progeny remained above

60% until 20 hours post injection, declining to ~30% then 13% over the next two 4-hour intervals. Rollers were virtually absent among

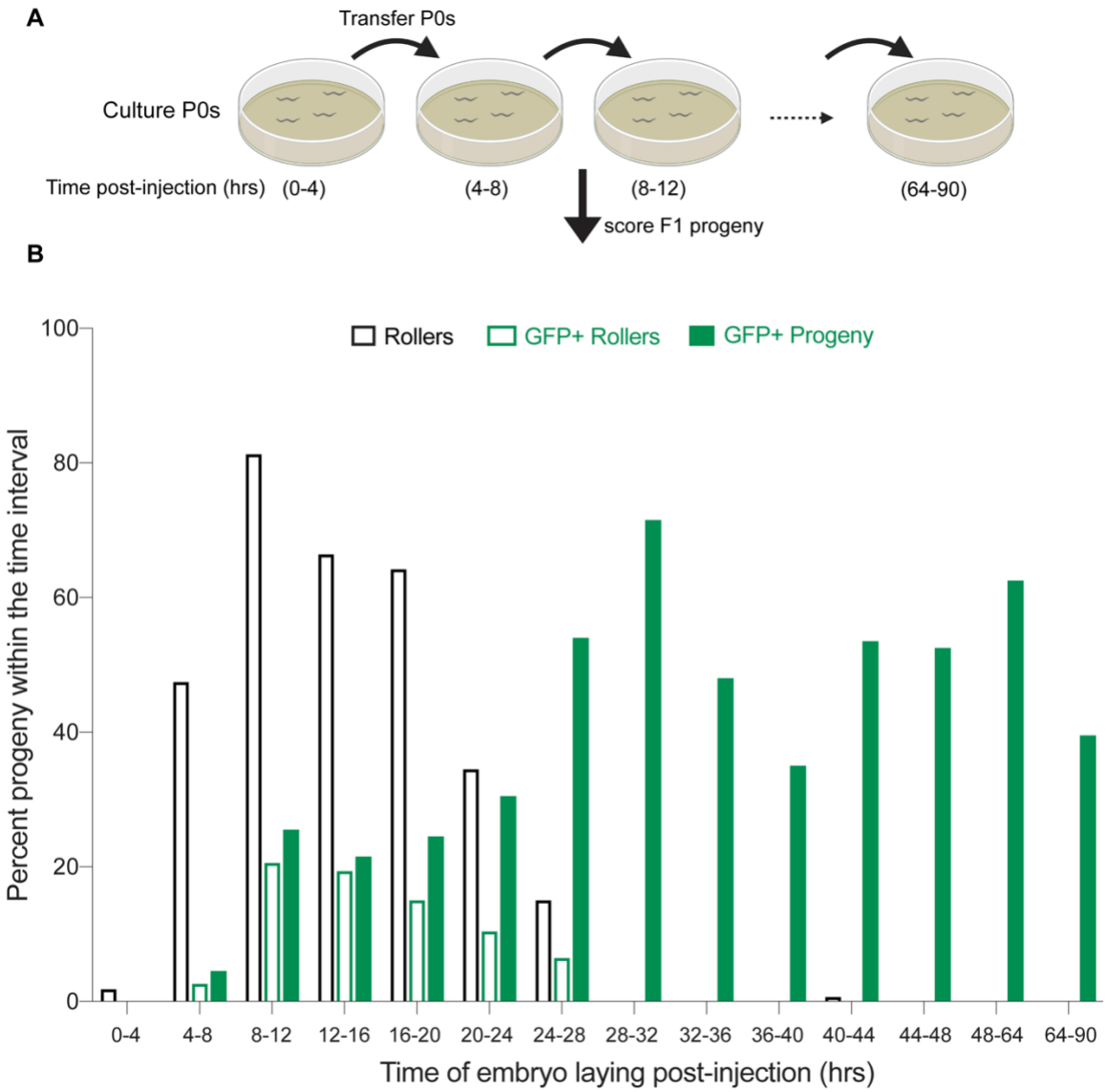


Figure 4.7. Editing occurs later in the brood after roller cohort. (A) Schematic representation of the experiment is shown. 4 Injected animals placed on a single plate were moved at indicated post injection time points and F1 embryos laid during the time-intervals were scored for GFP as adults. (B) Fraction of the progeny produced in each time window that are Rollers (open black bars), GFP+ Rollers (open green bars) and GFP+ progeny (Rollers and non-Rollers, solid green

bars) are plotted as percentage. Bars represent mean value of two replicates and each replicate consists of 4 P0 animals injected with 25 ng/µl of symmetric melted donors and 40 ng/µl of *rol-6* co-injection marker. Animals were cultured at 18°C-20°C.

progeny produced after 28 hours (Figure 4.6B). In striking contrast, the frequency of precision editing events was low within the first 24 hours and then appeared to plateau and remain high during the entire remainder of the brood (Figure 4.6B). For example, only 20% of the 306 progeny produced in the first 24 hours were GFP positive while an average of 54% were positive among the progeny produced thereafter (n=1327). Importantly, while GFP precision editing was less frequent within the first 24 hours (where Roller transgenics were found), precision editing was not under-represented within the Roller cohort. For example, we found that 24% of Rollers vs 20% of all animals produced in the first 24 hours were GFP positive (Figure 4.6B). Moreover, among GFP positive animals produced in this interval 60% were Rollers. Thus, the Roller marker positively correlates with *gfp* editing but does so within a cohort of progeny that precedes the optimal editing window for *gfp* insertion (See Discussion).

Donor purity is crucial for best HDR efficiencies

Although *rol-6* transformation precedes the optimal window of *gfp* insertion (as shown above), we nevertheless found that the *rol-6* marker provides a valuable troubleshooting metric (Figure 4.3). For example, while attempting to knock-in *gfp* at two different loci (*hrde-1* and *F53H1.1*), gfp insertions were unexpectedly rare. These experiments were conducted using melted TEG-modified donors[207], which typically yield as many as 100 GFP+ progeny per injected worm. However, despite

ideal injections that produced high numbers of Roller progeny, only 2 (average) Rollers were GFP positive per brood (spin-column, Figure 4.8, A and D). Scoring entire broods for GFP, we only obtained a maximum of 18 (*hrde-1*) (Figure 4.8A, P0# 2) and 13 (*F53H1.1*) (Figure 4.8D, P0#s 1 and 2) GFP-positive progeny per injected worm. The fraction of Rollers (spin-column, Figure 4.8, B and E) or non-Rollers (spin-column, Figure 4.8, C and F) expressing GFP stayed below 8% at both the loci. Because the number of Rollers per injected animal was near the optimal range, we reasoned that the injection quality was good, injected animals were healthy, and the injection mixture was non-toxic.

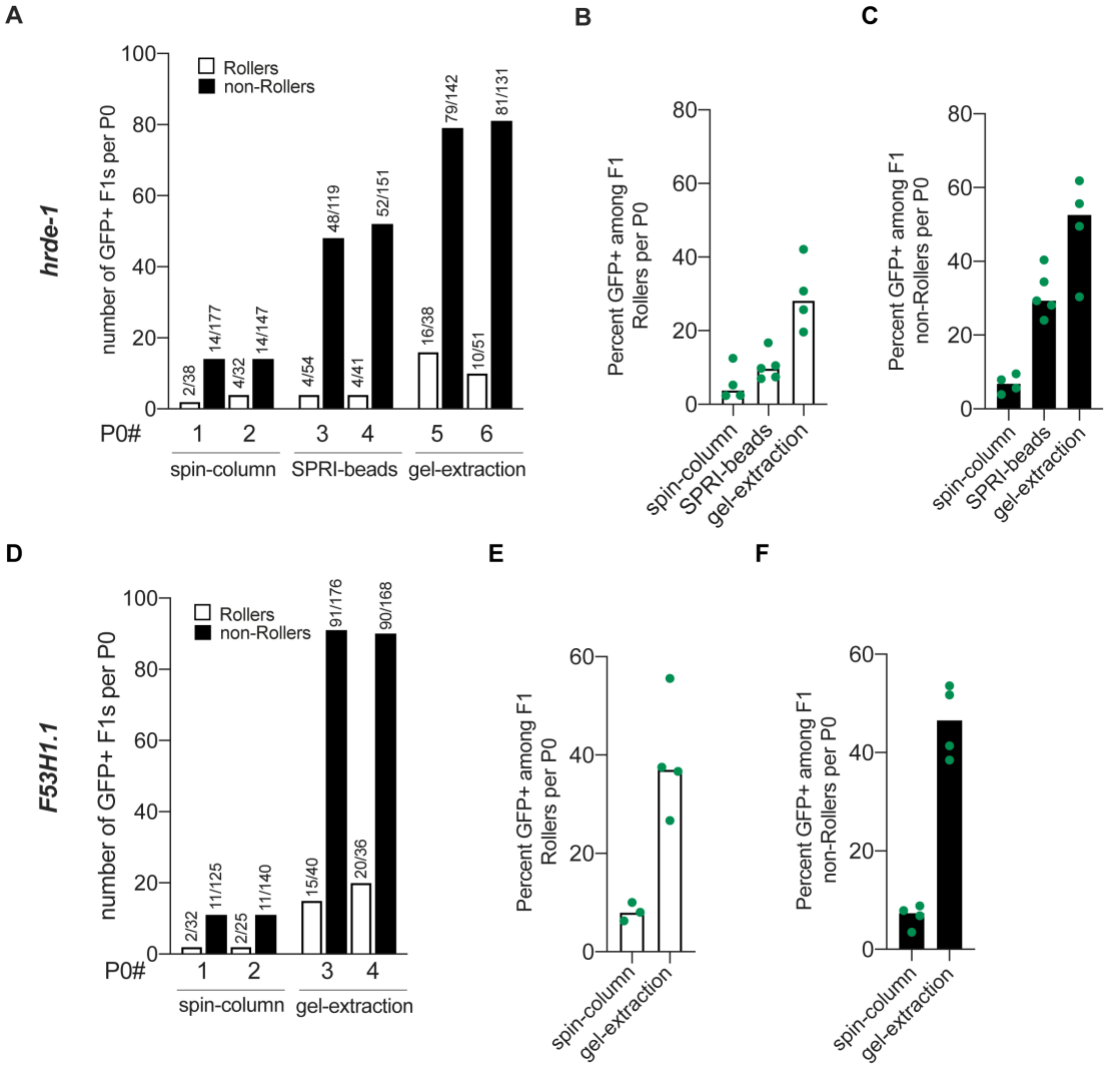


Figure 4.8. Purity of donor DNA is crucial for best HDR efficacy. HDR efficiencies of donors prepared by different methods of purification are plotted for *hrde-1* and *F53H1.1*. (A) Number of GFP+ F1 Rollers and non-Rollers from two representative broods are plotted. GFP+ animals among, (B) Rollers and (C) non-Rollers is plotted as percentage of animals scored in each cohort per brood. Similarly, (D-F) HDR efficiencies are plotted for *F53H1.1* locus. All the donors were

5′ TEG-modified and melted. Gel-purified donors were further cleaned-up with SPRI beads.

To understand why editing was so infrequent, we sequenced the target site in 25 randomly selected F1 rollers. In 21 of 25 Rollers, we identified non-wildtype sequences at the target site (Figure 4.9A), indicating that double-strand breaks were not the limitation. Importantly, none of these 21 Rollers contained *gfp* insertions (Figure 4.9A). Upon reading the sequencing trace, we found that 13 F1 animals contained a 15-bp insertion precisely where GFP sequences should have inserted (Figure 4.9, A and B). To our surprise, this short sequence perfectly matched a segment of the PCR oligo sequences (Figure 4.9C), and thus could be explained by insertion of a primer fragment or primer-dimer that was produced inadvertently during donor preparation. To test this possibility, we purified the *gfp* donors by size-exclusion using SPRI paramagnetic beads or by gel-extraction. Purifying the *hrde-1* donor with SPRI paramagnetic beads (optimized to exclude fragments smaller than 300 bp) modestly increased the percentage of GFP-positive progeny to 10% of F1 Rollers (n=212; Figure 4.8B) and 32% of non-Roller progeny (n=625; Figure 4.8C). By contrast, gel-purified *hrde-1* donor dramatically increased the percentage of GFP-positive progeny to 29% of F1 Rollers (n=163; Figure 4.8B) and 49% of non-Rollers (n=538; Figure 4.8C), with as many as 95 GFP-positive progeny from one injected worm (Figure 4.8A, P0#5). Similar results were obtained after gel purification of the *F53H1.1* donor (Figure 4.8, D-F). These findings demonstrate the utility of the Roller marker as a metric for troubleshooting the editing protocol and reveal the importance of removing PCR-based contaminants from donor preparations to achieve best knock-in efficiencies.

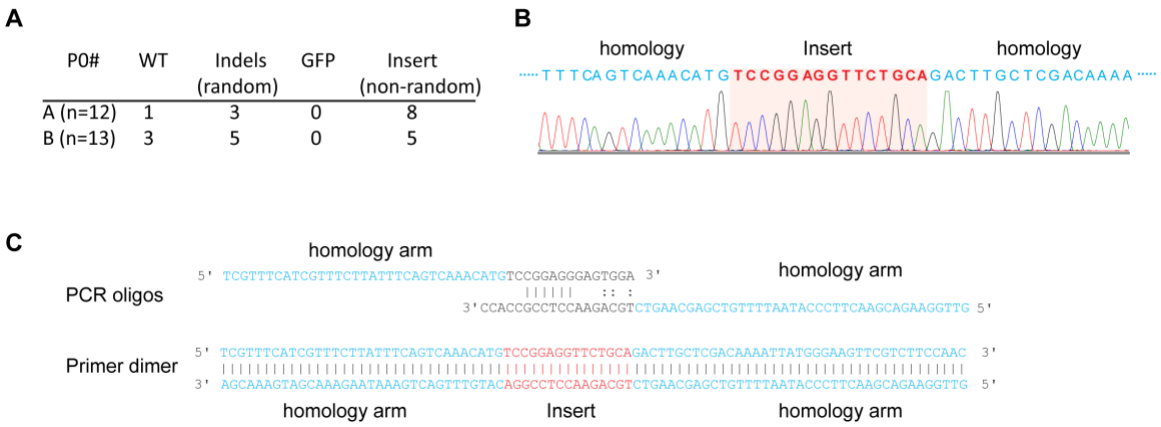


Figure 4.9. Short PCR contaminants outcompete dsDNA donors and integrate into the target locus efficiently (A) Insertions and deletions identified at *hrde-1* target site in F1 Rollers from two P0s (spin-column), (B) sanger sequencing trace of the 15bp non-random insert for a homozygous F2 animal. Partial homology arms of the donor are shown in blue and the sequence that got inserted into the genome is shown in red. (C) Schematic representation of predicted primer dimer formation is shown with 6bp perfect match and mismatched 3′ tails. Part of each oligo that is homologous to the PCR template plasmid is shown in black (linkers on either end of *gfp*) and the homology arms are shown in blue and the sequence (Insert) that would get inserted through HDR is shown in red. All the donors were 5′ TEG-modified and melted. Gel-purified donors were further cleaned-up with SPRI beads.

Discussion

We initiated these investigations to explore why long (~1-kb) DNA donors were less efficacious than short single-stranded oligonucleotide (ssODN) donors in *C. elegans.* We have shown that melting the donor DNA dramatically enhances precision editing, enabling efficient editing with shorter homology arms and at significantly lower donor DNA concentrations than previously recommended[208,211,212]. We show that as many as 100 precisely edited progeny can be obtained from a single injected animal, an editing efficiency of nearly 50% of post injection progeny, and far exceeding the typical frequency of progeny transformed with simple extrachromosomal arrays (Figure 4.3)[214].

Importantly, whereas the production of Roller transgenic progeny peaks during the first 24 hours post-injection, *gfp* edits peak after 24 hours and remain high through the remainder of the brood. Previous studies also reported that most *gfp*-edited animals are produced on the second day after injection[224]. These findings suggest that developmental differences between distal (less mature) and proximal (more mature) germ nuclei may favor formation or acquisition of distinct transgene types. For example, perhaps the large *rol-6* plasmid molecules are excluded from germ nuclei, and instead rapidly assemble into cytoplasmic extrachromosomal arrays that are swept by the germ plasm into developing oocytes, and only enter nuclei after fertilization (as previously suggested[214]). A size limitation on nuclear uptake may explain why we and others have found that donors over 2 kilo-basepairs yield few editing events (unpublished results)[210,219].

The observation that *gfp* editing peaks later, approximately 28 hours post injection, and then remains high, suggests either that proximal germ nuclei tend to exclude the donor, or that the pachytene nuclei are more receptive to recombination. Each gonad arm of a young adult worm contains hundreds of meiotic nuclei at the time of injection and each injected animal produces less than 200 post-injection progeny therefore it is likely that all the post-injection progeny come from the pachytene nuclei. Furthermore, based on an ovulation rate of 23 minutes[225], the appearance and persistence of GFP-positive progeny is consistent with editing in nuclei that were in pachytene (i.e., undergoing meiotic recombination) rather than in the mitotic zone at the time of injection. Whatever the reason for the HDR enhancement caused by melting the donor it is striking that the extrapolated rates of precision *gfp* insertion within these pachytene nuclei range as high as 70%.

Donor purity is crucial to achieve high knock-in efficiencies of long inserts. Contaminating primer dimers that contain homology arms can compromise HDR efficiency by integrating at the target site. Removing these contaminants by gel-extracting the donors dramatically increased *gfp* knock-in efficiencies. Similarly, as a time saving alternative to gel-extraction we found that purification using SPRI paramagnetic beads also improves HDR efficiencies, however using the optimal ratio of beads to PCR reaction was critical to removing the shorter contaminants (See Methods).

We do not know why melting the donor stimulates HDR. We obtained similar HDR rates across the entire range of donor concentrations, indicating that donor concentrations were saturated (or nearly so) at the lowest dose tested. Yet, melting the donor increased the HDR rate several fold at each concentration. Thus, melting stimulates recombination by acting on events or mechanisms that are independent of donor concentration. Conceivably, melting induces structural changes—e.g., denaturation bubbles caused by incomplete reannealing—that promote active nuclear uptake or directly stimulate repair. For example, single-stranded regions from incomplete re-annealing could promote strand invasion or act as damage signals that recruit trans-acting factors that facilitate HDR. Indeed, consistent with the idea that incomplete reannealing may be important preliminary studies utilizing slow cooling to promote better re-annealing, resulted in about half as many *gfp* insertions as fast cooled donors (Figure 4.10). However, further studies of this issue are required as the P-value in this initial study was not statistically significant. Interestingly, melting did not stimulate the already high HDR efficiency of a shorter 400 nt donor (Data not shown). Clearly more work is needed to fully explore and understand how donor-melting promotes HDR efficiency.

Undoubtedly, the high efficiencies of precision editing achieved here owe both to the easy access of worm pachytene germ cells to microinjection and to the remarkable receptiveness of these cells to HDR. A parallel study suggests that editing is enhanced even further when donor 5′ ends are modified with tri-ethylene glycol (TEG)[208]. Importantly, the combination of melting and TEG modifications,

which will be described in detail in the next chapter, increases the proportion of *gfp*-sized edits among the easily identified Roller progeny cohort by approximately twenty-fold from 1-2% to 20-40%. For experienced injectors, a single optimally injected animal can yield more than 100 GFP knock-ins (nearly two thirds of post-injection progeny), dramatically enhancing the ease and efficiency of genome editing. Given these high HDR efficiencies even researchers with little worm experience can now readily adopt this facile genetic animal model.

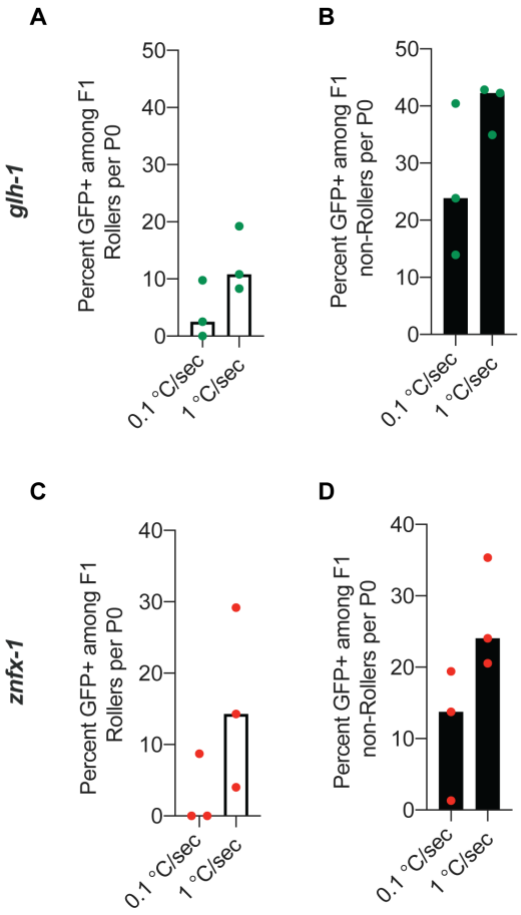


Figure 4.10 Quickly cooled donors act as better repair templates than slowly cooled donors. *gfp* (green dots) insertion efficiencies at *glh-1* loci are plotted for (A) Rollers and (B) non-Rollers using slow (0.1 ºC/sec) and quick (1 ºC/sec) cooled donors as percentage. (C and D) *mCherry* (red dots) insertion efficiencies at *znfx-1* locus. Each data point represents an F1 brood and bars represent median. Thermal cycler program for slow cooling: 95 ºC - 2:00 min; 85 ºC - 1:00 min; 75 ºC - 1:00 min, 65 ºC - 1:00 min, 55 ºC - 1:00 min, 45 ºC - 1:00 min, 35 ºC - 1:00 min, 25 ºC - 1:00 min, 4ºC- hold. Ramp down rate: 0.1 ºC/sec. P-value = 0.2 for panels

A, B, C and 0.1 for panel D (Mann-witney test, two-tailed). See Methods of chapter III for quick cooling conditions.

Acknowledgements. This work was funded by Howard Hughes Medical Institute (C.C.M.) and NIH R37 GM058800 (C.C.M). C.C.M is a Howard Hughes Medical Investigator. We thank Daniel Durning for technical assistance with preparation of *mCherry::znfx-1* injection mixes, Takao Ishidate for providing plasmid template for *linker-gfp-linker* donor PCRs and Darryl Conte, Jr. for critical review of this chapter's manuscript before publication.

Competing interests: Some of the findings described here are part of the patent applications filed by the University of Massachusetts Medical School on which the authors are inventors.

Preface to Chapter V

The contents of this chapter appeared in the following preprint

Ghanta, K.S., Chen, Z., Mir, A., Dokshin, G.A., Krishnamurthy, P.M., Yoon, Y., Gallant, J., Xu, P., Zhang, X.-O., Ozturk, A., et al. (2021). 5′ Modifications Improve Potency and Efficacy of DNA Donors for Precision Genome Editing. bioRxiv, 354480.

My contributions to this chapter are Figures 5.1, 5.2, 5.3 G-H, 5.4, 5.7- 5.14

Aamir Mir contributed to Figures 5.1, 5.2, 5.4, 5.5 and 5.6

Zexiang Chen contributed to Figures 5.3H, 5.4, and 5.13B

Gregoriy Dokshin contributed to optimization of conditions related to Figures 5.7 and 5.8

Ahmet Ozturk and Xiao-Ou Zhang contributed to bioinformatic analyses for Figure 5.14

Judith Gallant and Ping Xu contributed to microinjection and colony maintenance of mouse zygotes related to Figures 5.9, 5.10, 5.11 and 5.12

Yeonsoo Yeon contributed to design of mouse experiments

Pranathi Krishnamurthy contributed to generating chemically modified oligos

Chapter V: 5′ Modifications improve potency and efficacy of DNA donors for precision genome editing

Abstract

Nuclease-directed genome editing is a powerful tool for investigating physiology and has great promise as a therapeutic approach to correct mutations that cause disease. In its most precise form, genome editing can use cellular homology-directed repair (HDR) pathways to insert information from an exogenously supplied DNA repair template (donor) directly into a targeted genomic location. Unfortunately, particularly for long insertions, toxicity and delivery considerations associated with repair template DNA can limit HDR efficacy. Here, we explore chemical modifications to both double-stranded and single-stranded DNA-repair templates. We describe 5′-terminal modifications, including in its simplest form the incorporation of triethylene glycol (TEG) moieties, that consistently increase the frequency of precision editing in the germlines of animal models (*Caenorhabditis elegans*, mice) and in cultured human cells.

Introduction

Precision genome editing by HDR often requires cells to use exogenously supplied DNA templates (donors) to repair targeted double-strand breaks (DSBs). Maximizing precision genome editing, therefore, requires understanding both how cells respond to DSBs and to exogenous donors. These responses can be influenced by many variables, including cell-intrinsic factors (e.g., genetics, cell type, and cell cycle stage) and cell-extrinsic factors (e.g., donor length, strandedness, and chemistry)[132,136-139,145-147,157,226,227]. Each of these variables can influence the relative efficiency of HDR compared to competing DSB repair pathways, such as non-homologous end joining (NHEJ)[134,135,228,229].

In many organisms and cell types, high HDR efficiencies are readily achieved using short single-stranded oligodeoxynucleotide (ssODN) donor templates that permit single base changes or short insertions or deletions. However, HDR is frequently less efficient when longer double-stranded DNA (dsDNA) templates are used as donors. It is not known why longer DNA donors yield lower rates of HDR. In many cell types, high concentrations of dsDNA cause cytotoxicity, limiting the number of long donor molecules that can be safely delivered into cells. In addition, due to their size, long donor molecules may not transit the nuclear envelope as efficiently, reducing the effective concentration at the site of repair, or requiring cell division to gain access to the target locus. Moreover, end-joining ligation reactions assemble linear dsDNA molecules into

concatemers in eukaryotic cells[214,230-233], further limiting the number of individual donor molecules and their ability to diffuse to their DSB target sites.

In an effort to improve nuclear delivery and HDR efficacy, we incorporated 5′ modifications into the donor molecules, including a simple triethylene glycol (TEG) moiety, a 2′–*O*–methyl (2′OMe) RNA::TEG modification, and a peptide nucleic acid (PNA) comprising the SV40 nuclear localization signal (NLS) (see Methods). These 5′ modified donors increased the efficiency of templated repair by 2- to 5-fold in cultured mammalian cells as well as germline editing of *Caenorhabditis elegans,* and mouse *(Mus musculus)*. The modified donors exhibited a striking reduction in DNA ligation reactions including reduced self-ligation into concatemers and reduced sequence-independent ligation into cellular DSBs, suggesting that the 5′ modifications reduce the availability of 5′ ends for competing NHEJ reactions.

Material and Methods

Synthesis of PNA-NLS peptide. PNA oligomers were synthesized at 2µmol scale on Fmoc-PAL-PEG-PS solid support (Applied Biosystems) using an Expedite 8909 synthesizer. Fmoc/Bhoc-protected PNA monomers (Link Technologies) were dissolved to 0.2M in anhydrous *N*-methylpyrrolidinone. Amino acid monomers (Sigma Aldrich) and AEEA linker (Link Technologies) were dissolved to 0.2 M in anhydrous dimethylformamide. Coupling time was 8.5 min using HATU (Alfa Aesar) as activator; double coupling was performed on all PNA monomers and amino acids. PNAs were cleaved and deprotected by treating the resin with 400 µL of 19:1 TFA:*m*-Cresol for 90 min at room temperature. The resin was then removed with a PTFE centrifugal filter and PNAs were precipitated from cold diethyl ether and resuspended in deionized water. PNAs were purified by HPLC on a Waters XSelect CSH C18 5µm column at 60 °C, using gradients of acetonitrile in water containing 0.1% TFA, and were characterized on an Agilent 6530 Q-TOF LC/MS system with electrospray ionization. The PNA::NLS sequence used was GCGCTCGGCCCTTCC-[AEEA linker]-PKKKRK.

Synthesis of PEGylated oligos. PEG-modified oligonucleotides were synthesized using standard phosphoramidite methods on an ABI 394 synthesizer. Phosphoramidites were purchased from ChemGenes. Coupling times for 2′OMe-RNA and spacer phosphoramidites were extended to 5 min. Oligonucleotides were deprotected in concentrated aqueous ammonia at 55 °C for 16 h. Oligonucleotides were desalted using either Nap-10 (Sephadex) columns or Amicon ultrafiltration.

All the PEG-modified oligonucleotides were characterized on an Agilent 6530 Q-TOF LC/MS system with electrospray ionization. The 2′-OMe RNA sequence appended to the 5′-end of donor DNAs was GGAAGGGCCGAGCGC.

dsDNA Donor generation. Donor template sequences with the homology arms and the desired insert for knock-in (eg: gfp), were generated by PCR. PCR products were cloned into ZeroBlunt TOPO vector (Invitrogen, #450245) and plasmids were purified using Macherey-Nagel midi-prep kits (cat# 740412.50). Using the purified plasmids as templates and PEGylated oligos as primers, donor sequences were PCR amplified with Q5 (NEB, *C. elegans*) or Q5 or Phusion polymerase (NEB, mammalian). Before use in *C. elegans* microinjections, the resulting PEGylated PCR products were excised from 0.8-1% TAE agarose gel and purified using spin-columns (Omega, #D2501-02). For use in mammalian cells, the PEGylated long PCR products were purified using spin columns (Qiagen, # 28104) and short PCR products were gel-extracted (Omega, #D2501-02) and then purified again with Ampure XP beads.

Single Strand DNA donor generation. Long single stranded DNA donors were prepared using the protocol described by Li et al[234]. Briefly, the donor template containing the T7 promoter was amplified using standard PCR and purified using SPRI magnetic beads (Core Genomics). T7 *in vitro* transcription was performed using the HiScribe T7 High Yield RNA Synthesis kit (NEB) and the RNA was purified using the SPRI magnetic beads. Finally, the ssDNA donor was synthesized by TGIRT™-III (InGex) based reverse transcription using the

synthesized RNA as a template and a TEG-modified or unmodified DNA primer. We then performed base-treatment to remove RNA. The donor was again purified using SPRI beads.

Expression and purification of SpyCas9. The pMCSG7 vector containing the 6xHis-tagged 3xNLS SpyCas9 was a gift from Scot Wolfe at UMass Medical School. This construct was transformed into the Rosetta 2 DE3 strain of *E. coli* for protein production. Expression and purification of SpyCas9 was performed as described previously[39]. Briefly, cells were grown at 37°C to OD600 of 0.6, at which point 1 mM IPTG (Sigma) was added and the temperature was lowered to 18°C. Cells were grown overnight and harvested by centrifugation at 4,000 g. The protein was purified first by Ni^{2+} affinity chromatography, then by cation exchange and finally by size-exclusion chromatography.

Illumina sequencing (Mammalian cells)

Regions of interest were amplified from genomic DNA and sequenced on an Illumina MiniSeq platform. PCR1 ((98° C- 2min, 24 cycles of (98° C- 15sec, 64° C- 20sec, 72° C- 15sec), 72° C- 5min) was performed using 200ng gDNA, 1.25uL of 10uM forward and reverse primers that contain Illumina adapter sequences, 12.5uL NEBNext UltraII Q5 Master Mix, and water to bring the total volume to 25uL. PCR2 (98° C- 2min, 10 cycles of (98° C- 15sec, 64° C- 20sec, 72° C- 15sec), 72° C- 5min) was done using 1uL of unpurified PCR1 reaction mixture, 1.25 uL of 10uM forward and reverse primers that contain unique barcode sequences, 12.5uL

NEBNext UltraII Q5 Master Mix, and water to bring the total volume to 25uL. PCR2 products were first analyzed using 2% agarose gel electrophoresis, and then similar amounts were pooled based on the band intensities. Pooled PCR2 products were first purified by gel extraction (Qiagen) and purified again by PCR cleanup columns (Qiagen). Concentration of final purified library was determined by Qubit (High Sensitivity DNA assay). The integrity of library was confirmed by Agilent Tapestation using Agilent High Sensitivity D1000 ScreenTape kit. The library was then sequenced on an Illumina Miniseq platform according to the manufacturer's instructions using MiniSeq Mid Output Kit (300-cycles). Sequencing reads were demultiplexed using bcl2fastq2 (Illumina) and CRISPResso2[235] was used to align the reads and quantify editing efficiencies. Quantification window size was set as 30 to ensure the stringent analysis. HDR efficiency was calculated as percentage of (precise HDR reads) / (total reads).

Guide-Seq Experiment. Two phosphorothioate linkages were incorporated between the first three and the last three nucleotides in the dsODN tags. Unmodified dsODN does not contain any further modifications whereas modified dsODN contains 5′ TEG (SP9) modification (Integrated DNA Technologies). Sequencing libraries were prepared as previously described[236]. Data was processed and analyzed using the GUIDE-seq analysis software[236].

Cell culture and transfections. HEK293T cells were obtained from ATCC and were cultured in standard DMEM medium (Gibco, #11995) supplemented with 10% fetal bovine serum (FBS) (Sigma, #F0392). Human foreskin fibroblasts (HFF) were maintained in DMEM medium supplemented with 20% FBS. Chinese hamster ovary (CHO) cells (obtained from ATCC) were cultured in F-12K medium (Gibco 21127022) supplemented with 10% FBS, and K562 cells were cultured in IMDM medium (Gibco 12440053) supplemented with 10% FBS. Traffic Light Reporter Multi-Cas Variant 1 (TLR-MCV1) reporter cells were previously described[169]. Electroporations were performed using the Neon transfection system (ThermoFisher). SpyCas9 was delivered either as a plasmid or as protein. For plasmid delivery of Cas9 and sgRNA, appropriate amounts of plasmids were mixed in ~10 µl Neon buffer-R (ThermoFisher) followed by the addition of 100,000 cells. For RNP delivery of Cas9 (IDT), GFP-to-BFP assay (20 pmolCas9 and 25 pmol of crRNA-tracrRNA), EMX1-HEK293T (5pmol Cas9, 10pmol sgRNA (IDT)), EMX1-K562 (10pmol Cas9, 20pmol sgRNA), were mixed in 10 µl of buffer R. This mixture was incubated at room temperature for 30 minutes followed by the addition of 100,000 cells that were already resuspended in buffer R. This mixture was then electroporated using the 10 µl Neon tips. Electroporation parameters (pulse voltage, pulse width, number of pulses) were 1150 v, 20 ms, 2 pulses for HEK293T cells, 1650 v, 10 ms, 3 pulses for CHO cells, 1400 v, 30 ms, 1 pulse for HFF cells and 1600 v, 10 ms, 3 pulses for K562 cells. Electroporated cells were harvested for FACS analysis 48-72 hr post electroporation unless mentioned otherwise.

K562 GFP+ stable cell line generation. Lentiviral vector expressing EGFP was cloned using the Addgene plasmid #31482. The EGFP sequence was cloned downstream of the SFFV promoter using Gibson assembly. For lentivirus production, the lentiviral vector was co-transfected into HEK293T cells along with the packaging plasmids (Addgene 12260 & 12259) in 6-well plates using TransIT-LT1 transfection reagent (Mirus Bio) as recommended by the manufacturer. After 24 hours, the medium was aspirated from the transfected cells and replaced with fresh 1 ml of fresh DMEM media. The next day, the supernatant containing the virus from the transfected cells was collected and filtered through a 0.45 µm filter. 10 µl of the undiluted supernatant along with 2.5 µg of Polybrene was used to transduce ~1 million K562 cells in 6-well plates. The transduced cells were selected using media containing 2.5 µg/ml of puromycin. Less than 20% of the transduced cells survived, and these were then diluted into 96-well plates to select single clones. One of the K562 GFP+ clones was used for the analysis shown in this study. Cas9 was electroporated into the K562 GFP+ cells as RNP (20 pmol) with a crRNA targeting the *GFP* sequence. ssODN (66 nt) with or without end modifications was provided as donor template to convert the *GFP* coding sequence to the *BFP* coding sequence. % BFP (+) (HDR) and % GFP (-) BFP (-) (NHEJ) cells were quantified using flow cytometry.

Flow cytometry. The electroporated cells were analyzed on a MACSQuant VYB from Miltenyi Biotec. Cells were gated first based on forward and side scattering to select “live” cells and then for single cells. GFP-positive cells were identified using the blue laser (488 nm) and 525/50 nm filter whereas for the detection of mCherry positive cells, yellow laser (561 nm) and 615/20 nm filter were used. BFP-positive cells were identified using the violet laser (405 nm) and 450±50 nm filter.

Southern Blotting to visualize donor concatemers.

dsDNA donors (566bp) were prepared using DIG labeled dUTP nucleotide mix ((Sigma Aldrich # 11585550910). 1.5 pmol of gel-extracted DNA was nucleofected into HEK293T (100,000) cells (Cas9 or guideRNAs were not added to the mix). Nucleofected cells were collected at various time points and pellets were frozen at -80° C until processed for DNA extraction. Total DNA was extracted using buffered Phenol: Chloroform: Isoamyl Alcohol and quantified using Qubit (HS-DNA). Total DNA (genomic + exogenous) of 200ng (0 hr to 24 hr) or about 800ng (48 hr and 72 hr) was used for agarose gel (0.8%) electrophoresis. Higher amounts of DNA were loaded for the later time points to blot for roughly equal amounts of exogenous DNA and to account for the increase in total cell number over the time course. 200pg of 566bp and 800pg of 13kb DIG labelled PCR DNA were used as size markers. After electrophoresis agarose gel was treated with 0.25N HCl (depurination) for 10 min followed by three washes with distilled water. The gel

was then treated with denaturing solution (0.5M NaOH and 1.5M NaCl) for 20 min and another 30 min with fresh solution; followed by neutralization (2 washes 10 minutes each) with Alkaline transfer buffer (5xSSC with 10mM NaOH). Using Alkaline transfer buffer, DNA was then transferred for 3 hours with upward capillary action onto positively charged nylon membrane (Amersham Hybond N+, RPN303B). After transfer, membrane was soaked in 5xSSC for 10 min and UV crosslinked. Blots were then processed using DIG Wash and Block buffer set (Sigma Aldrich # 11585762001) according to the manufacturer's protocol. Briefly, membrane was blocked in 1x blocking solution with Maleic acid for 30 min, incubated with 1:20,000 Anti-Digoxigenin-AP, Fab fragments (Sigma Aldrich # 11093274910) in 1x blocking solution for 1 hour, washed twice with 1x wash buffer, incubated in 1x detection buffer and developed using CDP-star (Sigma Aldrich # 12041677001).

***C. elegans* microinjection and HDR screening.** Microinjections were performed using Cas9-RNPs as previously described[203]. dsDNA donors were generated by PCR; 25ng/µl of unmodified or end-modified dsDNA donors were used in each injection mixture. Donors were heated and quick-cooled as previously described[203]. Starting strain that is homozygous for 3XFLAG::GlyGlyGly::TEV::CSR-1 allele was used to knock-in *gfp* sequence between flag and glycine-linker. crRNA (CTATAAAGACGATGACGATA NGG) with PAM site in the glycine-linker and donor DNA with arms homologous to 35 bp

of 3xflag and 30 bp of 3xglycine-linker::tev flanking the gfp sequence were used. Loss of function WM702 (*eft3p::gfp(ne4807))* reporter strain was generated in EG6070 (oxSi221 [eft-3p::GFP + Cbr-unc-119(+)] II) strain background using CMG-48 and CMG-49 guides (See Table 5.1). *Rol-6 (su1006)* plasmid was used as co-injection marker. This marker plasmid forms episomal non-integrating extrachromosomal elements that transiently mark a subset of progeny by causing them to exhibit an easily scored Roller phenotype. Under the conditions used, high quality injections into both gonad arms yielded 20 to 40 Roller progenies from each injected animal. For each donor type entire F1 broods from four or more injected animals were scored and tabulated the total number of GFP positive progeny and the number of GFP positive Roller progeny.

Mouse Experiments

Strains and microinjection: All the mouse experiments were conducted according the UMMS Institute Animal Care and Use Committee (IACUC). C57BL/6J (Stock #000664) and Swiss Webster (Stock #SW) were obtained from Jackson Laboratory and Taconic respectively. All the animals were maintained in a 12 hr light/dark cycle. Superovulated females were mated, and their zygotes were collected at E0.5. Male pronuclei were injected with the injection mixtures described below. Finally, zygotes were transferred to pseudo pregnant recipients and allowed to go to term.

Donor preparation: Using plasmids as templates and either unmodified or end-modified oligos as primers, donor sequences were PCR amplified with Q5 polymerase (NEB). The resulting PCR products were excised from 0.8% TAE agarose gel and purified using spin-columns (Omega, #D2500). Gel-extracted DNA was further purified with 1.5X AMPure XP (Beckman Coulter) beads according to the manufacture's protocol and eluted in nuclease free water. Before use in microinjection mixes, dsDNA donors were subjected to heating and cooling protocol in thermal cycler as described previously[203].

Injection Mixture preparation: Injections mixes were prepared with the following final concentrations: S.p. Cas9 Protein (50 ng/µl) (IDT); S.p. Cas9 mRNA (50 ng/µl) (TriLink; L-7206); sgRNA (20 ng/µl) (IDT); dsDNA donor (1 ng/µl). Cas9 protein, sgRNA and TE (pH 7.5) were incubated at 37° C for 20 min. This mixture was then equally split into two tubes and the following components were added to each tube: Cas9 mRNA, dsDNA donor (either unmodified or 5′ 2′OMe-RNA::TEG modified), TE (pH 7.5) to bring the total volume to 50 µl. After pipetting well, the final injection mixtures were centrifuged at 14,000g for 2 min and 46 µl was taken from the top (to avoid particles that may clog the needles) and transferred to fresh tubes. All the steps were performed at room temperature. Mixtures were kept on ice and directly loaded into the needles for microinjection.

Genotyping: Tail clips of *Sox2-V5* founder animals were collected at P10, genotyped by PCR and Sanger sequenced to confirm precise insertion. To confirm

germline transmission, some of the HDR positive F0 animals were mated with WT animals and tail clips of F1 animals were genotyped.

Oligo Sequences. Sequences of all the guide RNAs used in this study are provided in Table 5.1 and sequences of all the oligos used are provided in Table 5.2.

Statistics. All the statistical analyses were performed using GraphPad Prism. The type of analysis performed, and the P-value information can be found in respective figure legends.

Data availability. All the data supporting the findings of this study are available within the paper and supplementary information. Any other data related to this manuscript are available upon reasonable request.

Table 5.1. Sequences of guide RNA spacers		
Name	**Guide sequence**	**Species**
Traffic Light Reporter_2.0	GAGACAAATCACCTGCCTCG	*eGFP*
GAPDH	gagagagaccctcactgctg	*H.spaiens*
SEC61B-1	CCCTCATCTCCAATATGGTA	*H.spaiens*
TOMM20	AATTGTAAGTGCTCAGAGCT	*H.spaiens*
EMX1	GAGTCCGAGCAGAAGAAGAA	*H.spaiens*
GFP-to-BFP	GCACTGCACGCCGTAGGTCA	*eGFP*
CMG-48 (*gfp*)	CCATCTAATTCAACAAGAAT	*C.elegans*
CMG-49*(gfp)*	CCTGAAAATTTAAATATGTA	*C.elegans*
CMG-67(gfp mutant(ne4807))	GTTGTCCTGTTGTTAGTTAG	*C.elegans*
CMG-33(Flag::linker::tev)	TATAAAGACGATGACGATAA	*C.elegans*

CMG-63 (*Sox2*)	TGCCCCTGTCGCACATGTGA	*Mus musculus*
CMG-89 (*Tyr*)	AACTGCGGAAACTCTAAGTT	*Mus musculus*

Table 5.2. Sequences of oligos	
Structure of 2'Ome-RNA::TEG oligos	2'OMeRNA(GGAAGGGCCGAGCGC) – TEG spacer – DNA(oligo)
Mammalian cell cultures	
TLR2.0-donor-F	GGGCCAAGAACAGATGGTCA
TLR2.0-donor-R	GGCGGATCTTGAAGTTCACC
Sec61-B-donor-F	GGGCCCACACTAAAGTTAGAG
Sec61-B-donor-R	GCGCCATTGGGATGTTCAG
TOMM20-donor-F	GACGCGTATTGGGATGATGA
TOMM20-donor-R	GCGCCATTGGGATACCTTAA
GAPDH-donor-F	CTCCTGCACCACCAACT
GAPDH-donor-R	TGGGGTTACAGGCGTGC
GFP_to_BFP_donor	GTGCCCTGGCCCACCCTCGTGACCACCCTGTCTCATGGAGTTCAGTGCTTCAGCCGCTACCCCGAC
05_emx1F_90bp	TGGCCCAGGTGAAGGTGTG
06_emx1R_90bp	GGTTGCCCACCCTAGTCATT
108_emx1_250hr_F	GCCCTGCCATCCCCTTCTGT
109_emx1_250hr_R	CCATTGCTTGTCCCTCTGTC
114_trac-250hr_F	GATAGCTTGTGCCTGTCCCT
115_trac-250hr_R	AGAACCTGGCCATTCCTGAA
EMX1 PCR1_F_NGS	ctacacgacgctcttccgatctGGCCTCCTGAGTTTCTCATCT
EMX1 PCR1_R_NGS	agacgtgtgctcttccgatctCAGCACTCTGCCCTCGT
gSeq_F_guideseq	G*T*TTAATTGAGTTGTCATATGTTAATAACGGT*A*T
gSeq_R_guideseq	A*T*ACCGTTATTAACATATGACAACTCAATTAA*A*C
C. elegans	

cmo17659_csr1 donor	CGATTGGAAGTAGAGGTTCT
cmo17660_csr1 donor	ATCATGATATTGACTATAAA
cmo-KG686F_eft3-gfp-donor	ATGAGTAAAGGAGAAGAACT
cmo-KG687R_eft3_gfp_donor	TATCACCTTCAAACTTGACT
Mouse	
cmo_KG993F_tyr-donor	AGGGGTGGATGACCGTGAGT
cmo_KG994R_tyr-donor	CTTATTCTTTTCGGAGACACTC
cmo_KG882F_tyr_genoyping	TTGTTGGCAAAAGAATGCTG
cmo_KG883R_tyr_genotyping	GCTTCATGGGCAAAATCAAT
cmo_KG884F_tyr_sequencing	GGATGGGTGATGGGAGTC
cmo_KG674F_sox2-v5donor	GCTGCGCCCAGTAGACTGCA
cmo_KG675R_sox2-v5donor	TCAGATTTTTCCTACTCTCC
cmo_KG823F_sox2F1-genotyping	ACATGATCAGCATGTACCTCC
cmo_KG824R_sox2R1-genotyping	TAATTTGGATGGGATTGGTGG

Results

End-modified DNA donors increase the efficiencies of HDR in mammalian cells

To examine the effects of donor end modifications on HDR in cultured mammalian cells, we took advantage of a modified traffic light reporter (TLR) comprising a "broken" GFP coding region followed by a frameshifted mCherry coding region[169,237]. Cas9 targets the "broken" GFP, which can only be made functional if precisely repaired by HDR, resulting in green fluorescence. If Cas9-mediated DSBs are imprecisely repaired by NHEJ, approximately one third of the imprecise repair events will restore the reading frame of mCherry, resulting in red fluorescence. Cas9 and single guide RNA (sgRNA) expression vectors and dsDNA donors with or without 5′ modifications were electroporated into HEK293T TLR cells (Figure 5.1A), followed by flow cytometry to determine the percentage of cells expressing either GFP or mCherry.

We first examined the performance of dsDNA donors modified with 15-nucleotide (nt) 2′OMe-RNA fused to triethylene glycol (RNA::TEG). Strikingly, the frequency of HDR increased with the amount of RNA::TEG-modified donor to a maximal 52% GFP+ cells at 1.2 pmol of donor before falling off at higher amounts of donor (Figure 5.1B). By contrast, a maximum HDR frequency of only 25% GFP+ cells was observed at 1.6 pmol of unmodified donor. Notably, 0.4 pmol RNA::TEG-modified donor was as efficient as 1.6 pmol unmodified donor, suggesting that the modified donor is ~4-fold more potent than the unmodified donor (Figure 5.1B).

The increase in GFP+ cells was accompanied by a corresponding reduction in mCherry+ cells (Figure 5.1C).

We next used the TLR assay to define features of the RNA::TEG moiety that promote maximal HDR. Nucleofection of 1.2 pmol donors modified with 2′OMe-RNA, TEG, or covalent RNA::TEG moieties all boosted HDR while reducing NHEJ events (Figure 5.1D and E). Increasing the length of the ethylene glycol moiety (3, 6, or 12 repeats) supported similar levels of HDR with or without the 2′OMe-RNA moiety (Figure 5.1F). Finally, donors with TEG modification at both 5′ ends yielded slightly better HDR efficiencies than donors with modification at only one of the two 5′ ends (Figure 5.1G). However, donors with RNA::TEG modification at both 5′ ends or at only one of the 5′ ends yielded similar HDR efficiencies (Figure 5.1G).

We reasoned that that the 2′O-Methyl RNA linker could be used to anneal PNA oligos attached to peptides that might enhance nuclear uptake. To test this idea, we produced complementary peptide-nucleic acid (PNA) oligos linked to a nuclear localization signal peptide or complementary PNA alone and tested these for HDR. Annealing these PNA oligos was well tolerated and did not diminish HDR, however neither did they enhance HDR (Figure 5.2A-D). Thus, further study will be needed to determine if RNA-TEG adapters can be used to append peptides or other molecules (e.g. CAS9 RNP) that stimulate HDR.

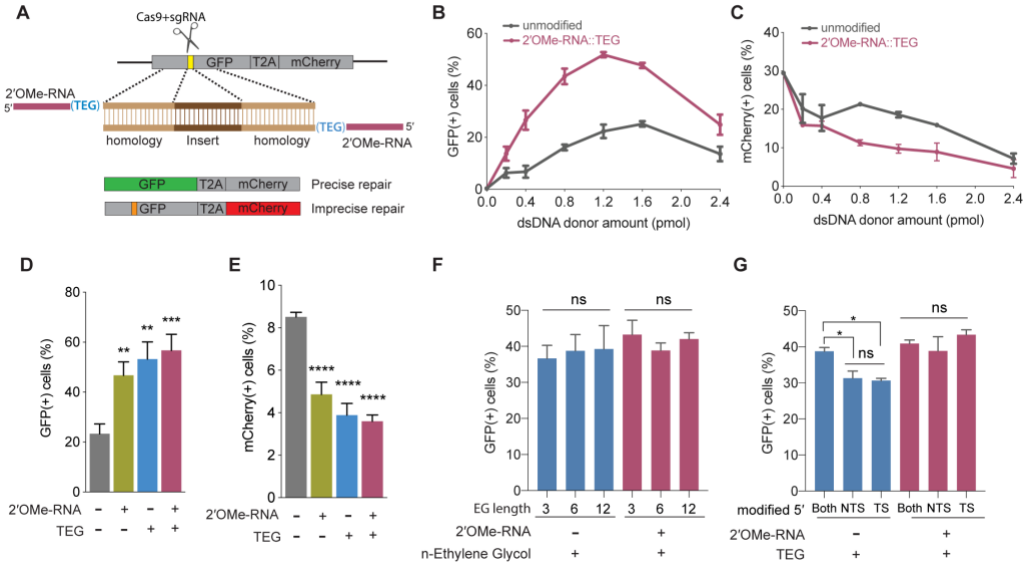


Figure 5.1. **5′ end-modified donors promote HDR in Traffic-Light Reporter (TLR) cells.** (A) Schematic showing the TLR assay to quantify HDR efficiencies using unmodified or end-modified dsDNA donors. Editing efficiencies plotted as percentage of (B) GFP+ (HDR) and (C) mCherry+ (NHEJ) HEK293T TLR cells at different amounts of unmodified, 2′OMe-RNA::TEG-modified dsDNA donors. Editing efficiencies plotted as percentage of (D) GFP+ (HDR) and (E) mCherry+ (NHEJ) HEK293T TLR cells at 1.2 pmol of dsDNA donors indicated. Percentage of GFP+ cells obtained with dsDNA donors modified with various lengths of ethylene glycol (F) and with modifications to only one end or both5′ ends of the donor. TS- target strand, NTS- non-target strand (G). Mean ± s.d for at least three independent replicates are plotted; two replicates for TEG-donor in panel G.

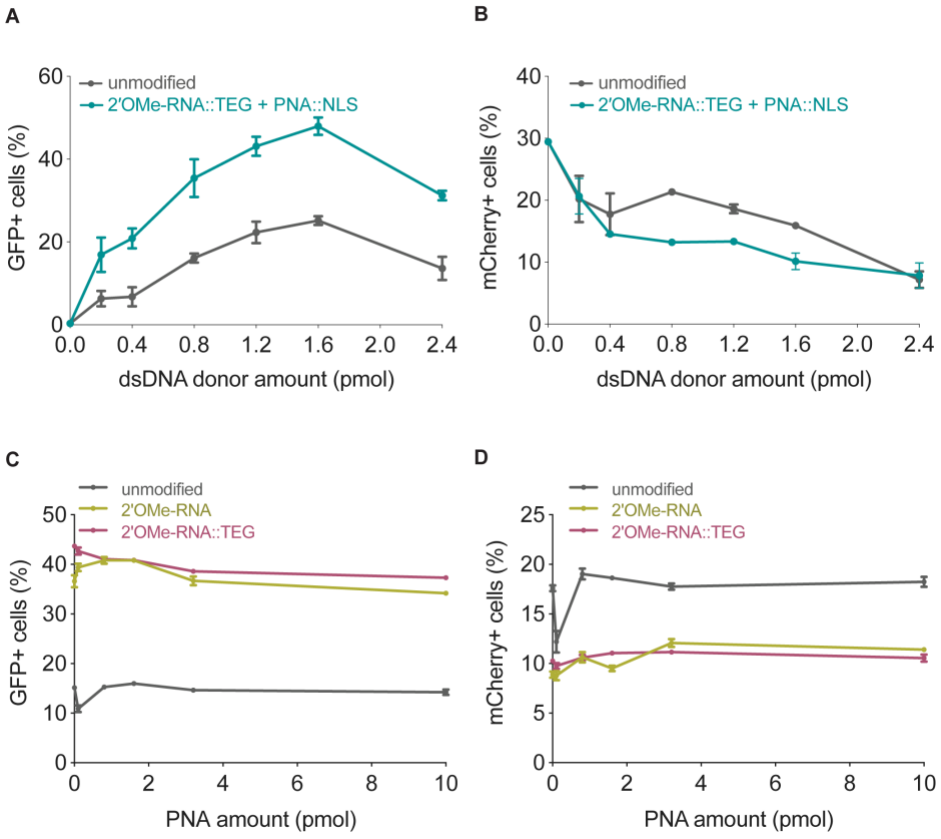


Figure 5.2. **2′OMe-RNA at 5′ ends of donors promote HDR in mammalian cells**. Editing efficacy plotted as percentage of (A) GFP+ (HDR) and (B) mCherry+ (NHEJ) HEK293T TLR cells at different amounts of unmodified, 2′OMe-RNA::TEG-modified and PNA::NLS-annealed dsDNA donors. Same unmodified controls are used in Figure 5.1 B and C. Addition of PNA (without NLS) to unmodified or end-modified donors does not further improve HDR efficiency in mammalian cells. 0.8 pmol of each type of donor was annealed to PNA (0.1 to 10 pmol). Editing efficiency was plotted as percentage of (C) GFP+ (HDR) cells and (D) mCherry+ (NHEJ) cells. Percentages were calculated by sorting the cells through flow cytometry (see Methods)

To explore the utility of TEG- and RNA::TEG-modified donors for repair at other genomic loci, we generated donors to integrate full-length *eGFP* at the endogenous *TOMM20*, *GAPDH*, and *SEC61B* loci (Figure 5.3A). We found that TEG or RNA::TEG donors consistently exhibited increased HDR levels in HEK293T cells as measured by the fraction of cells expressing eGFP at *TOMM20* (2-fold), at *GAPDH* (3-fold), and at SEC61B (5-fold) when compared to unmodified dsDNA donor (Figure 5.3B-D). RNA::TEG-modified donors also substantially increased HDR in two cell types that are less amenable to editing, increasing HDR at the *TOMM20* locus in human foreskin fibroblasts (HFF) cells (2.3-fold) and at the *Gapdh* locus in Chinese hamster ovary (CHO) cells (6-fold) (Figure 5.3E-F).

Next, to quantify the nature of repair outcomes (precise and imprecise), we employed deep sequencing assays. To facilitate sequencing across the repair site, we replaced a 12-nt sequence with a 9-nt sequence at the *EMX1* locus in HEK293T. We compared HDR efficiencies in this assay using unmodified, TEG-modified, and RNA::TEG-modified dsDNA donors with 90-base pair (bp) homology arms (Figure 5.3G). At 1.2 pmol and 2.4 pmol, RNA::TEG modified donors yielded two fold more precise edits compared to the unmodified donors. When even higher doses (5pmol) were used, the gap in efficacy between unmodified and RNA::TEG modified donors narrowed to just 16% (89.5% vs 72.8%) precise reads (Figure 5.3H). The *EMX1* donor with 90-bp homology arms also supported high levels of HDR in K562 cells across a broad dose range. Notably, low doses of donor supported higher levels of HDR in K562 cells than in HEK293T cells, suggesting

that K562 cells are more susceptible to editing (Figure 5.4). In this assay, donors modified with TEG alone exhibited no benefit over unmodified donors (Figure 5.3H and Figure 5.4). These results suggest that RNA::TEG modifications are more efficacious than TEG alone.

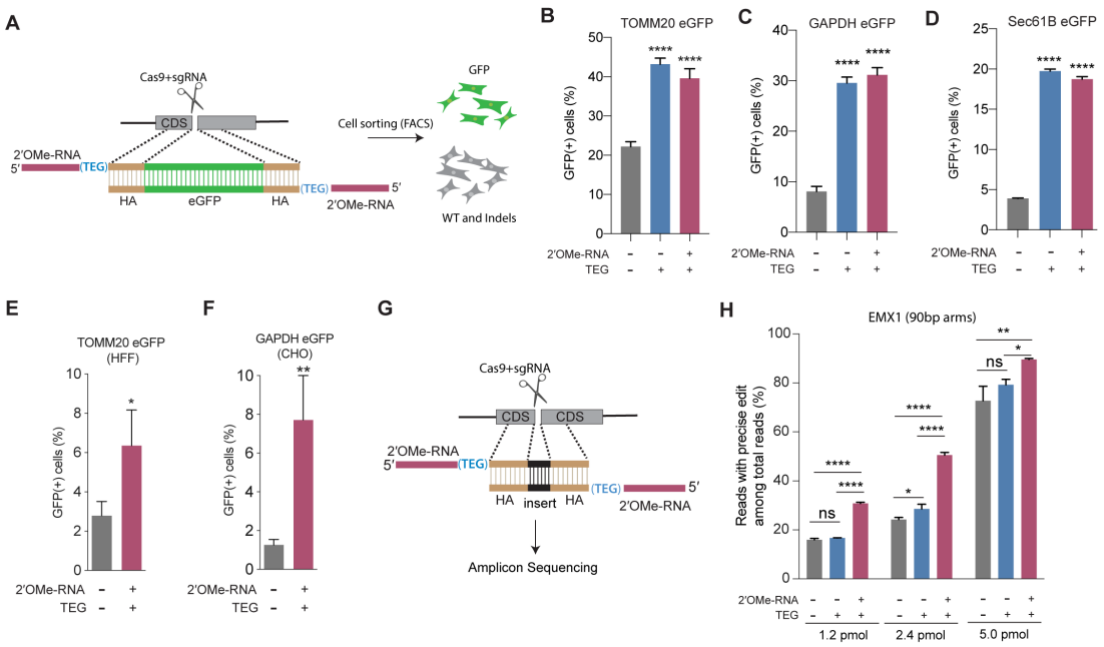


Figure 5.3 End-modified donors promote HDR at endogenous loci in mammalian cell cultures. (A) schematic representation of the 5′ modified donor design for eGFP insertion and FACS sorting is shown. Efficacy of eGFP integration at (B) *TOMM20* and (C) *GAPDH* (D) Sec61B loci in HEK293T cells using unmodified, TEG or 2′OMe-RNA::TEG-modified donors are plotted as percentage of GFP+ cells. Efficacy of eGFP integration at the (E) *TOMM20* locus in HFF (747 bp knock-in with ~1kb homology arms) and (F) *Gapdh* locus in CHO (1635 bp knock-in with ~800 bp homology arms) cells using dsDNA (500 ng) donors with and without 2′OMe-RNA::TEG modifications at the 5′ ends. (G) Schematic representation of the dsDNA donor design used for quantification with deep sequencing is shown. (H) Illumina sequencing reads with precise knock-in are plotted for dsDNA donors with 90bp homology arms at EMX1 locus in HEK293T cells. Mean ± s.d for at least three independent replicates are plotted.

P-values were calculated using one-way ANOVA and in all cases end-modified donors were compared to unmodified donor unless indicated otherwise (Tukey's multiple comparisons test; ****P < 0.0001; ***P < 0.001; **P < 0.01; *P< 0.05; ns-not significant).

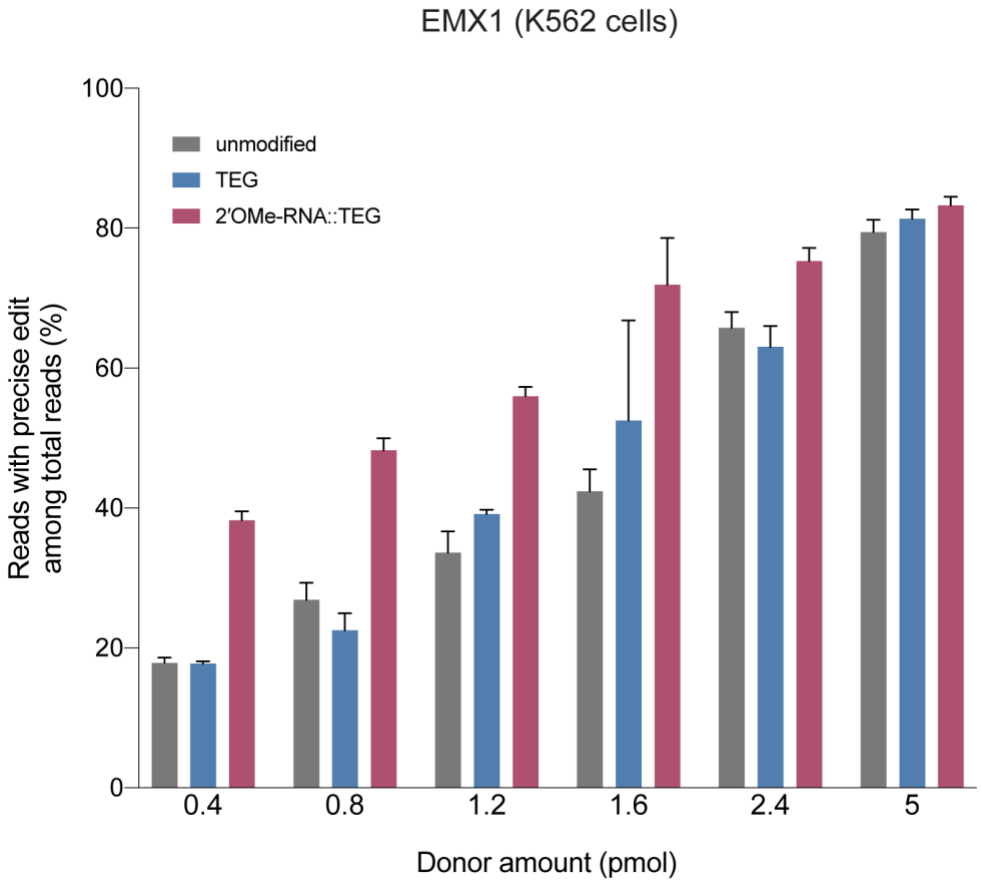


Figure 5.4. **RNA::TEG donors with short (90bp) homology arms are more potent than unmodified donors at EMX1 locus**. Fraction of precise reads is plotted as percentage of total Illumina reads obtained at various amounts of dsDNA donors into K562 cells. Cas9 RNPs and dsDNA donors were nucleofected into K562 cells and harvested after 3 days.

5′-modification increases potency of single-stranded DNA donors

The experiments described thus far employed dsDNA donors; however, long single-stranded DNA (ssDNA) or short single-stranded oligo deoxynucleotide (ssODN) donors are also widely used in many HDR editing protocols. We therefore decided to explore how 5′ end modifications affect single stranded donors of different lengths. Using the TLR assay, we found that addition of RNA::TEG at the 5′ end of a long (800-nt) ssDNA donor significantly boosted HDR compared to the unmodified ssDNA donor. The frequency of HDR increased with the dose of ssDNA donor, reaching maximal HDR (22.5% GFP(+)cells) at 6 pmol to 8 pmol donor amounts (Figure 5.5A, Figure 5.5B). The RNA::TEG-modified donor was greater than 4-fold more potent than the unmodified donor reaching a threshold of 16% GFP(+) cells at a concentration of approximately 2 pmol whereas achieving the same threshold of 16% required 8 pmol of unmodified donor (Figure 5.5A).

High yields of HDR in cultured mammalian cells have been achieved using short synthetic single-stranded oligo deoxynucleotide (ssODN) donors[156]. To test 5′-modified ssODNs for HDR efficacy, we used a sensitive GFP-to-BFP conversion assay in K562 cells. Precise editing converts a functional GFP sequence to blue fluorescent protein (BFP) sequence, producing cells that are GFP(-) and BFP(+). Imprecise editing produces cells that are both GFP(-) and BFP(-)[238]. Using 66 nt long ssODN donors and titrating the amount over a range of 0.01 to 40 pmol, we

found that RNA::TEG and unmodified donors produced similar maximal levels of HDR (47.5% to 52.8% BFP(+) cells). However, maximal HDR required 10-fold less RNA::TEG-modified ssODN than unmodified donors (Figure 5.5C). We also observed reduced levels of imprecise editing (GFP-negative and BFP-negative) as the frequency of HDR increased (Figure 5.5D). For both donor types, the decline in editing at higher doses correlated with the appearance of dead cells (data not shown), suggesting that dose-limiting toxicity scales with increased HDR potency.

The use of fully synthetic ssODN donors allowed us to explore additional modifications, including internal and 3′ modifications. Interestingly, 2′OMe-RNA, RNA::TEG, or TEG moieties at the 3′ terminus did not enhance HDR compared to unmodified ssODN, but they blocked the ability of 2′OMe-RNA, RNA::TEG, or TEG moieties at the 5′ end to enhance HDR (Figure 5.6). By contrast, HDR was neither enhanced nor impeded by phosphorothioate (PS) linkages placed at 5′ or 3′ terminal linkages at the doses tested (Figure 5.6). Taken together these findings suggest that the mechanism of HDR improvement requires an available 3′-OH.

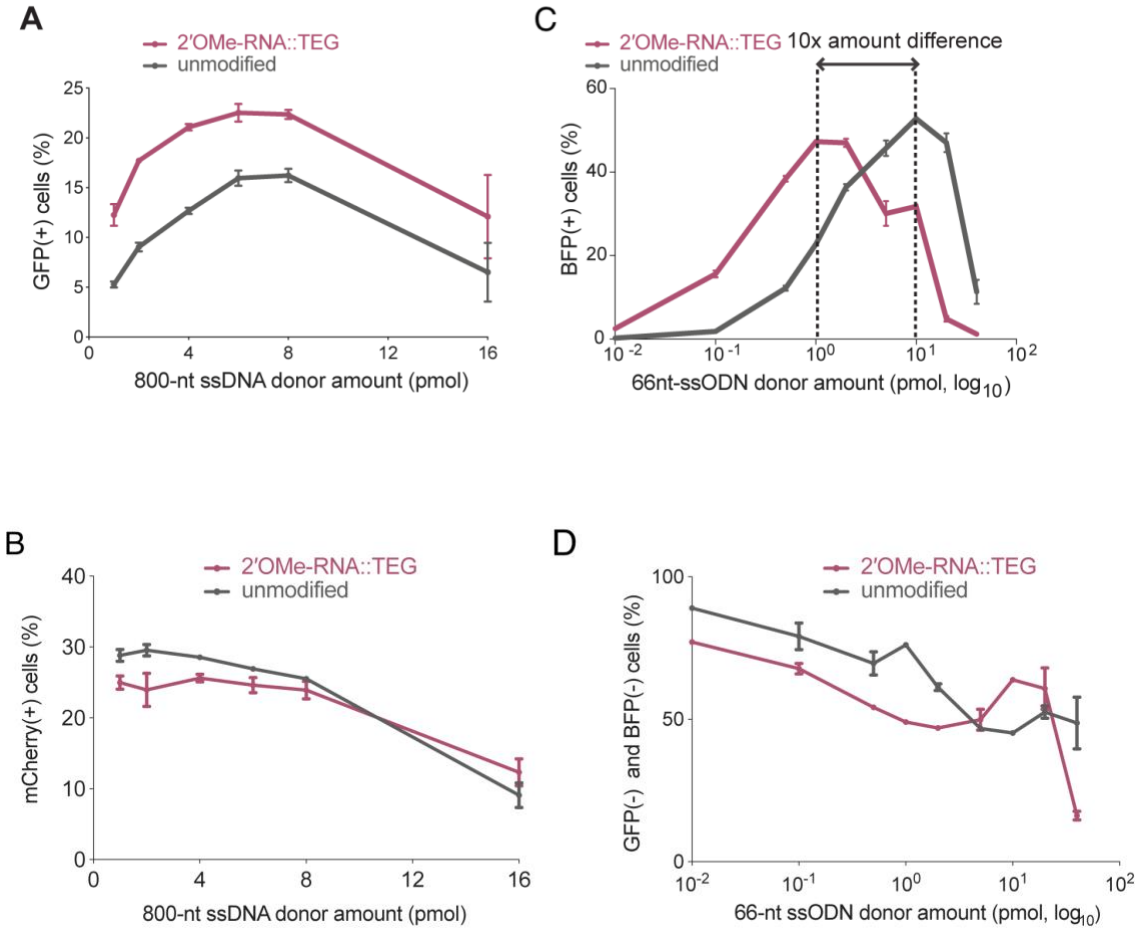


Figure 5.5. End-modifications increase potency of ssODN donors (A) Editing efficacy plotted as percentage of GFP+ (precise) and (B) mCherry(+) HEK293T TLR cells at different amounts of unmodified and 2′OMe-RNA::TEG-modified long ssDNA donors (800 nt). (C) Editing efficacy of GFP-to-BFP reporter conversion in K562 cells using different amounts of unmodified and 2′OMe-RNA::TEG-modified 66 nt ssODN donors plotted as percentage of BFP+ (HDR) and, (D) GFP(-) and BFP(-) (NHEJ) cells.

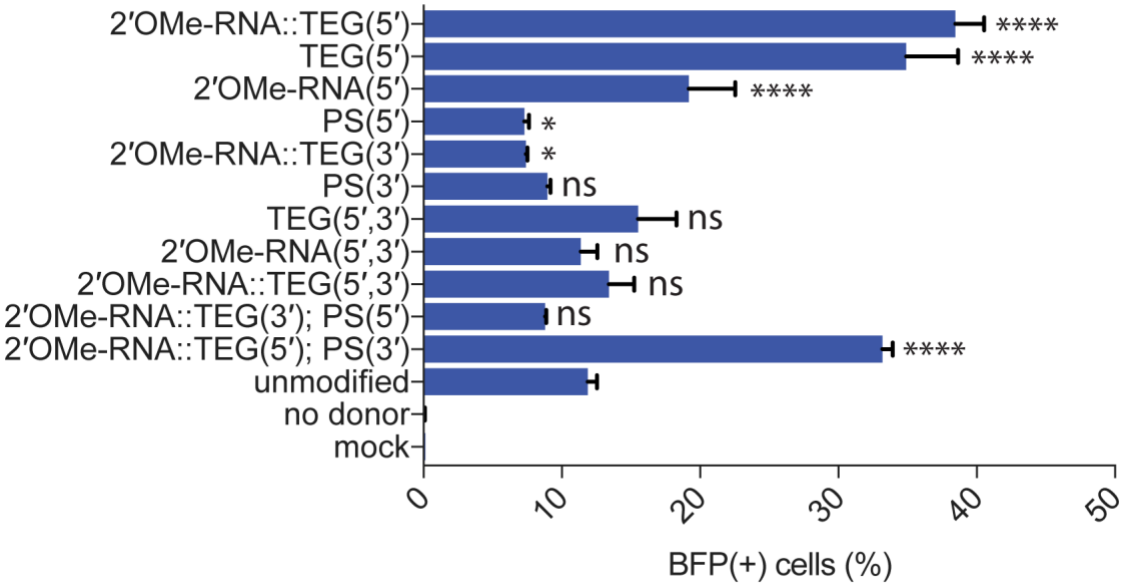


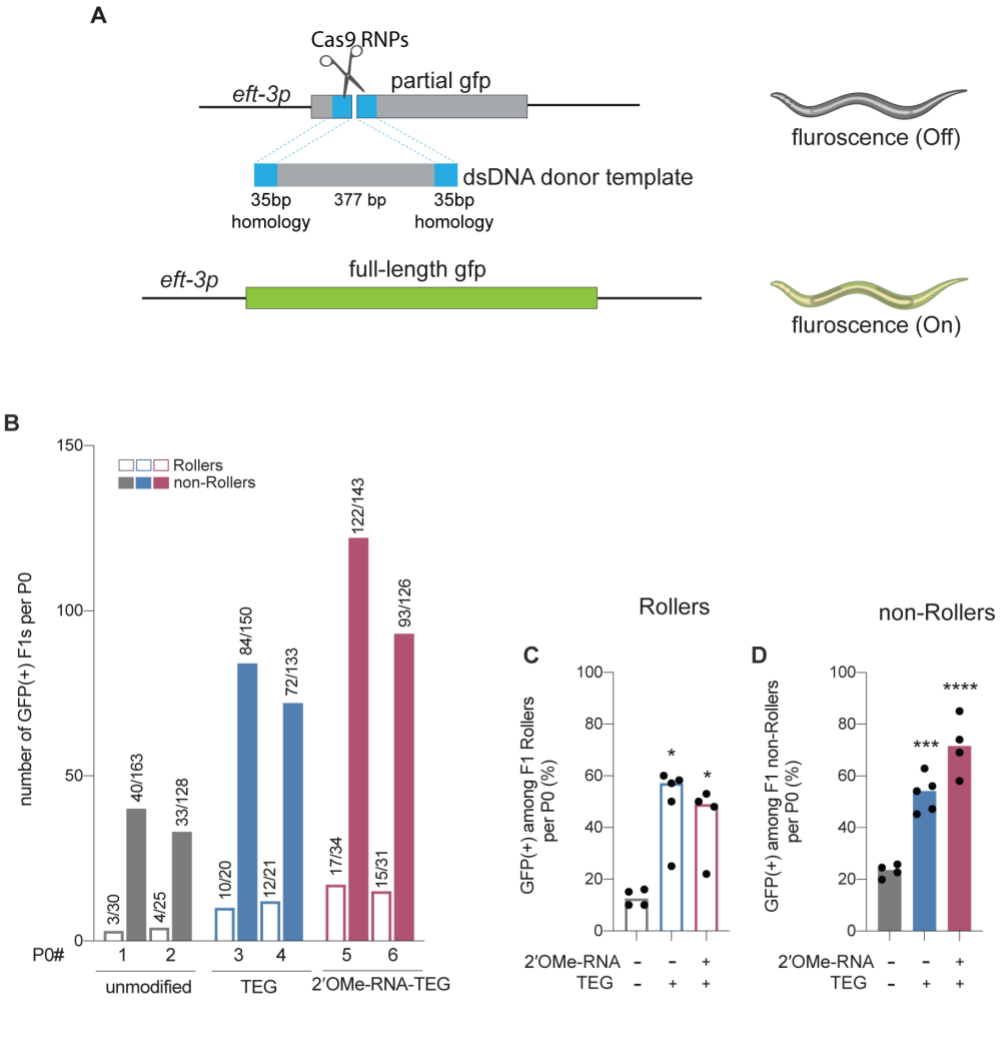
Figure 5.6. Effects of terminal and non-terminal modifications of ssODN donors on HDR efficacy. Editing efficacy of GFP-to-BFP conversion in K562 cells using 0.5 pmol of ssODN donors modified at the 5′ end alone, the 3′ end alone, or at both the 5′ and 3′ ends, with phosphorothioate (PS), TEG, 2′OMe-RNA, or 2′OMe-RNA::TEG, plotted as percentage of BFP(+) cells (HDR). Note that the PS modification is at the 5′ or 3′ internal linkages while TEG modifications are appended to the 5′ or 3′ terminus. All data points represent a mean of at least three independent replicates and all error bars represent standard deviation. P-values were calculated using one-way ANOVA and in all cases end-modified donors were compared to the unmodified donor (Tukey's multiple comparisons test; ****P < 0.0001; ***P < 0.001; **P < 0.01; *P< 0.05; ns- not significant). Mean ± s.d for at least three independent replicates are plotted. P-values were calculated using one-

way ANOVA and in all cases end-modified donors were compared to unmodified donor unless indicated otherwise (Tukey's multiple comparisons test; ****P < 0.0001; ***P < 0.001; **P < 0.01; *P< 0.05; ns- not significant).

5′-modified donors promote precision germline editing in *C. elegans*

Efficient genome editing in *C. elegans* can be achieved by directly injecting mixtures of Cas9 ribonucleoprotein (RNP) complex and donor into the syncytial ovary[208,212,217], producing dozens of independent precision editing events among the progeny of each injected animal[203]. We designed unmodified, TEG-modified, and RNA::TEG-modified donors to insert *gfp* at the *csr-1* locus or to correct *eft-3p::gfp* reporter that contains partial sequence of *gfp* (see Methods; Figure 5.7A and Figure 5.8A). To monitor injection quality, we co-injected a plasmid encoding the transformation marker *rol-6(su1006)*, which produces the Roller phenotype. The TEG- and RNA::TEG-modified donors produced about twice as many GFP(+) progeny per injected animal than did the unmodified donor (Figure 5.7B and E, two representative broods per donor). Among the Roller cohort, which was previously shown to exhibit lower editing efficiency[203], end-modified donors increased the fraction of GFP(+) Roller progeny by several fold. For example, whereas the unmodified *eft-3* donor produced only 12.6% GFP-positive Rollers, the TEG- and RNA::TEG-modified *eft-3* donors produced 57.1% and 49% GFP-positive Rollers (Figure 5.7C). Similarly, GFP::CSR-1(+) Rollers increased from 8.8% (unmodified) to 28% (TEG) and 32.8% (RNA::TEG) (Figure 5.8B and C). TEG- and RNA::TEG-modified *eft-3* and *csr-1* donors produced >50% GFP(+) non-Roller progeny compared to roughly 22% (*eft-3*) and 30% (*csr-1*) GFP(+) non-Rollers produced by the unmodified donor (Figure 5.7D and 5.8D). Every GFP(+) animal tested

transmitted the edit to the next generation (Figure 5.7E). Thus, compared to the unmodified donors, the 5′-TEG and 5′-RNA::TEG donors substantially increase the frequency of *gfp* insertion by HDR in the *C. elegans* germline. Strikingly, end-modified donors frequently yielded more than 100 independent GFP(+) F1 progeny from a single injected hermaphrodite.



E

dsDNA donor type	GFP(+) F1s cloned	F2 transmission (%)
unmodified	40	40 (100)
TEG	39	39 (100)
2'OMe-RNA-TEG	36	36 (100)

Figure 5.7. Modified donors promote precise editing in *C. elegans*. (A) Schematic showing end-modified dsDNA donors (25ng/µl) with short (~35bp) homology arms to insert part *gfp* into partial *gfp* deletion. (B) Number of GFP expressing animals among entire F1 brood of two representative P0 animals for each donor type are plotted for *eft-3p* reporter locus. Fraction of F1 animals

expressing GFP among (C) Roller and (D) non-Roller cohorts are plotted as percentage for *eft-3p* locus. (E) GFP-positive F1 animals were cloned and their progeny (F2s) were scored for GFP expression. Number of F1s that produced GFP expressing F2 in a mendelian fashion are shown under F2 transmission column. Open bars (Rollers) and closed bars represent (non-Rollers) median. Number of GFP expressing animals among total number of animals scored per cohort are shown above the bars. n $\geq$4 broods for each donor condition. P-values were calculated using one-way ANOVA and in all cases end-modified donors were compared to unmodified donors (Tukey's multiple comparisons test; ****P < 0.0001; ***P < 0.001; **P < 0.01; *P< 0.05; ns- not significant).

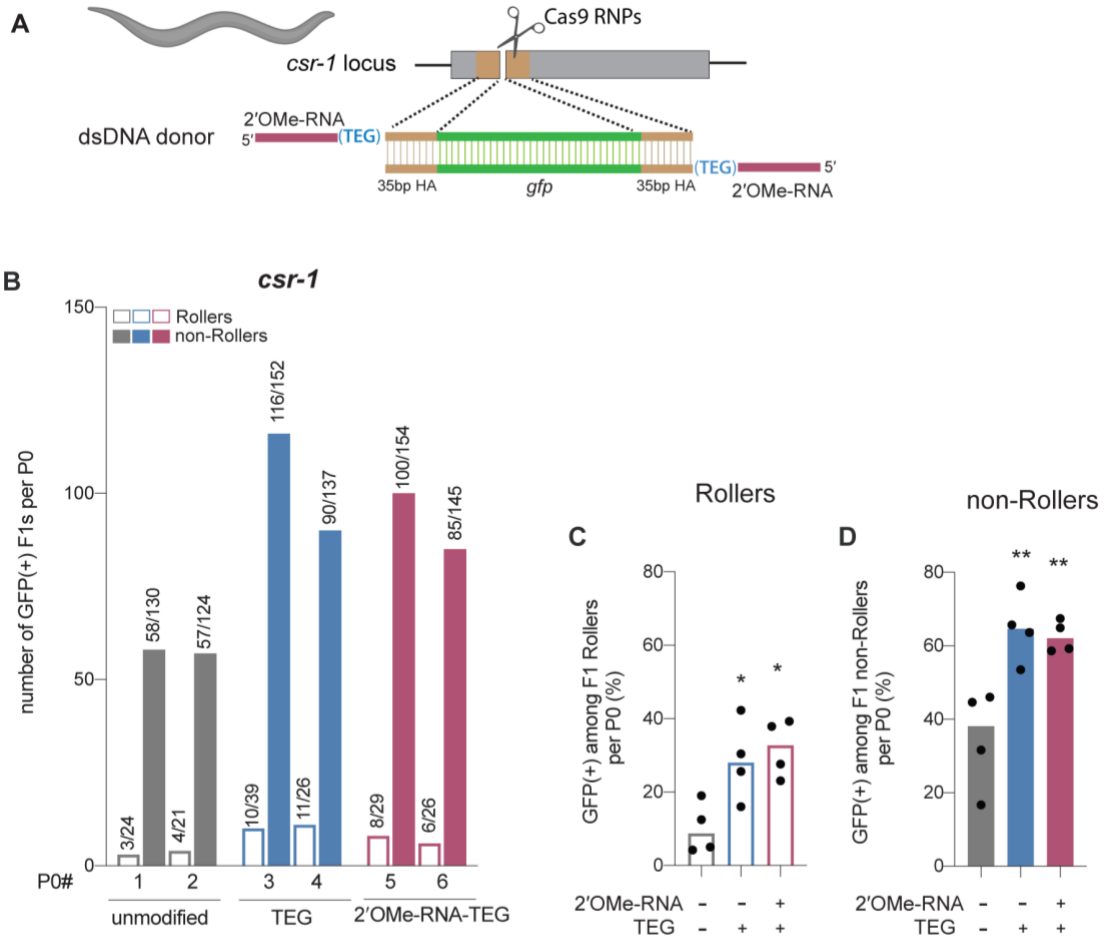


Figure 5.8. **Modified donors promote precise editing in *C. elegans*.** (A) Schematic showing end-modified dsDNA donors (25ng/µl) with short (~35bp) homology arms to insert *gfp* at the *csr-1* locus. (B) Number of GFP expressing animals among entire F1 brood of two representative P0 animals for each donor type are plotted. Fraction of F1 animals expressing GFP among (C) Roller and (D) non-Roller cohorts are plotted as percentage for *eft-3p* locus. Open bars (Rollers) and closed bars represent (non-Rollers) median. Number of GFP expressing

animals among total number of animals scored per cohort are shown above the bars. n ≥4 broods for each donor condition. P-values were calculated using one-way ANOVA and in all cases end-modified donors were compared to unmodified donors (Tukey's multiple comparisons test; ****P < 0.0001; ***P < 0.001; **P < 0.01; *P< 0.05; ns- not significant).

5′-modified donors promote precision editing in mouse zygotes

To test whether RNA::TEG-modified donors enhance precise editing in mouse zygotes, we targeted the *Tyrosinase (Tyr)* and *Sox2* loci. First, we sought to convert the coat color of Swiss-Webster albino (Tyr^{c}) mice to a pigmented phenotype ($Tyr^{c\text{-}cor}$*; cor: corrected)* using a donor to replace the serine 103 codon (TCT) with a cysteine (TGC) codon. The donor also introduces six silent mutations to prevent the guide RNA from directing cleavage of the edited locus (Figure 5.9A). We injected unmodified or RNA::TEG-modified donors with Cas9 RNPs into zygotes, transferred the embryos into pseudo-pregnant females, and quantified the repair efficiency by phenotyping the coat color of founder (F0) mice. The RNA::TEG-modified donor yielded more than twice as many pigmented F0 mice (37.9% uniform or mosaic) compared to unmodified donor (17.4%) (Figure 5.9A, Figure 5.10A). Strikingly, most (92%) of the edited founders produced by the RNA::TEG-modified donor had uniformly pigmented coats, whereas only 62.5% of the edited F0 produced by the unmodified donor had a uniformly pigmented coat color (Figure 5.9B; Figure 5.10A), suggesting that the RNA::TEG-modified donor promotes editing during early zygotic divisions. Representative images of F0 litters with dark coat color are shown in Figure 5.9C. We confirmed that F0 mice with pigmented coat transmitted the corrected $Tyr^{c\text{-}cor}$ allele to F1 pups (Figure 5.10B and C). Taken together, these results show that RNA::TEG donors are at least two-fold more efficient than unmodified donors in mouse zygote editing.

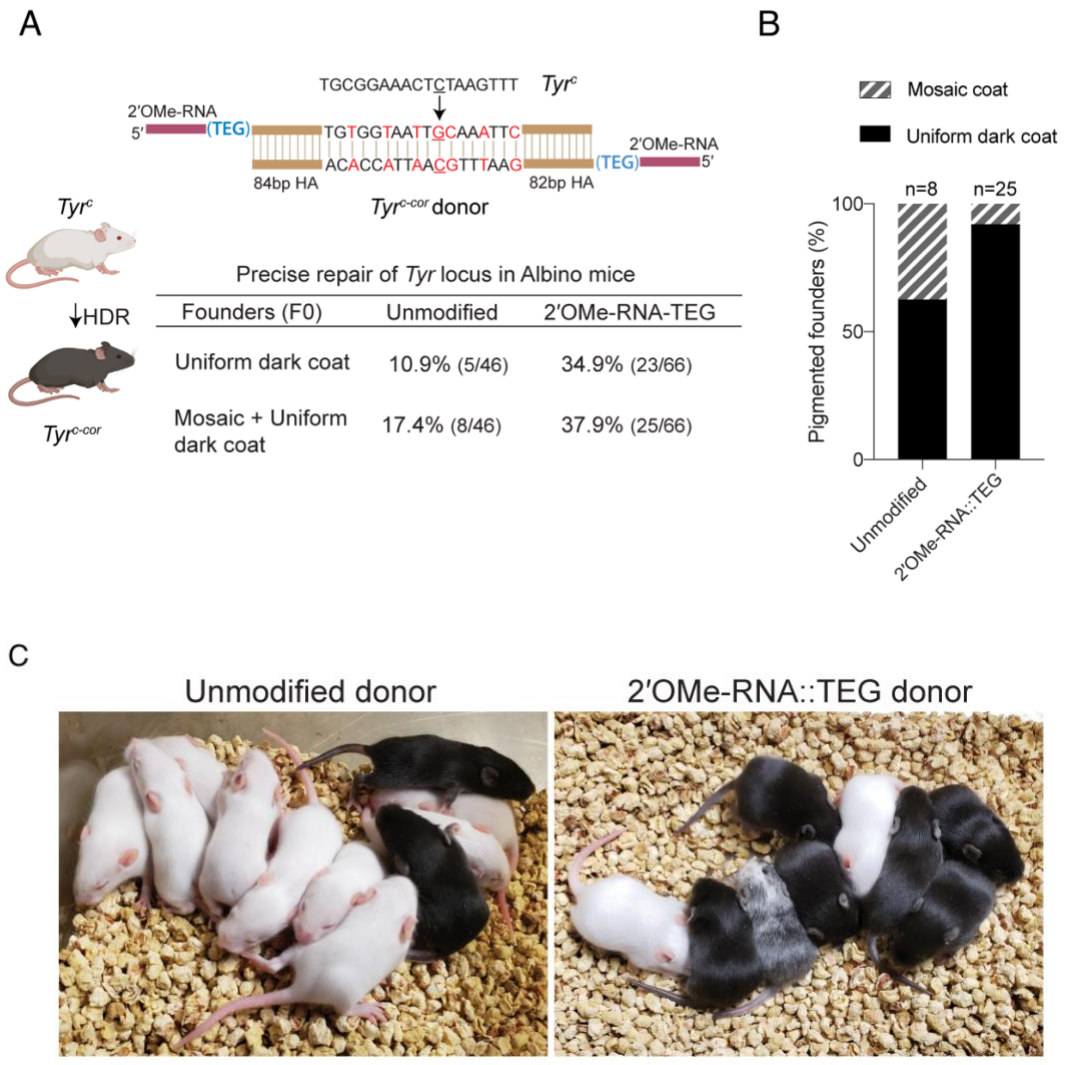


Figure 5.9. **2′-OMe-RNA-TEG donors promote precise editing in mouse zygotes**. (A). Design of the dsDNA donors to precisely convert the coat color of albino mice (*Tyr*^{C}) to pigmented (*Tyr*$^{C\text{-}Cor}$) by editing C to G (underscored) along with six silent mutations (in red) is shown. Percentages of F0 founder mice with black coat are shown. (B) Percentages of animals among HDR positive F0s that have uniform dark coat or mosaic coat color are plotted for unmodified and 5′ modified donors. (C) Representative pictures of 10 days old F0 mice with

pigmented (HDR) or white (wt or indel) coat color are shown. One mosaic mouse (third from left) can be seen among the pups obtained with end-modified donor. HA: Homology Arms.

A

Microinjection information for targetted editing of the *Tyr* locus in Swiss-webster mice

Donor type	Zygotes transferred	Pups born	Pups with dark eyes (P2)	Pups alive (P10)	Pups with coat pigment (P10)	Coat color (% among P10 animals)		
						Uniform-HDR (Black/Agouti)	Mosaic HDR	Non-HDR (Albino)
Unmodified	169	56	7 (12.5%)	46	8 (17.4%)	5 (10.9%)	3 (6.5%)	38 (82.6%)
2′OMe-RNA-TEG	198	68	24 (35.3%)	66	25 (37.9%)	23 (34.9%)	2 (3.0%)	41 (62.1%)

Only pups that were alive on P10 were inlcuded in the analysis. Among the pups that died before P10 one pup (2′OMe-RNA-TEG group) exhibited dark eye phenotype.

B

Germline transmission of *Tyr*$^{c\text{-}cor}$ allele to F1 generation

Donor Type	F0 Mouse ID# (gender)	Number of F1 pups	number of F1 pups with TYR$^{C\text{-}Cor}$ (%)
Unmodified	1 (female)	11	5 (45.5)
	16 (male)	7	2 (28.5)
	20 (female)	14	8 (57.1)
	51 (male)	14	10 (71.4)
	54 (female)	15	8 (53.3)
	55 (female)	15	8 (53.3)
	59 (male)	17	6 (35.3)
	60 (male)	9	8 (88.9)
2′OMe-RNA-TEG	1 (male)	11	7 (63.6)
	6 (male)	12	6 (50)
	18 (female)	14	5 (35.7)
	22 (female)	14	6 (42.3)
	23 (female)	14	5 (35.7)
	24 (female)	13	9 (69.2)
	41 (male)	8	4 (50)
	42 (male)	9	4 (44.4)

C

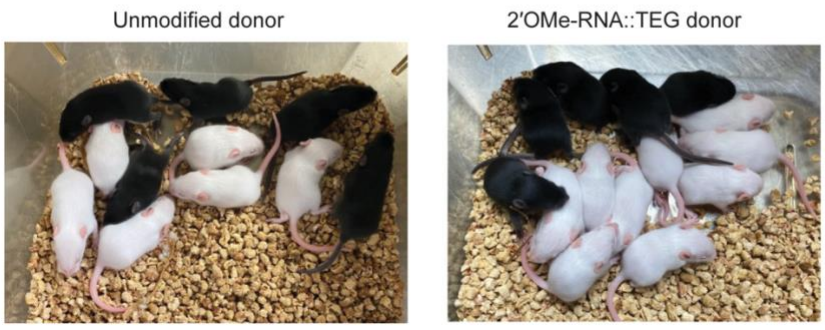
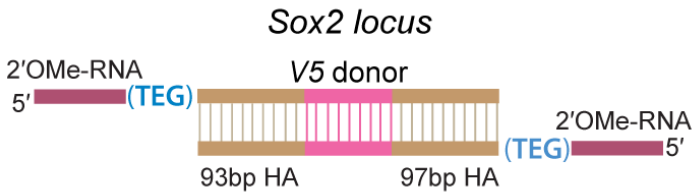


Figure 5.10. **Microinjection information for editing at the Tyr locus in Albino mice** (A) Microinjection information and HDR efficiencies obtained for unmodified and 2′OMe-RNA::TEG modified donors. (B) Germline transmission of the edited

$Tyr^{c\text{-}cor}$ allele was confirmed by crossing some of the pigmented F0 mice to Swiss webster mice (Tyr^{c}) and phenotyping their F1 progeny (C) Representative images of F1 litters obtained from crosses for germline transmission tests.

Next, we sought to insert a sequence encoding an in-frame V5 epitope immediately before the stop codon at the 3′ end of the *Sox2* locus (Figure 5.11A). We injected unmodified or RNA::TEG-modified donors with Cas9 RNPs into zygotes, transferred the embryos into pseudo-pregnant mice, and genotyped F0 progeny by PCR across the *Sox2* target site and Sanger sequencing. The V5 tag was precisely inserted into the *Sox2* locus in only 5.7% (n=35) of F0 animals from the injection with unmodified donor. By contrast, the RNA::TEG-modified donor resulted in precise insertion of V5 in 33.3% (n=24) of the F0 animals—a greater than 5-fold increase in precise editing (Figure 5.11B and Figure 5.12A). All of the V5-positive founders tested (one F0 from the unmodified donor and six F0s from RNA::TEG-modified donor) transmitted the *Sox2::V5* allele to F1 progeny and the insertion was confirmed by Sanger sequencing (Figure 5.12B and C). Thus the 5′-RNA::TEG modification greatly improves the efficiency of precise genome editing in vertebrate model systems.

A



B

Precise insertion of C-terminal V5 tag at *Sox-2* locus

Founders (F0)	Unmodified	2′OMe-RNA-TEG
V5 insertion	5.7% (2/35)	33.3% (8/24)

Figure 5.11. **5′ modified donors improve knock-in efficiencies at the Sox2 locus**. Donor design to knock-in V5 tag at the C-terminus of Sox2 is shown. (B) Percentage of founder animals containing perfect *V5* insertion at *Sox2* locus are shown for each donor type. HA: Homology Arms.

A

Microinjection information for *V5* tag insertion at the *Sox2* locus

Donor type	Zygotes transferred	Pups born	Pups analyzed	HDR (% among analyzed)
Unmodified	214	36	35	2 (5.7%)
2′OMe-RNA-TEG	209	27	24	8 (33.3%)

1 pup from the unmodified donor group and 3 pups from 2′OMe-RNA-TEG donor group died at P3 and were not included in the analysis).

B

Germline transmission of *Sox2::V5* to F1 generation

Donor Type	F0 Mouse ID# (gender)	number of F1 pups	number of F1 pups with *Sox2::V5* (%)
Unmodified	6 (male)	18	14 (77.8)
2′OMe-RNA-TEG	10 (male)	10	5 (50)
	11 (male)	10	3 (30)
	12 (female)	6	4 (66.7)
	16 (female)	1	1 (100)
	17 (female)	5	3 (60)
	26 (male)	6	4 (66.7)

C

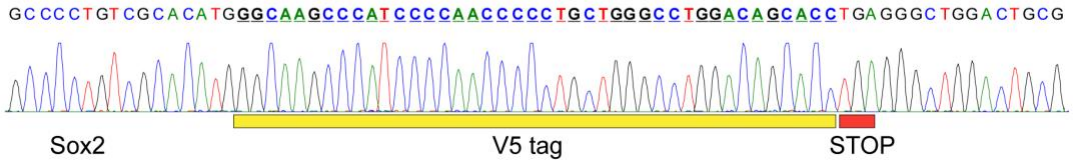


Figure 5.12. Microinjection information for editing at the Sox2 locus in mouse zygotes (A) Microinjection information and HDR efficiencies obtained using unmodified and 2′OMe-RNA::TEG modified donors are shown. (B) Germline transmission rates of the Sox2::V5 allele was confirmed by crossing the HDR positive F0 mice with WT mice and genotyping the F1 pups. (C). Sanger sequencing trace of Sox2::V5 allele in F1 mice.

5′-modifications suppress donor concatenation

Upon delivery into animal cells or embryos, linear DNA molecules are known to form extensive homology-mediated and ligation-dependent concatemers (Figure 5.13A)[214,230,231]. We reasoned that 5′ modifications to the donor might suppress the formation of concatemers, thereby making linear donors more available for HDR. To test this idea, we nucleofected 566 bp dsDNA donors into HEK293T cells, harvested cells over a course of 3 days, and assessed the formation of concatemers by Southern blot analysis. We found that the unmodified dsDNA formed concatemers within 1 hour after nucleofection. These concatemers were composed of two to several copies of the DNA, inferred from the presence of a ladder of bands on the Southern blot (Figure 5.13B). Concatemers of up to ten copies were present within 3 hours after nucleofection and peaked in abundance by 12 hours. Concatemer levels declined over the next 12 hours but persisted at low levels until at least 72 hours after nucleofection. By contrast, the TEG-modified DNA showed a marked delay in the formation and levels of multimers (Figure 5.13B). Dimers and trimers gradually formed over the first 12 to 24 hours but were present at much lower levels than those formed by unmodified DNA. At late time points—24, 48, and 72 hours after transfection—we observed a greater fraction of TEG-modified DNA monomers than unmodified monomers (Figure 5.13B). These results suggest that the 5′-TEG modification significantly suppresses concatemer formation.

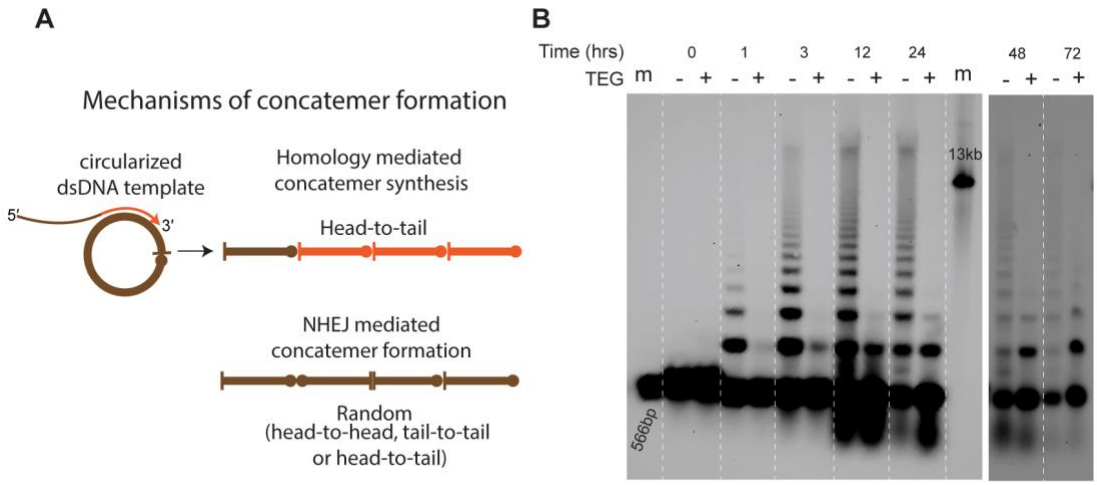


Figure 5.13. End-modifications suppress formation of donor concatemers. (A) Model for mechanisms of concatemer formation is shown. Concatemers form through homology mediated synthesis of new molecules that contain all the monomers oriented in the same fashion or by end-ligation reactions between the donor molecules in a random fashion. Monomers in NHEJ mediated concatemer can be ligated head-to-tail, tail-to-tail or head-to-head (B) Southern blot of unmodified and TEG modified dsDNA (566bp) nucleofected into HEK293T cells and collected at indicated time points. 1pmol dsDNA was nucleofected into HEK293 cells and cells were harvested at indicated time points. Concatemerization of unmodified DNA is visualized as ladders; 566bp DNA and 13kb long DNA are used as size markers (m)

End-modifications suppress direct ligation of short DNA into DSBs

To determine if TEG modification suppresses the direct ligation of TEG-modified linear molecules into chromosomal DSBs, we performed GUIDE-seq analyses[236], which measures the incorporation of short (34nt) dsDNA into on-target and off-target DSBs (Figure 5.14A). We targeted the *ARHGEF9* locus, previously characterized for off- target editing[110]. Strikingly, the TEG-modified DNA produced 19-fold fewer GUIDE-seq reads (genome wide) than did the unmodified DNA (Figure 5.14B). The number of TEG-modified DNA insertions obtained at the on-target cut site in the *ARHGEF9* locus and at the top 6 off target sites were dramatically reduced, ranging from 15-fold to 6-fold lower compared to insertions of the unmodified DNA (Figure 5.14C). Taken together these data suggest that TEG-modifications suppress direct ligation of donor molecules both to each other and to chromosomal DSBs.

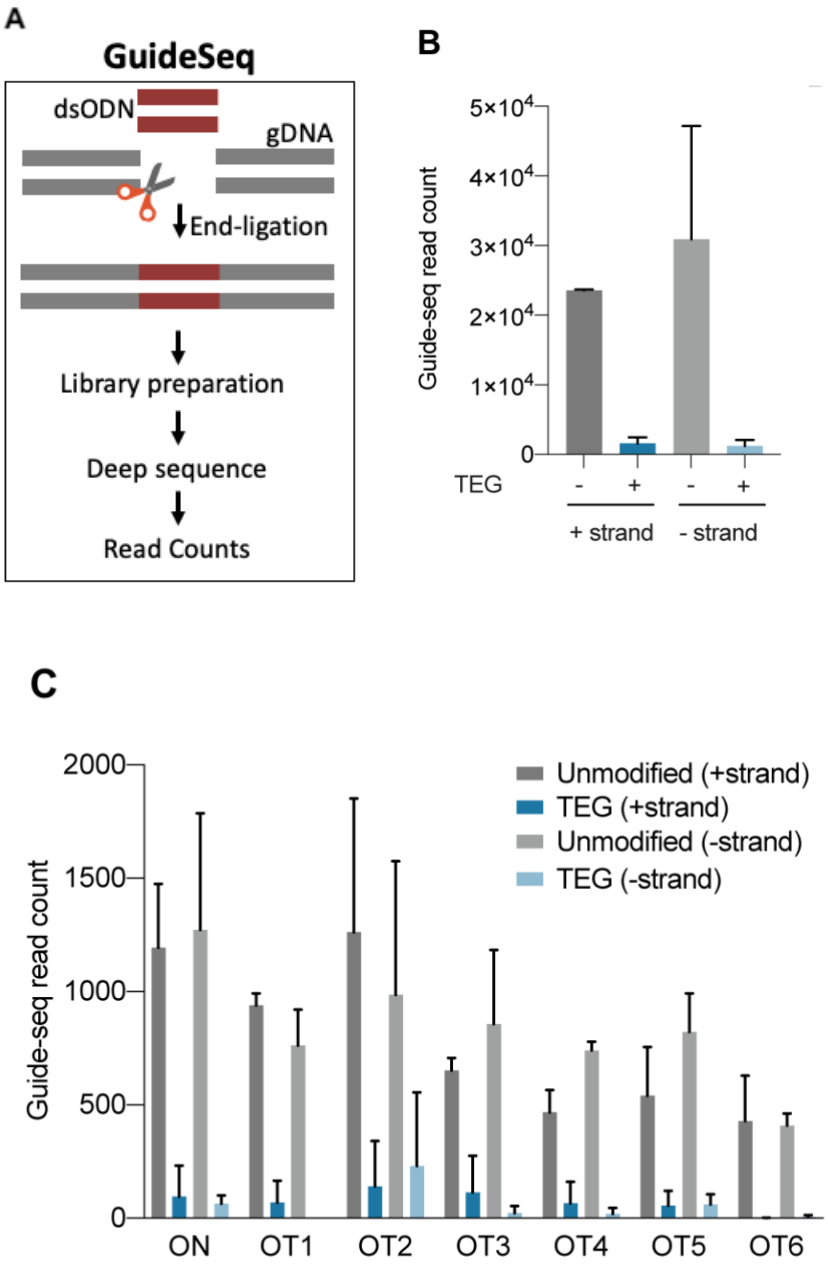


Figure 5.14. **End-modified donors suppress donor integration at off-target loci** (A) Schematic showing the Guideseq protocol. dsODN (34bp) was nucleofected into HEK293T cells with or without TEG modifications on the 5′

ends of the DNA. Number of Guide-seq reads with unmodified and TEG modified dsDNA integration for, (B) whole genome and (C) on-target (*ARHGEF9)* and six previously validated off-target loci are plotted. Data from two biological replicates is shown.

Discussion

Here we have explored how several types of chemical modifications to the repair template DNA affect the efficiency of precise homology-dependent repair. In mammalian cells, donors containing simple modifications such as TEG or 2′OMe-RNA::TEG on their 5′ ends improved HDR efficacy. These modifications increased the potency of single- and double-stranded DNA (long and short) donors, allowing efficient editing at significantly lower amounts. Modifying the ends of the donors suppressed concatemer formation and significantly reduced random integration of short dsDNA at chromosomal DSBs.

End modifications affected long and short donors differently in mammalian cells. On long donors end modification caused a ~2-to-5-fold increase in HDR frequency (total efficacy) compared to unmodified donors and did so without changing the donor concentration where efficacy reached its plateau. In contrast, on short donors end modifications did not increase the maximal efficacy of HDR, but instead dramatically reduced the amount of donor required to reach that maximal level. Put another way, long DNA donors exhibited both increased potency and maximal efficacy when modified, while short ssODN and dsDNA donors exhibited increased potency but no increase in maximal efficacy. This difference requires further study but could be explained if shorter donors and longer DNA donors experience different dose-limiting barriers. For example, the dose-limiting toxicity of ssODNs could be driven by total number of free DNA ends per cell, while longer molecules could encounter dose-limiting toxicity driven by

total DNA mass. Consistent with this idea, unmodified long dsDNA donors begin to plateau in efficacy at nearly 4-fold more mass, but ~10-fold lower molar amounts than ssODNs. When end-modified, both types of donor exhibit similar maximal efficacy in the 1 to 2 pmol range.

RNA::TEG-modified donors significantly increased the levels of precision editing in model organisms (*C. elegans* and mice). In all animal models, high HDR efficiencies were achieved using end-modified dsDNA donors, that in some cases approached efficiencies previously observed for ssODN donors[208,239]. Importantly, precise insertions were obtained with relatively short homology arms. For example, in mouse zygote injections, we used donors with homology arms of less than 90 bp, similar to typical arm lengths used for ssODN donors[162] and at relatively low concentrations (1 ng/µl).

How do end-modifications help increase the efficacy of the donors? Our findings suggest that they do so, in part, by suppressing non-homologous end-joining reactions. In several systems dsDNA donors have been shown to quickly form extrachromosomal arrays[214,230,231] and may do so directly in the cytoplasm[222]. For example, DNA delivered into the cytoplasm of the *C. elegans* gonadal syncytium gains entry into oocytes over a 24 hour period in a manner more consistent with cytoplasmic flow than with direct nuclear uptake by germ nuclei[203], and transformants established in this way have been shown to contain concatenated arrays of injected DNA, several hundred kilobases in length, which then partition to progeny in a non-Mendelian fashion as extrachromosomal

elements[214,215]. Integration of similar extrachromosomal arrays into the host genome have been reported in zebrafish and mouse zygotes[5,232,233]. Thus, the suppression of donor concatemer formation by 5′ modified donors could increase the effective molar amounts of donor available for precise repair of the target double strand break. Similarly, once in the nucleus, the suppression of direct ligation to chromosomal DNA through end-joining reactions could further increase precision repair. Perhaps consistent with suppression of concatenation as a major mechanism of action, it is intriguing that modification of a single end was nearly as effective as modifications to both ends of the donor. In principle, a single end modification would limit concatenation to dimer formation. Similarly, modification of a single end could prevent donors from ligating into circles which might then concatenate further through HDR.

In addition to increasing the amount of available donor molecules, another possible benefit of suppressing end-joining reactions is that the free ends of the donor might then be available to participate in the HDR mechanism (for example, by assembling elements of the DSB repair machinery directly on the free 3′-end of the donor). We found that a free unmodified 3′ end was required for efficient HDR. Thus, by suppressing ligation, the 5′ modification in effect maintains available 3′ ends, perhaps to prime repair synthesis.

In previous studies, fluorescent and amine modifications to the 5′ and 3′ termini of ssODN donors did not improve HDR efficacy over unmodified donors[100]. However, these studies were performed using doses 50-fold higher than the

optimal dose for modified donors determined here. Similarly, phosphorothioate (PS) linkages were shown to improve HDR at doses much higher than the optimal dose for modified ssODNs in our study[157]. In our study, ssODNs with PS linkages did not improve HDR at doses where RNA::TEG- and TEG-modified donors were most efficacious. While our study was in preparation[207], three studies explored donors with 5′-end modifications. One study showed that the addition of biotin improved HDR and favored single copy insertion in the rice fish medaka[240]. The biotin moiety was attached to the donor via a polyethylene glycol (PEG) linker, but the study did not explore donors with PEG alone. Yu et al. (2020) showed that PEG10 with a 6-carbon linker boosted precise GFP insertions in vertebrate cells similar to those reported here for TEG- and RNA::TEG-modified donors, and at similar concentrations to those we employed[241]. The third study describes the suppression of NHEJ-mediated insertions using donors with 5′-Biotin::PEG or 5′-ssDNA::PEG moieties[242]. Our studies are in agreement with these findings and extend them to additional modifications and to *in vivo* genome-editing applications in animal systems.

We do not understand why donors modified with TEG and RNA::TEG performed similarly in *C. elegans*, while RNA::TEG was consistently superior to TEG alone in human cells. The *C. elegans* system is unique in that it targets meiotic pachytene nuclei that are actively engaged in HDR. Perhaps donors must persist longer to engage the DSB repair machinery in mitotic cells. The RNA::TEG modification might therefore facilitate editing in mitotic cells by providing better

protection from nuclease activity compared to TEG alone. PS linkages are known to protect against nuclease activity[157], and it will therefore be interesting to explore whether a combination of internal (e.g., PS linkages) and terminal (e.g., 5′-RNA::TEG or 5′-TEG) modifications can further increase HDR efficacy. Indeed, our results should incite the search for additional chemistries that could boost donor stability while still allowing the donor to serve as a template for repair polymerases; some such studies are underway in our laboratories. Future studies will also need to explore whether the incorporation of donor chemistries will synergize with other methods that stimulate HDR[132,134,135,139,243-245].

Acknowledgements. This work was funded by Howard Hughes Medical Institute (C.C.M.), NIH R37GM058800-23 (C.C.M), institutional funds from University of Massachusetts Medical School (E.J.S and J.K.W), NIH UG3 TR002668 (E.J.S and J.K.W), NIH/NHLBI R35 HL140017 (N.D.L) and NIH/OD R21 OD030004 (N.D.L). We thank Darryl Conte, Jr. for critical reading of the manuscript, Scot Wolfe for sharing the Cas9 expression plasmid and Julia Rembetsy-Brown for maintaining *Sox2::V5* mouse colony. Biorender was used to generate some of the graphics in figures. We thank the UMass Medical School Deep Sequencing Core for their assistance with the Illumina and PacBio sequencing.

Competing interests: The authors (K.S.G, A.M, G.A.D, H.G, J.K.W, E.J.S and C.C.M) have a patent application pending related to the findings described. Craig C. Mello is a co-founder and Scientific Advisory Board member of CRISPR Therapeutics, and Erik J. Sontheimer is a co-founder and Scientific Advisory Board member of Intellia Therapeutics.

Chapter VI: Discussion

We started these studies to improve genome editing efficiencies of CRISPR/Cas9 technology. First, using *C. elegans* as the model system we have shown that our genome editing protocols are highly efficient. Using the methods developed here, investigators including beginners should be able to easily implement the genome editing protocols. We found that high doses of Cas9 produced unexpected phenotypes such as embryonic lethality and sterility that are not associated with the target locus. These observations motivated us to include metrics that could track toxicity. To this end, we used a visible dominant co- injection marker, *rol-6 (su1006)*, to report on the viability of the F1 progeny. Previous studies on DNA transformation show that properly injected hermaphrodite produces about 30-50 Roller progeny[214]. In our experiments, at high doses of Cas9 the number of F1 Rollers decreased significantly compared to Roller alone. Therefore, our data clearly shows that high doses of Cas9 RNPs are toxic to the animals and lead to significant reduction in brood sizes. Presumably, the progeny that receive the largest amounts of injection mixture are eliminated due to toxicity thereby leading to fewer Rollers. With doses used in our optimized protocol, more than 90% of F1 progeny were edited at the target locus(indels). Leveraging these high editing efficiencies, we show that high frequencies of precise inserts can be achieved with ssODN donors. We designed asymmetric donor strategy to improve HDR efficiencies with long (1kb) dsDNA donors (Chapter 3).

Our efforts to improve knock-in efficiencies of long donors (~1kb) have significantly lowered the barrier to precisely editing the germline of *C. elegans* (Chapter 4). We have shown that melting and quickly cooling the double stranded DNA donors dramatically increases HDR efficiencies. We have also shown that nearly 50% of the post-injection cohort express GFP (knock-in). Several CRISPR protocols recommend high doses (300-700 ng/µl) of donor for efficient HDR[210,211,213]. However, while carefully optimizing the donor doses we have found that high DNA concentrations are toxic, reduce brood sizes and compromise HDR efficiencies. Our studies show that as low as 10-25 ng/µl of melted donor is sufficient and ideal to achieve high frequency of GFP insertions. To our knowledge these are the highest HDR efficiencies achieved with 1kb long dsDNA donors in *C. elegans.* Our protocols allow even novice researchers to adopt this powerful genetic system.

While exploring the properties of chemically modified donors (see below), we serendipitously found that heating and quickly cooling (melting) the dsDNA donors is critical to increase HDR efficiencies in *C. elegans* germline. We do not yet know why melting increases HDR efficiencies. We speculate that the changes in DNA topology due to melting promote recombination. Nuclear DNA exists in a highly packaged stage and in complex with histone proteins. DNA and histone modifications such as methylation and acetylation add another layer of complexity to the chromatin structure. Protein complexes bound to chromatin at the homologous DNA may facilitate homology search, template unwinding and

repair. In contrast, exogenous DNA donors generated by PCRs do not contain any modifications or proteins to facilitate homologous recombination. Indeed, studies in mammalian cells have shown that preassembling donor templates with histones H2A and H2B stimulates HDR[246]. It is possible that melting the donor DNA circumvents the need for some proteins on the repair template. Localized opening of double stranded DNA is required for biological processes such as DNA replication and transcription[247,248]. Denaturing bubbles in dsDNA can also function as binding sites for regulatory proteins[248,249]. For example, formation of denaturing bubbles could recruit repair proteins or promote easier strand invasion without the need for a helicase to unwind the duplex. To this end, we found that slowly cooling the donors to promote better annealing has resulted in lower efficiencies compared to quickly cooling. These results suggest that imperfect annealing may play a role in increasing recombination efficiencies. Furthermore, local denaturation induced by melting and protein binding could also lead to compaction of donors[250] that facilitates passive diffusion of donor DNA across the nuclear pore complexes (NPC) while untreated linear may not gain nuclear access. Future biochemical and single molecule imaging studies would help understand the exact mechanism of melted donors. Experimental validation of these models and better understanding of the mechanism of increased HDR with melted donors will be helpful to improve editing protocols in other model systems.

Precision genome editing using repair templates has been inefficient in mammalian model systems. In chapter 5, using mammalian cells, we address HDR inefficiencies by chemically modifying the donor templates. We initiated these studies to explore ways to promote nuclear entry of long donors. Studies have shown that animals cells have size limits for nuclease exclusion and[129,251].

For efficient precision editing, Cas9 RNPs and donor DNA should gain access to the nuclear DNA and donor should be available at the DSB for repair (spatially and temporally). Recombinant Cas9 protein contains an NLS peptide that enables active nuclear import of Cas9 RNPs. However, longer donors likely cannot enter the nucleus because of size limits imposed by nuclear pore complexes thereby reducing the effective concentrations at the site of repair. We hypothesized that the disparity in availability of ssODN and dsDNA donor molecules inside germ nuclei could account for the differences in observed HDR efficiencies. Previous studies have shown that nuclear pore complexes in animal cells do not permit passive diffusion of DNA more than 300 bp[129]. In contrast, short donors can freely diffuse into the nucleus[252]. Using mammalian cell cultures it has also been demonstrated that the addition of an NLS enhances nuclear uptake of plasmid DNA following transfection[251].

Therefore, to increase potency of long dsDNA donors we attached an SV40 peptide containing the core nuclear localization signal (NLS) to the donor molecule, reasoning that the modification might promote nuclear uptake and retention. To attach an NLS peptide to a long donor DNA, we first conjugated 15-

nucleotide 2′-O-methyl (2′OMe) RNA adapters via a triethylene glycol (TEG) linkage to the 5′ ends of two target-locus specific synthetic ~20-nucleotide DNA oligonucleotides. The DNA sequences in these molecules serve as PCR primers to amplify the donor from plasmid containing the homology arms and insert (e.g., GFP sequence) for in-frame insertion into the target gene of interest. In addition, we synthesized an NLS peptide linked to a peptide nucleic acid (PNA) complementary to the 15nt 2′OMe-RNA adapter. Attachment of the NLS to the donor was then achieved by simply annealing the PNA::NLS molecules to the 2′OMe-RNA adapters on the ends of the PCR product.

Surprisingly, we found that the modified donors improved HDR even without PNA-NLS attachments suggesting that the blocking 5′ ends of the donors is critical for improving efficiencies. While annealing PNA-NLS to the donors by heating and cooling (melting) the mixtures we serendipitously found that this melting step is required to potentiate the dsDNA donors for editing in *C. elegans* germline (discussed above). Our data shows that TEG modifications and 2′OMe-RNA:TEG modifications at the 5′ termini of ssODNs or dsDNA donors increases HDR efficiencies. Chemical modifications not only increased HDR but also increased the potency of the donors (single stranded or double stranded). We show that the modified donors are efficient in *C. elegans* germline and mouse zygote injections. In *C. elegans*, combination of melted donors and chemical modifications remarkably produced more than 100 GFP expressing progeny from several individual injected hermaphrodites. Similarly, HDR increase in mouse

zygote injections show that chemically modified donors are potent and efficient. Mouse genome editing results also suggest that end-modified donors may reduce moscaism in the founder animals. It is possible that the end-modified donors function as repair templates during early cell divisions of embryogenesis or distribute across a greater number of cells in the zygotes. Importantly, we achieved HDR improvements in repair efficiencies with significantly lower concentrations of donor templates than previously used[162]. Also, our data clearly shows that optimally designed end-modified dsDNA donors are as potent as commonly used ssODNs and do not cause toxicity in zygotes[245].

Increasing the on-target efficacies without increasing off-target integration is important, especially for therapeutic purposes. Indeed, our data suggest that end-modifications reduce concatemer formation and random integration of donors into off-target loci. It is surprising to find that donors concatemerized into multicopy arrays in less than an hour in human cells. Exogenous dsDNA has been shown to readily form long concatemers in mammalian cells and many model organisms[214,230-233]. It is likely that the factors involved in ligation of free ends at DSBs are also responsible for the formation of the extra-chromosomal concatemers. Folger et al have shown that as few as 25 molecules of exogenous DNA per cell can form concatemers and both HDR and end-joining pathways contribute to their formation depending on the amount of the DNA[231]. However, this issue of concatemer formation has been largely ignored in the genome engineering field. These concatemers may reduce the effective concentration of

the donor at the double strand break (DSB) and may increase cytotoxicity. Furthermore, shorter concatemers that could traverse into the nucleus could potentially act as templates for repair which may lead to multicopy insertion of the knock-in sequence or ligate directly into the DSBs. Indeed, such tandem repeats and direct integration of the donor DNA were observed in the context of genome engineering[240,253]. We noted that although end-modifications suppressed concatemer formation, they do form at much later time points compared to unmodified donors. It is possible that the modifications get cleaved off by nucleases exposing the reactive ends of the nucleotides to end-joining ligases. Thus, suppression of concatemers and increase in amounts of monomers could contribute to increased HDR.

Although we initiated these studies with a hypothesis that NLS conjugated donors could increase HDR efficiencies, we found that VS40-NLS was not required for improved HDR efficiencies and RNA-TEG modifications increase efficacy of the donors. We have not explored if RNA-TEG modifications alone can promote nuclear entry. Future studies using DNA Fluorescence In-Situ Hybridization (FISH) studies could help quantity the number of donor molecules in the nucleus and in the cytoplasm. Different classes of NLS peptides are recognized by different importin pathways[254]. Using NLS peptides that can be targeted by other importin pathways (in addition to the canonical pathway) could increase nuclear import of donor DNA. For example, it has been shown that fusing both SV40-NLS and Nucleoplasmin long NLS to Cas12a has significantly

improved nuclear import of the nuclease and efficiencies of genome editing[180]. More experiments are needed to assess the potencies donors conjugated to various NLS peptides and how they influence the efficiencies of homology directed repair.

While our work was in progress other groups also have explored 5′ modified donors and reported similar observations[240-242]. Increase in HDR efficacy by end-modified donors in multiple model organisms and cell types demonstrates reproducibility and generality of our methodology. Importantly, this method does not involve chemical or genetics inhibition of endogenous competing end-joining pathways which could cause unforeseen toxicity.

In summary, the methods developed in these studies dramatically increase efficiencies of precision genome editing. By simply modulating the donor DNA significant HDR efficiencies can be achieved. We have shown that our technology increases HDR efficiencies in multiple model organisms (*c. elegans*, mice and human cells).

Future directions for genome editing in *C. elegans*

Genetic Screens

Chemical mutagenesis has been traditionally used to preform forward genetic screens and to identify special alleles such as temperature sensitive (TS) alleles. With the high HDR efficiencies achieved here it is now feasible to perform targeted mutagenesis. Using a library of donor templates generated by error

prone PCRs, endogenous wildtype alleles can be replaced with mutant versions. A variety of mutations across the gene can be easily isolated from a few injected P0 animals. Similarly, multiplexed reverse genetic screens can be performed using pools of guide RNAs. A recent study used this approach to identify the targets of mir-35-42 family of microRNAs[255].

Epigenome editing

Chemical modifications such as methylation and acetylation to histone residues in chromatin regulate gene expression. Targeted recruitment of chromatin modifiers provides the ability to control gene expression and enables the study of gene functions. Using fusion proteins of dCas9 and histone modifiers (Krüppel-associated box (KRAB) or VP64), gene expression can be either suppressed or induced[256-259]. Targeted epigenetic modifiers in *C. elegans* will allow the researchers to turn the gene expression on or off and the study of effects of chromatin marks at the target locus.

Challenges with long inserts

We have achieved consistent and efficient knock-in of up to 2.5kb long donors (data not shown). Yet, knocking-in inserts longer than 3kb has been challenging. For efficient templated repair availability of monomeric donor molecules is necessary. As discussed above, long donors readily concatenate into large extrachromosomal arrays that may reduce efficiencies. Long donors may not gain access to the nuclear DNA while Cas9 RNPs can readily induce double stranded breaks. Lack of donor availability, spatially and temporally, results in

poor editing efficiencies. This problem could be circumvented if the donor is available during nuclear envelope breakdown which allows its diffusion into the nuclei. Presence of homozygous alleles of knock-in (e.g., gfp) in F1 progeny of injected animals suggests that Cas9 and donor are active in zygotes. Future developments to activate Cas9 specifically in late oocytes or early embryos may increase repair efficiencies with long donors. Furthermore, the minimum number of moles of longer donors needed for efficient HDR increases the mass of the DNA and may cause toxicity. We have developed an *unc-119* based selection strategy that addresses the HDR inefficiencies associated with long donors (will be described elsewhere). Although, this method doesn't increase absolute efficiencies, positive selection easily identifies the progeny with the *knock-in.*

Future directions for genome editing in vertebrate systems

End-modified donors have significantly improved HDR efficiencies with all types of donors and lengths. However, several challenges still exist. For example, our results show that more than 50% of HEK293K/ TLR cells can be precisely edited with end-modified donors. In contrast, although end-modified donors improved HDR in HFFs by several fold, maximal HDR efficiencies remained low (<10%) (Chapter 5). Similarly, non-transformed primary cells are resistant to editing compared to transformed cells. Precision editing without toxicity or random integration of the donors into the genome is of utmost importance to bring CRISPR technology to help patients.

End-modified donors have increased HDR efficiencies in mouse zygote injections. Achieving precise insertion of longer (greater than 1kb) edits is still a challenge. One study has shown improvements in HDR by using biotinylated donors and injecting into 2-cell embryos[173]. *In vivo* correction of mutant alleles has great therapeutic potential. In future studies, we aim to test the efficiencies of chemically modified long donors in mouse zygote editing and *in vivo* disease models.

High doses of exogenous DNA, particularly dsDNA, results in cell death (chapter 4 and 5). However, the mechanism behind toxicity is not clear. Two pathways that could cause DNA dependent toxicity have been largely ignored in the genome editing field. Cellular DNA is compartmentalized either in the nucleus or in mitochondria. Presence of DNA in the cytoplasm can be detected as a threat to the cell. DNA sensors such as AIM2, TLR9 and cGAS/STING recognize cytosolic DNA and trigger the release of cytokines that further activate downstream adaptive immune responses[260,261]. Cytokine release triggered by AIM2 pathway eventually leads to cell death[262]. Whether the toxicity and cell death in CRISPR applications is caused by the cytoplasmic DNA sensors remains to be explored. Since DNA recognition by the cytosolic sensors is thought to be sequence independent[261], it is likely that donor DNA is recognized by the cytoplasmic sensors. Chemical modifications to donor DNA that could evade these sensors may reduce genotoxicity. Transiently inhibiting the DNA sensors may also increase tolerability to exogenous DNA and improve HDR

efficiencies. Such a strategy could be easily implemented in *ex vivo* editing applications. For example, to insert CAR antigen into T-cells.

Similarly, the ends of freely floating donor DNA in the nucleus could be perceived as chromosomal DSBs. Therefore, it Is possible that the free ends of the donor molecules trigger DNA damage response pathways. Even with low donor doses, each cell receives thousands of molecules that could amplify the damage signal by several orders of magnitude. Whether donor copies recruit DNA repair factors onto their ends has not been explored. Perhaps modifications such as 2′OMe-RNA moieties at the 5′ ends of the donors could suppress end-recognition by the repair factors. Studies using single molecule DNA FISH and immunofluorescence could address some of these questions.

Conclusions

In conclusion our methodologies have dramatically improved HDR efficiencies in *C. elegans* germline genome editing and in mammalian systems. The chemical modifications we have used here laid foundation for future advancements in the field of modified donor research. We are in search of other modifications that can further improve donor stability, evade the immune system, gain nuclear entry in non-diving cells and yet do not cause toxicity to the cells. Using chemical screens, we hope to find novel modifications that could further increase the efficacy and potency of the donors especially for *in vivo* genome editing and enable the scientific community in understanding biological problems using targeted genome editing.

References

1 Auerbach, C. The chemical production of mutations. The effect of chemical mutagens on cells and their genetic material is discussed. *Science* **158**, 1141-1147, doi:10.1126/science.158.3805.1141 (1967).

2 Brenner, S. The genetics of Caenorhabditis elegans. *Genetics* **77**, 71-94 (1974).

3 Coulondre, C. & Miller, J. H. Genetic studies of the lac repressor. III. Additional correlation of mutational sites with specific amino acid residues. *J Mol Biol* **117**, 525-567, doi:10.1016/0022-2836(77)90056-0 (1977).

4 Jackson, D. A., Symons, R. H. & Berg, P. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of Escherichia coli. *Proc Natl Acad Sci U S A* **69**, 2904-2909, doi:10.1073/pnas.69.10.2904 (1972).

5 Costantini, F. & Lacy, E. Introduction of a rabbit beta-globin gene into the mouse germ line. *Nature* **294**, 92-94, doi:10.1038/294092a0 (1981).

6 Cohen, S. N., Chang, A. C., Boyer, H. W. & Helling, R. B. Construction of biologically functional bacterial plasmids in vitro. *Proc Natl Acad Sci U S A* **70**, 3240-3244, doi:10.1073/pnas.70.11.3240 (1973).

7 Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A. & Kucherlapati, R. S. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* **317**, 230-234, doi:10.1038/317230a0 (1985).

8 Kim, Y. G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A* **93**, 1156-1160, doi:10.1073/pnas.93.3.1156 (1996).

9 Pavletich, N. P. & Pabo, C. O. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science* **252**, 809-817, doi:10.1126/science.2028256 (1991).

10 Kim, H. J., Lee, H. J., Kim, H., Cho, S. W. & Kim, J. S. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res* **19**, 1279-1288, doi:10.1101/gr.089417.108 (2009).

11 Bae, K. H. *et al.* Human zinc fingers as building blocks in the construction of artificial transcription factors. *Nat Biotechnol* **21**, 275-280, doi:10.1038/nbt796 (2003).

12 Greisman, H. A. & Pabo, C. O. A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. *Science* **275**, 657-661, doi:10.1126/science.275.5300.657 (1997).

13 Bibikova, M., Golic, M., Golic, K. G. & Carroll, D. Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics* **161**, 1169-1175 (2002).

14 Morton, J., Davis, M. W., Jorgensen, E. M. & Carroll, D. Induction and repair of zinc-finger nuclease-targeted double-strand breaks in Caenorhabditis elegans somatic cells. *Proc Natl Acad Sci U S A* **103**, 16370-16375, doi:10.1073/pnas.0605633103 (2006).
15 Bibikova, M., Beumer, K., Trautman, J. K. & Carroll, D. Enhancing gene targeting with designed zinc finger nucleases. *Science* **300**, 764, doi:10.1126/science.1079512 (2003).
16 Porteus, M. H. & Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. *Science* **300**, 763, doi:10.1126/science.1078395 (2003).
17 Urnov, F. D. *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646-651, doi:10.1038/nature03556 (2005).
18 Kim, H. *et al.* Surrogate reporters for enrichment of cells with nuclease-induced mutations. *Nat Methods* **8**, 941-943, doi:10.1038/nmeth.1733 (2011).
19 Boch, J. & Bonas, U. Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol* **48**, 419-436, doi:10.1146/annurev-phyto-080508-081936 (2010).
20 Boch, J. *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509-1512, doi:10.1126/science.1178811 (2009).
21 Joung, J. K. & Sander, J. D. TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol* **14**, 49-55, doi:10.1038/nrm3486 (2013).
22 Miller, J. C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* **29**, 143-148, doi:10.1038/nbt.1755 (2011).
23 Huang, P. *et al.* Heritable gene targeting in zebrafish using customized TALENs. *Nat Biotechnol* **29**, 699-700, doi:10.1038/nbt.1939 (2011).
24 Lei, Y. *et al.* Efficient targeted gene disruption in Xenopus embryos using engineered transcription activator-like effector nucleases (TALENs). *Proc Natl Acad Sci U S A* **109**, 17484-17489, doi:10.1073/pnas.1215421109 (2012).
25 Cheng, Z. *et al.* Conditional targeted genome editing using somatically expressed TALENs in C. elegans. *Nat Biotechnol* **31**, 934-937, doi:10.1038/nbt.2674 (2013).
26 Lo, T. W. *et al.* Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. *Genetics* **195**, 331-348, doi:10.1534/genetics.113.155382 (2013).
27 Mojica, F. J., Diez-Villasenor, C., Soria, E. & Juez, G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol Microbiol* **36**, 244-246, doi:10.1046/j.1365-2958.2000.01838.x (2000).

28 van Belkum, A., Scherer, S., van Alphen, L. & Verbrugh, H. Short-sequence DNA repeats in prokaryotic genomes. *Microbiol Mol Biol Rev* **62**, 275-293 (1998).
29 Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**, 2551-2561, doi:10.1099/mic.0.28048-0 (2005).
30 Jansen, R., Embden, J. D., Gaastra, W. & Schouls, L. M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* **43**, 1565-1575, doi:10.1046/j.1365-2958.2002.02839.x (2002).
31 Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. 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* **1**, 7, doi:10.1186/1745-6150-1-7 (2006).
32 Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709-1712, doi:10.1126/science.1138140 (2007).
33 Brouns, S. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960-964, doi:10.1126/science.1159689 (2008).
34 Marraffini, L. A. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**, 1843-1845, doi:10.1126/science.1165771 (2008).
35 Garneau, J. E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67-71, doi:10.1038/nature09523 (2010).
36 Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602-607, doi:10.1038/nature09886 (2011).
37 Makarova, K. S. *et al.* Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**, 467-477, doi:10.1038/nrmicro2577 (2011).
38 Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A* **109**, E2579-2586, doi:10.1073/pnas.1208507109 (2012).
39 Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821, doi:10.1126/science.1225829 (2012).
40 Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62-67, doi:10.1038/nature13011 (2014).
41 Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* **31**, 839-843, doi:10.1038/nbt.2673 (2013).

42 Chapman, J. R., Taylor, M. R. & Boulton, S. J. Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell* **47**, 497-510, doi:10.1016/j.molcel.2012.07.029 (2012).
43 Sung, P. & Klein, H. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol* **7**, 739-750, doi:10.1038/nrm2008 (2006).
44 Lieber, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* **79**, 181-211, doi:10.1146/annurev.biochem.052308.093131 (2010).
45 Daley, J. M., Palmbos, P. L., Wu, D. & Wilson, T. E. Nonhomologous end joining in yeast. *Annu Rev Genet* **39**, 431-451, doi:10.1146/annurev.genet.39.073003.113340 (2005).
46 Difilippantonio, M. J. *et al.* DNA repair protein Ku80 suppresses chromosomal aberrations and malignant transformation. *Nature* **404**, 510-514, doi:10.1038/35006670 (2000).
47 Li, G. C. *et al.* Ku70: a candidate tumor suppressor gene for murine T cell lymphoma. *Mol Cell* **2**, 1-8, doi:10.1016/s1097-2765(00)80108-2 (1998).
48 Ferguson, D. O. *et al.* The nonhomologous end-joining pathway of DNA repair is required for genomic stability and the suppression of translocations. *Proc Natl Acad Sci U S A* **97**, 6630-6633, doi:10.1073/pnas.110152897 (2000).
49 Boulton, S. J. & Jackson, S. P. Saccharomyces cerevisiae Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways. *EMBO J* **15**, 5093-5103 (1996).
50 Mimitou, E. P. & Symington, L. S. Ku prevents Exo1 and Sgs1-dependent resection of DNA ends in the absence of a functional MRX complex or Sae2. *EMBO J* **29**, 3358-3369, doi:10.1038/emboj.2010.193 (2010).
51 Yannone, S. M. *et al.* Coordinate 5' and 3' endonucleolytic trimming of terminally blocked blunt DNA double-strand break ends by Artemis nuclease and DNA-dependent protein kinase. *Nucleic Acids Res* **36**, 3354-3365, doi:10.1093/nar/gkn205 (2008).
52 Grawunder, U. *et al.* Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* **388**, 492-495, doi:10.1038/41358 (1997).
53 Ma, Y. *et al.* A biochemically defined system for mammalian nonhomologous DNA end joining. *Mol Cell* **16**, 701-713, doi:10.1016/j.molcel.2004.11.017 (2004).
54 Mahajan, K. N., Nick McElhinny, S. A., Mitchell, B. S. & Ramsden, D. A. Association of DNA polymerase mu (pol mu) with Ku and ligase IV: role for pol mu in end-joining double-strand break repair. *Mol Cell Biol* **22**, 5194-5202, doi:10.1128/mcb.22.14.5194-5202.2002 (2002).

55 Liang, F., Romanienko, P. J., Weaver, D. T., Jeggo, P. A. & Jasin, M. Chromosomal double-strand break repair in Ku80-deficient cells. *Proc Natl Acad Sci U S A* **93**, 8929-8933, doi:10.1073/pnas.93.17.8929 (1996).
56 Ma, J. L., Kim, E. M., Haber, J. E. & Lee, S. E. Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences. *Mol Cell Biol* **23**, 8820-8828, doi:10.1128/mcb.23.23.8820-8828.2003 (2003).
57 Sartori, A. A. *et al.* Human CtIP promotes DNA end resection. *Nature* **450**, 509-514, doi:10.1038/nature06337 (2007).
58 Wang, H. & Xu, X. Microhomology-mediated end joining: new players join the team. *Cell Biosci* **7**, 6, doi:10.1186/s13578-017-0136-8 (2017).
59 Xie, A., Kwok, A. & Scully, R. Role of mammalian Mre11 in classical and alternative nonhomologous end joining. *Nat Struct Mol Biol* **16**, 814-818, doi:10.1038/nsmb.1640 (2009).
60 Paull, T. T. & Gellert, M. The 3' to 5' exonuclease activity of Mre 11 facilitates repair of DNA double-strand breaks. *Mol Cell* **1**, 969-979, doi:10.1016/s1097-2765(00)80097-0 (1998).
61 Truong, L. N. *et al.* Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proc Natl Acad Sci U S A* **110**, 7720-7725, doi:10.1073/pnas.1213431110 (2013).
62 Rass, E. *et al.* Role of Mre11 in chromosomal nonhomologous end joining in mammalian cells. *Nat Struct Mol Biol* **16**, 819-824, doi:10.1038/nsmb.1641 (2009).
63 Sfeir, A. & Symington, L. S. Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? *Trends Biochem Sci* **40**, 701-714, doi:10.1016/j.tibs.2015.08.006 (2015).
64 Iyer, S. *et al.* Precise therapeutic gene correction by a simple nuclease-induced double-stranded break. *Nature* **568**, 561-565, doi:10.1038/s41586-019-1076-8 (2019).
65 Surralles, J. *et al.* Molecular cross-talk among chromosome fragility syndromes. *Genes Dev* **18**, 1359-1370, doi:10.1101/gad.1216304 (2004).
66 Hickson, I. D. RecQ helicases: caretakers of the genome. *Nat Rev Cancer* **3**, 169-178, doi:10.1038/nrc1012 (2003).
67 Keeney, S., Giroux, C. N. & Kleckner, N. Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* **88**, 375-384, doi:10.1016/s0092-8674(00)81876-0 (1997).
68 McKim, K. S. & Hayashi-Hagihara, A. mei-W68 in Drosophila melanogaster encodes a Spo11 homolog: evidence that the mechanism for initiating meiotic recombination is conserved. *Genes Dev* **12**, 2932-2942, doi:10.1101/gad.12.18.2932 (1998).

69 Dernburg, A. F. *et al.* Meiotic recombination in C. elegans initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. *Cell* **94**, 387-398, doi:10.1016/s0092-8674(00)81481-6 (1998).
70 Romanienko, P. J. & Camerini-Otero, R. D. The mouse Spo11 gene is required for meiotic chromosome synapsis. *Mol Cell* **6**, 975-987, doi:10.1016/s1097-2765(00)00097-6 (2000).
71 Keeney, S. & Kleckner, N. Covalent protein-DNA complexes at the 5' strand termini of meiosis-specific double-strand breaks in yeast. *Proc Natl Acad Sci U S A* **92**, 11274-11278, doi:10.1073/pnas.92.24.11274 (1995).
72 Liu, J., Wu, T. C. & Lichten, M. The location and structure of double-strand DNA breaks induced during yeast meiosis: evidence for a covalently linked DNA-protein intermediate. *EMBO J* **14**, 4599-4608 (1995).
73 Mimitou, E. P. & Symington, L. S. Nucleases and helicases take center stage in homologous recombination. *Trends Biochem Sci* **34**, 264-272, doi:10.1016/j.tibs.2009.01.010 (2009).
74 Neale, M. J., Pan, J. & Keeney, S. Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature* **436**, 1053-1057, doi:10.1038/nature03872 (2005).
75 Liu, Y. & West, S. C. Happy Hollidays: 40th anniversary of the Holliday junction. *Nat Rev Mol Cell Biol* **5**, 937-944, doi:10.1038/nrm1502 (2004).
76 Garcia, V., Phelps, S. E., Gray, S. & Neale, M. J. Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1. *Nature* **479**, 241-244, doi:10.1038/nature10515 (2011).
77 Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* **34**, 184-191, doi:10.1038/nbt.3437 (2016).
78 Doench, J. G. *et al.* Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol* **32**, 1262-1267, doi:10.1038/nbt.3026 (2014).
79 Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80-84, doi:10.1126/science.1246981 (2014).
80 Moreno-Mateos, M. A. *et al.* CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat Methods* **12**, 982-988, doi:10.1038/nmeth.3543 (2015).
81 Farboud, B. & Meyer, B. J. Dramatic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. *Genetics* **199**, 959-971, doi:10.1534/genetics.115.175166 (2015).
82 Xu, H. *et al.* Sequence determinants of improved CRISPR sgRNA design. *Genome Res* **25**, 1147-1157, doi:10.1101/gr.191452.115 (2015).
83 Wang, D. *et al.* Optimized CRISPR guide RNA design for two high-fidelity Cas9 variants by deep learning. *Nat Commun* **10**, 4284, doi:10.1038/s41467-019-12281-8 (2019).

84 Horlbeck, M. A. *et al.* Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *Elife* **5**, doi:10.7554/eLife.12677 (2016).
85 Uusi-Makela, M. I. E. *et al.* Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (Danio rerio). *PLoS One* **13**, e0196238, doi:10.1371/journal.pone.0196238 (2018).
86 Kim, H. K. *et al.* SpCas9 activity prediction by DeepSpCas9, a deep learning-based model with high generalization performance. *Sci Adv* **5**, eaax9249, doi:10.1126/sciadv.aax9249 (2019).
87 Kim, H. K. *et al.* Deep learning improves prediction of CRISPR-Cpf1 guide RNA activity. *Nat Biotechnol* **36**, 239-241, doi:10.1038/nbt.4061 (2018).
88 Kim, S., Kim, D., Cho, S. W., Kim, J. & Kim, J. S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res* **24**, 1012-1019, doi:10.1101/gr.171322.113 (2014).
89 Liang, X. *et al.* Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *J Biotechnol* **208**, 44-53, doi:10.1016/j.jbiotec.2015.04.024 (2015).
90 Merkle, F. T. *et al.* Efficient CRISPR-Cas9-mediated generation of knockin human pluripotent stem cells lacking undesired mutations at the targeted locus. *Cell Rep* **11**, 875-883, doi:10.1016/j.celrep.2015.04.007 (2015).
91 Yang, S., Li, S. & Li, X. J. Shortening the Half-Life of Cas9 Maintains Its Gene Editing Ability and Reduces Neuronal Toxicity. *Cell Rep* **25**, 2653-2659 e2653, doi:10.1016/j.celrep.2018.11.019 (2018).
92 Han, H. A., Pang, J. K. S. & Soh, B. S. Mitigating off-target effects in CRISPR/Cas9-mediated in vivo gene editing. *J Mol Med (Berl)* **98**, 615-632, doi:10.1007/s00109-020-01893-z (2020).
93 Houseley, J. & Tollervey, D. The many pathways of RNA degradation. *Cell* **136**, 763-776, doi:10.1016/j.cell.2009.01.019 (2009).
94 Liu, G. & Gack, M. U. Distinct and Orchestrated Functions of RNA Sensors in Innate Immunity. *Immunity* **53**, 26-42, doi:10.1016/j.immuni.2020.03.017 (2020).
95 Wienert, B., Shin, J., Zelin, E., Pestal, K. & Corn, J. E. In vitro-transcribed guide RNAs trigger an innate immune response via the RIG-I pathway. *PLoS Biol* **16**, e2005840, doi:10.1371/journal.pbio.2005840 (2018).
96 Braasch, D. A. *et al.* RNA interference in mammalian cells by chemically-modified RNA. *Biochemistry* **42**, 7967-7975, doi:10.1021/bi0343774 (2003).
97 Deleavey, G. F. & Damha, M. J. Designing chemically modified oligonucleotides for targeted gene silencing. *Chem Biol* **19**, 937-954, doi:10.1016/j.chembiol.2012.07.011 (2012).
98 Behlke, M. A. Chemical modification of siRNAs for in vivo use. *Oligonucleotides* **18**, 305-319, doi:10.1089/oli.2008.0164 (2008).

99 Mir, A. *et al.* Heavily and fully modified RNAs guide efficient SpyCas9-mediated genome editing. *Nat Commun* **9**, 2641, doi:10.1038/s41467-018-05073-z (2018).
100 Lee, K. *et al.* Synthetically modified guide RNA and donor DNA are a versatile platform for CRISPR-Cas9 engineering. *Elife* **6**, doi:10.7554/eLife.25312 (2017).
101 Yin, H. *et al.* Structure-guided chemical modification of guide RNA enables potent non-viral in vivo genome editing. *Nat Biotechnol* **35**, 1179-1187, doi:10.1038/nbt.4005 (2017).
102 Rahdar, M. *et al.* Synthetic CRISPR RNA-Cas9-guided genome editing in human cells. *Proc Natl Acad Sci U S A* **112**, E7110-7117, doi:10.1073/pnas.1520883112 (2015).
103 Hendel, A. *et al.* Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat Biotechnol* **33**, 985-989, doi:10.1038/nbt.3290 (2015).
104 Kelley, M. L., Strezoska, Z., He, K., Vermeulen, A. & Smith, A. Versatility of chemically synthesized guide RNAs for CRISPR-Cas9 genome editing. *J Biotechnol* **233**, 74-83, doi:10.1016/j.jbiotec.2016.06.011 (2016).
105 Vakulskas, C. A. *et al.* A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat Med* **24**, 1216-1224, doi:10.1038/s41591-018-0137-0 (2018).
106 Ryan, D. E. *et al.* Improving CRISPR-Cas specificity with chemical modifications in single-guide RNAs. *Nucleic Acids Res* **46**, 792-803, doi:10.1093/nar/gkx1199 (2018).
107 Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**, 822-826, doi:10.1038/nbt.2623 (2013).
108 Kleinstiver, B. P. *et al.* High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490-495, doi:10.1038/nature16526 (2016).
109 Yin, H. *et al.* Partial DNA-guided Cas9 enables genome editing with reduced off-target activity. *Nat Chem Biol* **14**, 311-316, doi:10.1038/nchembio.2559 (2018).
110 Amrani, N. *et al.* NmeCas9 is an intrinsically high-fidelity genome-editing platform. *Genome Biol* **19**, 214, doi:10.1186/s13059-018-1591-1 (2018).
111 Edraki, A. *et al.* A Compact, High-Accuracy Cas9 with a Dinucleotide PAM for In Vivo Genome Editing. *Mol Cell* **73**, 714-726 e714, doi:10.1016/j.molcel.2018.12.003 (2019).
112 Rueda, F. O. *et al.* Mapping the sugar dependency for rational generation of a DNA-RNA hybrid-guided Cas9 endonuclease. *Nat Commun* **8**, 1610, doi:10.1038/s41467-017-01732-9 (2017).

113 Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol* **32**, 279-284, doi:10.1038/nbt.2808 (2014).
114 Pavel-Dinu, M. *et al.* Gene correction for SCID-X1 in long-term hematopoietic stem cells. *Nat Commun* **10**, 1634, doi:10.1038/s41467-019-09614-y (2019).
115 Kocak, D. D. *et al.* Increasing the specificity of CRISPR systems with engineered RNA secondary structures. *Nat Biotechnol* **37**, 657-666, doi:10.1038/s41587-019-0095-1 (2019).
116 Tang, W., Hu, J. H. & Liu, D. R. Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation. *Nat Commun* **8**, 15939, doi:10.1038/ncomms15939 (2017).
117 Chen, B. *et al.* Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479-1491, doi:10.1016/j.cell.2013.12.001 (2013).
118 Grevet, J. D. *et al.* Domain-focused CRISPR screen identifies HRI as a fetal hemoglobin regulator in human erythroid cells. *Science* **361**, 285-290, doi:10.1126/science.aao0932 (2018).
119 Li, Y. *et al.* Genomic Surveillance of Streptococcus pyogenes Strains Causing Invasive Disease, United States, 2016-2017. *Front Microbiol* **11**, 1547, doi:10.3389/fmicb.2020.01547 (2020).
120 Charlesworth, C. T. *et al.* Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat Med* **25**, 249-254, doi:10.1038/s41591-018-0326-x (2019).
121 Datsenko, K. A. *et al.* Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat Commun* **3**, 945, doi:10.1038/ncomms1937 (2012).
122 Simhadri, V. L. *et al.* Prevalence of Pre-existing Antibodies to CRISPR-Associated Nuclease Cas9 in the USA Population. *Mol Ther Methods Clin Dev* **10**, 105-112, doi:10.1016/j.omtm.2018.06.006 (2018).
123 Wagner, D. L. *et al.* High prevalence of Streptococcus pyogenes Cas9-reactive T cells within the adult human population. *Nat Med* **25**, 242-248, doi:10.1038/s41591-018-0204-6 (2019).
124 OMIM. doi:https://www.omim.org/statistics/geneMap.
125 Nami, F. *et al.* Strategies for In Vivo Genome Editing in Nondividing Cells. *Trends Biotechnol* **36**, 770-786, doi:10.1016/j.tibtech.2018.03.004 (2018).
126 Orthwein, A. *et al.* A mechanism for the suppression of homologous recombination in G1 cells. *Nature* **528**, 422-426, doi:10.1038/nature16142 (2015).
127 Orthwein, A. *et al.* Mitosis inhibits DNA double-strand break repair to guard against telomere fusions. *Science* **344**, 189-193, doi:10.1126/science.1248024 (2014).

128 Lees-Miller, S. P. DNA double strand break repair in mitosis is suppressed by phosphorylation of XRCC4. *PLoS Genet* **10**, e1004598, doi:10.1371/journal.pgen.1004598 (2014).
129 Ludtke, J. J., Zhang, G., Sebestyen, M. G. & Wolff, J. A. A nuclear localization signal can enhance both the nuclear transport and expression of 1 kb DNA. *J Cell Sci* **112 (Pt 12)**, 2033-2041 (1999).
130 Auer, T. O., Duroure, K., De Cian, A., Concordet, J. P. & Del Bene, F. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res* **24**, 142-153, doi:10.1101/gr.161638.113 (2014).
131 Suzuki, K. *et al.* In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* **540**, 144-149, doi:10.1038/nature20565 (2016).
132 Lin, S., Staahl, B. T., Alla, R. K. & Doudna, J. A. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife* **3**, e04766, doi:10.7554/eLife.04766 (2014).
133 Lomova, A. *et al.* Improving Gene Editing Outcomes in Human Hematopoietic Stem and Progenitor Cells by Temporal Control of DNA Repair. *Stem Cells* **37**, 284-294, doi:10.1002/stem.2935 (2019).
134 Chu, V. T. *et al.* Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotechnol* **33**, 543-548, doi:10.1038/nbt.3198 (2015).
135 Maruyama, T. *et al.* Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol* **33**, 538-542, doi:10.1038/nbt.3190 (2015).
136 Canny, M. D. *et al.* Inhibition of 53BP1 favors homology-dependent DNA repair and increases CRISPR-Cas9 genome-editing efficiency. *Nat Biotechnol* **36**, 95-102, doi:10.1038/nbt.4021 (2018).
137 Riesenberg, S. & Maricic, T. Targeting repair pathways with small molecules increases precise genome editing in pluripotent stem cells. *Nat Commun* **9**, 2164, doi:10.1038/s41467-018-04609-7 (2018).
138 Robert, F., Barbeau, M., Ethier, S., Dostie, J. & Pelletier, J. Pharmacological inhibition of DNA-PK stimulates Cas9-mediated genome editing. *Genome Med* **7**, 93, doi:10.1186/s13073-015-0215-6 (2015).
139 Wienert, B. *et al.* Timed inhibition of CDC7 increases CRISPR-Cas9 mediated templated repair. *Nat Commun* **11**, 2109, doi:10.1038/s41467-020-15845-1 (2020).
140 Ward, J. D. Rapid and precise engineering of the Caenorhabditis elegans genome with lethal mutation co-conversion and inactivation of NHEJ repair. *Genetics* **199**, 363-377, doi:10.1534/genetics.114.172361 (2015).
141 Beumer, K. J. *et al.* Efficient gene targeting in Drosophila by direct embryo injection with zinc-finger nucleases. *Proc Natl Acad Sci U S A* **105**, 19821-19826, doi:10.1073/pnas.0810475105 (2008).

142 Gao, Y. *et al.* A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell* **95**, 891-902, doi:10.1016/s0092-8674(00)81714-6 (1998).
143 Frank, K. M. *et al.* Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV. *Nature* **396**, 173-177, doi:10.1038/24172 (1998).
144 Srivastava, M. *et al.* An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. *Cell* **151**, 1474-1487, doi:10.1016/j.cell.2012.11.054 (2012).
145 Yu, C. *et al.* Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell* **16**, 142-147, doi:10.1016/j.stem.2015.01.003 (2015).
146 Pinder, J., Salsman, J. & Dellaire, G. Nuclear domain 'knock-in' screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. *Nucleic Acids Res* **43**, 9379-9392, doi:10.1093/nar/gkv993 (2015).
147 Jayathilaka, K. *et al.* A chemical compound that stimulates the human homologous recombination protein RAD51. *Proc Natl Acad Sci U S A* **105**, 15848-15853, doi:10.1073/pnas.0808046105 (2008).
148 Zhang, Y., Zhang, Z. & Ge, W. An efficient platform for generating somatic point mutations with germline transmission in the zebrafish by CRISPR/Cas9-mediated gene editing. *J Biol Chem* **293**, 6611-6622, doi:10.1074/jbc.RA117.001080 (2018).
149 Yang, H. *et al.* Methods Favoring Homology-Directed Repair Choice in Response to CRISPR/Cas9 Induced-Double Strand Breaks. *Int J Mol Sci* **21**, doi:10.3390/ijms21186461 (2020).
150 Paulsen, B. S. *et al.* Ectopic expression of RAD52 and dn53BP1 improves homology-directed repair during CRISPR-Cas9 genome editing. *Nat Biomed Eng* **1**, 878-888, doi:10.1038/s41551-017-0145-2 (2017).
151 Shao, S. *et al.* Enhancing CRISPR/Cas9-mediated homology-directed repair in mammalian cells by expressing Saccharomyces cerevisiae Rad52. *Int J Biochem Cell Biol* **92**, 43-52, doi:10.1016/j.biocel.2017.09.012 (2017).
152 Yuan, J. & Chen, J. N terminus of CtIP is critical for homologous recombination-mediated double-strand break repair. *J Biol Chem* **284**, 31746-31752, doi:10.1074/jbc.M109.023424 (2009).
153 Charpentier, M. *et al.* CtIP fusion to Cas9 enhances transgene integration by homology-dependent repair. *Nat Commun* **9**, 1133, doi:10.1038/s41467-018-03475-7 (2018).
154 Tran, N. T. *et al.* Enhancement of Precise Gene Editing by the Association of Cas9 With Homologous Recombination Factors. *Front Genet* **10**, 365, doi:10.3389/fgene.2019.00365 (2019).

155 Davis, L. & Maizels, N. Two Distinct Pathways Support Gene Correction by Single-Stranded Donors at DNA Nicks. *Cell Rep* **17**, 1872-1881, doi:10.1016/j.celrep.2016.10.049 (2016).
156 Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L. & Corn, J. E. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol* **34**, 339-344, doi:10.1038/nbt.3481 (2016).
157 Renaud, J. B. *et al.* Improved Genome Editing Efficiency and Flexibility Using Modified Oligonucleotides with TALEN and CRISPR-Cas9 Nucleases. *Cell Rep* **14**, 2263-2272, doi:10.1016/j.celrep.2016.02.018 (2016).
158 Okamoto, S., Amaishi, Y., Maki, I., Enoki, T. & Mineno, J. Highly efficient genome editing for single-base substitutions using optimized ssODNs with Cas9-RNPs. *Sci Rep* **9**, 4811, doi:10.1038/s41598-019-41121-4 (2019).
159 Lanza, D. G. *et al.* Comparative analysis of single-stranded DNA donors to generate conditional null mouse alleles. *BMC Biol* **16**, 69, doi:10.1186/s12915-018-0529-0 (2018).
160 Zhao, P., Zhang, Z., Ke, H., Yue, Y. & Xue, D. Oligonucleotide-based targeted gene editing in C. elegans via the CRISPR/Cas9 system. *Cell Res* **24**, 247-250, doi:10.1038/cr.2014.9 (2014).
161 Yang, H. *et al.* One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* **154**, 1370-1379, doi:10.1016/j.cell.2013.08.022 (2013).
162 Quadros, R. M. *et al.* Easi-CRISPR: a robust method for one-step generation of mice carrying conditional and insertion alleles using long ssDNA donors and CRISPR ribonucleoproteins. *Genome Biol* **18**, 92, doi:10.1186/s13059-017-1220-4 (2017).
163 Yoshimi, K. *et al.* ssODN-mediated knock-in with CRISPR-Cas for large genomic regions in zygotes. *Nat Commun* **7**, 10431, doi:10.1038/ncomms10431 (2016).
164 Nguyen, D. N. *et al.* Polymer-stabilized Cas9 nanoparticles and modified repair templates increase genome editing efficiency. *Nat Biotechnol* **38**, 44-49, doi:10.1038/s41587-019-0325-6 (2020).
165 Mazina, O. M., Keskin, H., Hanamshet, K., Storici, F. & Mazin, A. V. Rad52 Inverse Strand Exchange Drives RNA-Templated DNA Double-Strand Break Repair. *Mol Cell* **67**, 19-29 e13, doi:10.1016/j.molcel.2017.05.019 (2017).
166 Meers, C. *et al.* Genetic Characterization of Three Distinct Mechanisms Supporting RNA-Driven DNA Repair and Modification Reveals Major Role of DNA Polymerase zeta. *Mol Cell* **79**, 1037-1050 e1035, doi:10.1016/j.molcel.2020.08.011 (2020).
167 Li, S. *et al.* Precise gene replacement in rice by RNA transcript-templated homologous recombination. *Nat Biotechnol* **37**, 445-450, doi:10.1038/s41587-019-0065-7 (2019).

168 Chandramouly, G. *et al.* Poltheta reverse transcribes RNA and promotes RNA-templated DNA repair. *Sci Adv* **7**, doi:10.1126/sciadv.abf1771 (2021).
169 Iyer, S. *et al.* Efficient Homology-directed Repair with Circular ssDNA Donors. *bioRxiv*, 864199, doi:10.1101/864199 (2019).
170 Wilde, J. J. *et al.* Efficient embryonic homozygous gene conversion via RAD51-enhanced interhomolog repair. *Cell* **184**, 3267-3280 e3218, doi:10.1016/j.cell.2021.04.035 (2021).
171 Wang, D. *et al.* Cas9-mediated allelic exchange repairs compound heterozygous recessive mutations in mice. *Nat Biotechnol* **36**, 839-842, doi:10.1038/nbt.4219 (2018).
172 Carlson-Stevermer, J. *et al.* Assembly of CRISPR ribonucleoproteins with biotinylated oligonucleotides via an RNA aptamer for precise gene editing. *Nat Commun* **8**, 1711, doi:10.1038/s41467-017-01875-9 (2017).
173 Gu, B., Posfai, E. & Rossant, J. Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. *Nat Biotechnol* **36**, 632-637, doi:10.1038/nbt.4166 (2018).
174 Ma, M. *et al.* Efficient generation of mice carrying homozygous double-floxp alleles using the Cas9-Avidin/Biotin-donor DNA system. *Cell Res* **27**, 578-581, doi:10.1038/cr.2017.29 (2017).
175 Roche, P. J. R. *et al.* Double-Stranded Biotinylated Donor Enhances Homology-Directed Repair in Combination with Cas9 Monoavidin in Mammalian Cells. *CRISPR J* **1**, 414-430, doi:10.1089/crispr.2018.0045 (2018).
176 Savic, N. *et al.* Covalent linkage of the DNA repair template to the CRISPR-Cas9 nuclease enhances homology-directed repair. *Elife* **7**, doi:10.7554/eLife.33761 (2018).
177 Li, G., Wang, H., Zhang, X., Wu, Z. & Yang, H. A Cas9-transcription factor fusion protein enhances homology-directed repair efficiency. *J Biol Chem* **296**, 100525, doi:10.1016/j.jbc.2021.100525 (2021).
178 Aird, E. J., Lovendahl, K. N., St Martin, A., Harris, R. S. & Gordon, W. R. Increasing Cas9-mediated homology-directed repair efficiency through covalent tethering of DNA repair template. *Commun Biol* **1**, 54, doi:10.1038/s42003-018-0054-2 (2018).
179 Lovendahl, K. N., Hayward, A. N. & Gordon, W. R. Sequence-Directed Covalent Protein-DNA Linkages in a Single Step Using HUH-Tags. *J Am Chem Soc* **139**, 7030-7035, doi:10.1021/jacs.7b02572 (2017).
180 Liu, P. *et al.* Enhanced Cas12a editing in mammalian cells and zebrafish. *Nucleic Acids Res* **47**, 4169-4180, doi:10.1093/nar/gkz184 (2019).
181 Finn, J. D. *et al.* A Single Administration of CRISPR/Cas9 Lipid Nanoparticles Achieves Robust and Persistent In Vivo Genome Editing. *Cell Rep* **22**, 2227-2235, doi:10.1016/j.celrep.2018.02.014 (2018).

182 Hirakawa, M. P., Krishnakumar, R., Timlin, J. A., Carney, J. P. & Butler, K. S. Gene editing and CRISPR in the clinic: current and future perspectives. *Biosci Rep* **40**, doi:10.1042/BSR20200127 (2020).
183 Frangoul, H. *et al.* CRISPR-Cas9 Gene Editing for Sickle Cell Disease and beta-Thalassemia. *N Engl J Med* **384**, 252-260, doi:10.1056/NEJMoa2031054 (2021).
184 Musunuru, K. *et al.* In vivo CRISPR base editing of PCSK9 durably lowers cholesterol in primates. *Nature* **593**, 429-434, doi:10.1038/s41586-021-03534-y (2021).
185 Frokjaer-Jensen, C., Davis, M. W., Ailion, M. & Jorgensen, E. M. Improved Mos1-mediated transgenesis in C. elegans. *Nat Methods* **9**, 117-118, doi:10.1038/nmeth.1865 (2012).
186 Frokjaer-Jensen, C. *et al.* Single-copy insertion of transgenes in Caenorhabditis elegans. *Nat Genet* **40**, 1375-1383, doi:10.1038/ng.248 (2008).
187 Dickinson, D. J., Ward, J. D., Reiner, D. J. & Goldstein, B. Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination. *Nat Methods* **10**, 1028-1034, doi:10.1038/nmeth.2641 (2013).
188 Friedland, A. E. *et al.* Heritable genome editing in C. elegans via a CRISPR-Cas9 system. *Nat Methods* **10**, 741-743, doi:10.1038/nmeth.2532 (2013).
189 Katic, I. & Grosshans, H. Targeted heritable mutation and gene conversion by Cas9-CRISPR in Caenorhabditis elegans. *Genetics* **195**, 1173-1176, doi:10.1534/genetics.113.155754 (2013).
190 Waaijers, S. *et al.* CRISPR/Cas9-targeted mutagenesis in Caenorhabditis elegans. *Genetics* **195**, 1187-1191, doi:10.1534/genetics.113.156299 (2013).
191 Chiu, H., Schwartz, H. T., Antoshechkin, I. & Sternberg, P. W. Transgene-free genome editing in Caenorhabditis elegans using CRISPR-Cas. *Genetics* **195**, 1167-1171, doi:10.1534/genetics.113.155879 (2013).
192 Kim, H. *et al.* A co-CRISPR strategy for efficient genome editing in Caenorhabditis elegans. *Genetics* **197**, 1069-1080, doi:10.1534/genetics.114.166389 (2014).
193 Norris, A. D., Kim, H. M., Colaiacovo, M. P. & Calarco, J. A. Efficient Genome Editing in Caenorhabditis elegans with a Toolkit of Dual-Marker Selection Cassettes. *Genetics* **201**, 449-458, doi:10.1534/genetics.115.180679 (2015).
194 Arribere, J. A. *et al.* Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in Caenorhabditis elegans. *Genetics* **198**, 837-846, doi:10.1534/genetics.114.169730 (2014).
195 Shirayama, M. *et al.* piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. *Cell* **150**, 65-77, doi:10.1016/j.cell.2012.06.015 (2012).

196 Kelly, W. G., Xu, S., Montgomery, M. K. & Fire, A. Distinct requirements for somatic and germline expression of a generally expressed Caernorhabditis elegans gene. *Genetics* **146**, 227-238 (1997).
197 Kim, J. K. *et al.* Functional genomic analysis of RNA interference in C. elegans. *Science* **308**, 1164-1167, doi:10.1126/science.1109267 (2005).
198 Stiernagle, T. Maintenance of C. elegans. *WormBook*, 1-11, doi:10.1895/wormbook.1.101.1 (2006).
199 Fire, A. Integrative transformation of Caenorhabditis elegans. *EMBO J* **5**, 2673-2680 (1986).
200 Mello, C. & Fire, A. DNA transformation. *Methods Cell Biol* **48**, 451-482 (1995).
201 Dickinson, D. J., Pani, A. M., Heppert, J. K., Higgins, C. D. & Goldstein, B. Streamlined Genome Engineering with a Self-Excising Drug Selection Cassette. *Genetics* **200**, 1035-1049, doi:10.1534/genetics.115.178335 (2015).
202 Schwartz, M. L. & Jorgensen, E. M. SapTrap, a Toolkit for High-Throughput CRISPR/Cas9 Gene Modification in Caenorhabditis elegans. *Genetics* **202**, 1277-1288, doi:10.1534/genetics.115.184275 (2016).
203 Ghanta, K. S. & Mello, C. C. Melting dsDNA Donor Molecules Greatly Improves Precision Genome Editing in Caenorhabditis elegans. *Genetics* **216**, 643-650, doi:10.1534/genetics.120.303564 (2020).
204 Ebbing, A., Shang, P., Geijsen, N. & Korswagen, H. Extending the CRISPR toolbox for C. elegans: Cpf1 as an alternative gene editing system for AT-rich sequences. *MicroPubl Biol* **2017**, doi:10.17912/W2237D (2017).
205 Labun, K. *et al.* CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res* **47**, W171-W174, doi:10.1093/nar/gkz365 (2019).
206 Concordet, J. P. & Haeussler, M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res* **46**, W242-W245, doi:10.1093/nar/gky354 (2018).
207 Ghanta, K. S. *et al.* 5′ Modifications Improve Potency and Efficacy of DNA Donors for Precision Genome Editing. *bioRxiv*, 354480, doi:10.1101/354480 (2021).
208 Dokshin, G. A., Ghanta, K. S., Piscopo, K. M. & Mello, C. C. Robust Genome Editing with Short Single-Stranded and Long, Partially Single-Stranded DNA Donors in Caenorhabditis elegans. *Genetics* **210**, 781-787, doi:10.1534/genetics.118.301532 (2018).
209 Prior, H., Jawad, A. K., MacConnachie, L. & Beg, A. A. Highly Efficient, Rapid and Co-CRISPR-Independent Genome Editing in Caenorhabditis elegans. *G3 (Bethesda)* **7**, 3693-3698, doi:10.1534/g3.117.300216 (2017).
210 Farboud, B., Severson, A. F. & Meyer, B. J. Strategies for Efficient Genome Editing Using CRISPR-Cas9. *Genetics* **211**, 431-457, doi:10.1534/genetics.118.301775 (2019).

211 Paix, A., Folkmann, A. & Seydoux, G. Precision genome editing using CRISPR-Cas9 and linear repair templates in C. elegans. *Methods* **121-122**, 86-93, doi:10.1016/j.ymeth.2017.03.023 (2017).
212 Paix, A., Folkmann, A., Rasoloson, D. & Seydoux, G. High Efficiency, Homology-Directed Genome Editing in Caenorhabditis elegans Using CRISPR-Cas9 Ribonucleoprotein Complexes. *Genetics* **201**, 47-54, doi:10.1534/genetics.115.179382 (2015).
213 Vicencio, J., Martinez-Fernandez, C., Serrat, X. & Ceron, J. Efficient Generation of Endogenous Fluorescent Reporters by Nested CRISPR in Caenorhabditis elegans. *Genetics* **211**, 1143-1154, doi:10.1534/genetics.119.301965 (2019).
214 Mello, C. C., Kramer, J. M., Stinchcomb, D. & Ambros, V. Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. *EMBO J* **10**, 3959-3970 (1991).
215 Stinchcomb, D. T., Shaw, J. E., Carr, S. H. & Hirsh, D. Extrachromosomal DNA transformation of Caenorhabditis elegans. *Mol Cell Biol* **5**, 3484-3496, doi:10.1128/mcb.5.12.3484 (1985).
216 Brinkman, E. K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* **42**, e168, doi:10.1093/nar/gku936 (2014).
217 Cho, S. W., Lee, J., Carroll, D., Kim, J. S. & Lee, J. Heritable gene knockout in Caenorhabditis elegans by direct injection of Cas9-sgRNA ribonucleoproteins. *Genetics* **195**, 1177-1180, doi:10.1534/genetics.113.155853 (2013).
218 Richardson, C. D. *et al.* CRISPR-Cas9 genome editing in human cells occurs via the Fanconi anemia pathway. *Nat Genet* **50**, 1132-1139, doi:10.1038/s41588-018-0174-0 (2018).
219 Paix, A., Schmidt, H. & Seydoux, G. Cas9-assisted recombineering in C. elegans: genome editing using in vivo assembly of linear DNAs. *Nucleic Acids Res* **44**, e128, doi:10.1093/nar/gkw502 (2016).
220 Silva-Garcia, C. G. *et al.* Single-Copy Knock-In Loci for Defined Gene Expression in Caenorhabditis elegans. *G3 (Bethesda)* **9**, 2195-2198, doi:10.1534/g3.119.400314 (2019).
221 Tzur, Y. B. *et al.* Heritable custom genomic modifications in Caenorhabditis elegans via a CRISPR-Cas9 system. *Genetics* **195**, 1181-1185, doi:10.1534/genetics.113.156075 (2013).
222 Forbes, D. J., Kirschner, M. W. & Newport, J. W. Spontaneous formation of nucleus-like structures around bacteriophage DNA microinjected into Xenopus eggs. *Cell* **34**, 13-23, doi:10.1016/0092-8674(83)90132-0 (1983).
223 Ebbing, A., Shang, P., Geijsen, N. & Korswagen, H. (2017).
224 Paix, A. *et al.* Scalable and versatile genome editing using linear DNAs with microhomology to Cas9 Sites in Caenorhabditis elegans. *Genetics* **198**, 1347-1356, doi:10.1534/genetics.114.170423 (2014).

225 McCarter, J., Bartlett, B., Dang, T. & Schedl, T. On the control of oocyte meiotic maturation and ovulation in Caenorhabditis elegans. *Dev Biol* **205**, 111-128, doi:10.1006/dbio.1998.9109 (1999).
226 Nambiar, T. S. *et al.* Stimulation of CRISPR-mediated homology-directed repair by an engineered RAD18 variant. *Nat Commun* **10**, 3395, doi:10.1038/s41467-019-11105-z (2019).
227 Rees, H. A., Yeh, W. H. & Liu, D. R. Development of hRad51-Cas9 nickase fusions that mediate HDR without double-stranded breaks. *Nat Commun* **10**, 2212, doi:10.1038/s41467-019-09983-4 (2019).
228 Frank-Vaillant, M. & Marcand, S. Transient stability of DNA ends allows nonhomologous end joining to precede homologous recombination. *Mol Cell* **10**, 1189-1199, doi:10.1016/s1097-2765(02)00705-0 (2002).
229 Mao, Z., Bozzella, M., Seluanov, A. & Gorbunova, V. Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA Repair (Amst)* **7**, 1765-1771, doi:10.1016/j.dnarep.2008.06.018 (2008).
230 Perucho, M., Hanahan, D. & Wigler, M. Genetic and physical linkage of exogenous sequences in transformed cells. *Cell* **22**, 309-317, doi:10.1016/0092-8674(80)90178-6 (1980).
231 Folger, K. R., Wong, E. A., Wahl, G. & Capecchi, M. R. Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules. *Mol Cell Biol* **2**, 1372-1387, doi:10.1128/mcb.2.11.1372 (1982).
232 Stuart, G. W., McMurray, J. V. & Westerfield, M. Replication, integration and stable germ-line transmission of foreign sequences injected into early zebrafish embryos. *Development* **103**, 403-412 (1988).
233 Lacy, E., Roberts, S., Evans, E. P., Burtenshaw, M. D. & Costantini, F. D. A foreign beta-globin gene in transgenic mice: integration at abnormal chromosomal positions and expression in inappropriate tissues. *Cell* **34**, 343-358, doi:10.1016/0092-8674(83)90369-0 (1983).
234 Li, H. *et al.* Design and specificity of long ssDNA donors for CRISPR-based knock-in. *bioRxiv*, 178905, doi:10.1101/178905 (2019).
235 Clement, K. *et al.* CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat Biotechnol* **37**, 224-226, doi:10.1038/s41587-019-0032-3 (2019).
236 Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* **33**, 187-197, doi:10.1038/nbt.3117 (2015).
237 Certo, M. T. *et al.* Tracking genome engineering outcome at individual DNA breakpoints. *Nat Methods* **8**, 671-676, doi:10.1038/nmeth.1648 (2011).
238 Glaser, A., McColl, B. & Vadolas, J. GFP to BFP Conversion: A Versatile Assay for the Quantification of CRISPR/Cas9-mediated Genome Editing. *Mol Ther Nucleic Acids* **5**, e334, doi:10.1038/mtna.2016.48 (2016).

239 Kan, Y., Ruis, B., Takasugi, T. & Hendrickson, E. A. Mechanisms of precise genome editing using oligonucleotide donors. *Genome Res* **27**, 1099-1111, doi:10.1101/gr.214775.116 (2017).
240 Gutierrez-Triana, J. A. *et al.* Efficient single-copy HDR by 5' modified long dsDNA donors. *Elife* **7**, doi:10.7554/eLife.39468 (2018).
241 Yu, Y. *et al.* An efficient gene knock-in strategy using 5'-modified double-stranded DNA donors with short homology arms. *Nat Chem Biol* **16**, 387-390, doi:10.1038/s41589-019-0432-1 (2020).
242 Canaj, H. *et al.* Deep profiling reveals substantial heterogeneity of integration outcomes in CRISPR knock-in experiments. *bioRxiv*, 841098, doi:10.1101/841098 (2019).
243 Gutschner, T., Haemmerle, M., Genovese, G., Draetta, G. F. & Chin, L. Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair. *Cell Rep* **14**, 1555-1566, doi:10.1016/j.celrep.2016.01.019 (2016).
244 Yang, D. *et al.* Enrichment of G2/M cell cycle phase in human pluripotent stem cells enhances HDR-mediated gene repair with customizable endonucleases. *Sci Rep* **6**, 21264, doi:10.1038/srep21264 (2016).
245 Ling, X. *et al.* Improving the efficiency of precise genome editing with site-specific Cas9-oligonucleotide conjugates. *Sci Adv* **6**, eaaz0051, doi:10.1126/sciadv.aaz0051 (2020).
246 Cruz-Becerra, G. & Kadonaga, J. T. Enhancement of homology-directed repair with chromatin donor templates in cells. *Elife* **9**, doi:10.7554/eLife.55780 (2020).
247 van Erp, T. S., Cuesta-Lopez, S. & Peyrard, M. Bubbles and denaturation in DNA. *Eur Phys J E Soft Matter* **20**, 421-434, doi:10.1140/epje/i2006-10032-2 (2006).
248 Choi, C. H. *et al.* DNA dynamically directs its own transcription initiation. *Nucleic Acids Res* **32**, 1584-1590, doi:10.1093/nar/gkh335 (2004).
249 Traverso, J. J., Manoranjan, V. S., Bishop, A. R., Rasmussen, K. O. & Voulgarakis, N. K. Allostery through protein-induced DNA bubbles. *Sci Rep* **5**, 9037, doi:10.1038/srep09037 (2015).
250 Farge, G. *et al.* Protein sliding and DNA denaturation are essential for DNA organization by human mitochondrial transcription factor A. *Nat Commun* **3**, 1013, doi:10.1038/ncomms2001 (2012).
251 Branden, L. J., Mohamed, A. J. & Smith, C. I. A peptide nucleic acid-nuclear localization signal fusion that mediates nuclear transport of DNA. *Nat Biotechnol* **17**, 784-787, doi:10.1038/11726 (1999).
252 Leonetti, J. P., Mechti, N., Degols, G., Gagnor, C. & Lebleu, B. Intracellular distribution of microinjected antisense oligonucleotides. *Proc Natl Acad Sci U S A* **88**, 2702-2706, doi:10.1073/pnas.88.7.2702 (1991).
253 Skryabin, B. V. *et al.* Pervasive head-to-tail insertions of DNA templates mask desired CRISPR-Cas9-mediated genome editing events. *Sci Adv* **6**, eaax2941, doi:10.1126/sciadv.aax2941 (2020).

254 Stewart, M. Molecular mechanism of the nuclear protein import cycle. *Nat Rev Mol Cell Biol* **8**, 195-208, doi:10.1038/nrm2114 (2007).
255 Yang, B., Schwartz, M. & McJunkin, K. In vivo CRISPR screening for phenotypic targets of the mir-35-42 family in C. elegans. *Genes Dev* **34**, 1227-1238, doi:10.1101/gad.339333.120 (2020).
256 Nakamura, M., Gao, Y., Dominguez, A. A. & Qi, L. S. CRISPR technologies for precise epigenome editing. *Nat Cell Biol* **23**, 11-22, doi:10.1038/s41556-020-00620-7 (2021).
257 Amabile, A. *et al.* Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. *Cell* **167**, 219-232 e214, doi:10.1016/j.cell.2016.09.006 (2016).
258 Thakore, P. I. *et al.* Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat Methods* **12**, 1143-1149, doi:10.1038/nmeth.3630 (2015).
259 Black, J. B. *et al.* Targeted Epigenetic Remodeling of Endogenous Loci by CRISPR/Cas9-Based Transcriptional Activators Directly Converts Fibroblasts to Neuronal Cells. *Cell Stem Cell* **19**, 406-414, doi:10.1016/j.stem.2016.07.001 (2016).
260 Sun, L., Wu, J., Du, F., Chen, X. & Chen, Z. J. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* **339**, 786-791, doi:10.1126/science.1232458 (2013).
261 Motwani, M., Pesiridis, S. & Fitzgerald, K. A. DNA sensing by the cGAS-STING pathway in health and disease. *Nat Rev Genet* **20**, 657-674, doi:10.1038/s41576-019-0151-1 (2019).
262 Fernandes-Alnemri, T., Yu, J. W., Datta, P., Wu, J. & Alnemri, E. S. AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. *Nature* **458**, 509-513, doi:10.1038/nature07710 (2009).