Title

To CRISPR and beyond: the evolution of genome editing in stem cells

Journal

Regenerative Medicine, 11(8)

ISSN

1746-0751

Authors

Chen, Kuang-Yui

Knoepfler, Paul S

Publication Date

2016-12-01

DOI

10.2217/rme-2016-0107

Peer reviewed

Perspective

Regenerative Medicine

To CRISPR and beyond: the evolution of genome editing in stem cells

The goal of editing the genomes of stem cells to generate model organisms and cell lines for genetic and biological studies has been pursued for decades. There is also exciting potential for future clinical impact in humans. While recent, rapid advances in targeted nuclease technologies have led to unprecedented accessibility and ease of gene editing, biology has benefited from past directed gene modification via homologous recombination, gene traps and other transgenic methodologies. Here we review the history of genome editing in stem cells (including via zinc finger nucleases, transcription activator-like effector nucleases and CRISPR–Cas9), discuss recent developments leading to the implementation of stem cell gene therapies in clinical trials and consider the prospects for future advances in this rapidly evolving field.

Kuang-Yui Chen[1,2,3] & Paul S Knoepfler*[,1,2,3]

[1]Department of Cell Biology and Human Anatomy, University of California Davis School of Medicine, Davis, CA, USA

[2]Institute of Pediatric Regenerative Medicine, Shriners Hospital For Children Northern California, Sacramento, CA, USA

[3]Genome Center, University of California Davis, Davis, CA, USA

*Author for correspondence: knoepfler@ucdavis.edu

First draft submitted: 11 August 2016; Accepted for publication: 11 October 2016; Published online: 1 December 2016

Keywords: CRISPR-Cas9 • gene therapy • genome editing • stem cells • TALEN • ZFN

The ability to generate defined alterations in the genomes of stem cells has been an integral part of genetic and biological studies for more than three decades, but the technologies involved have changed substantially during this period. In addition the rate of technological change in genetic engineering is perceived to have greatly accelerated just in the last few years. A genome editing methodology that is state of the art today may be out of date within a few years. For instance, transcription activator-like effector nucleases (TALENs) technology was overtaken in just a few years after the cracking of its targeting code by the more facile and accessible clustered regularly interspaced short palindrome repeat (CRISPR)–CRISPR associated protein 9 (Cas9) technology. The relatively new CRISPR–Cas9 approach even as it is rapidly evolving and overtaking other methods to become the dominant genetic modification methodology today could be largely replaced by future, perhaps even unanticipated technologies.

The emergence of new methodological approaches has also changed the speed and accessibility of genome engineering as a research approach available to the scientific community. The use of CRISPR–Cas9 has proven capable of relatively quickly yielding discrete genomic changes in stem cells as small as point mutations termed 'gene edits' on a routine basis, an accomplishment not previously regularly attainable. By contrast, past approaches attempting to create such discrete modifications including defined 'knock-ins' were laborious, often left behind large alterations such as antibiotic selection cassettes that would require a second step for removal [1], and sometimes failed to work at all.

Thousands of laboratories today have access to and have ordered (for just a nominal fee in many cases at the nonprofit repository Addgene) the needed CRISPR–Cas9 materials such as plasmids for introducing genetic alterations into stem cells [2]. In parallel the

Future Medicine part of fsg

number of publications mentioning CRISPR–Cas9 in various forms has exhibited an almost exponential increase in the past few years. Together these developments are a clear indication that unlike in past years, now almost any laboratory can conduct gene targeting research, which is likely to speed the advancement of knowledge greatly.

In past years, making genetically modified mice even with relatively large, imprecise changes such as knockouts of particular genes depended on introduction of DNA via homologous recombination into mouse embryonic stem cells (mESCs) and subsequent breeding to obtain the correct genetic background to facilitate gene knockout/knockin (using recombinase technologies such as the Cre–Lox system). This was such a slow and expensive process that these considerations became impediments to the research and some individual laboratories could not participate in this kind of work, but this technology transformed mouse genetics nonetheless. Today making a knockout or transgenic mouse or other model organism via newer methods such as zinc finger nucleases (ZFNs), TALENs and CRISPR–Cas9 is a relatively much quicker process due to the ability of these engineered molecules to target specific genes and cut the DNA, stimulating DNA repair mechanisms resulting in gene disruption, insertion and mutation at a high rate. Both in the past using methods dependent on homologous recombination and now with CRISPR and other genetic modification approaches, the production of genetically engineered mice relies upon introducing gene targeting materials and in turn the desired genetic changes specifically first into mESCs and ultimately zygotes. The modified mESC were and still are then used via embryo manipulation to produce mutant embryos and ultimately mice. However, in principle CRISPR and other genetic engineering tools such as TALENs can be used in any kind of species of stem or germ cell, including human.

The increasing capability to produce genetic changes specifically in human stem and germ cells has stimulated a great deal of excitement both scientifically and in terms of potential translational applications to address clinically challenging genetic diseases, particularly those that are monogenic. At the same time this technological advancement of genetic engineering of human cells has sparked intense debate about emerging societal and bioethical issues. The use of genetic modification technology in human pluripotent stem and germ cells, particularly if this research ultimately could advance outside of the laboratory into a clinical or reproductive setting, is exciting but also highly controversial.

In this review, we provide a comprehensive overview of the evolution of genetic engineering technologies as they have been used in stem cells. We discuss the current state of the field, comparing the advantages and disadvantages of the most important technologies, and we discuss important potential future developments in this area.

Building a better mouse gene trap & other early mutation strategies

There are a large number of spontaneous mutant strains of mice and for a century or more a variety of different kinds have been collected and in some cases studied [3]. The realization that distinct phenotypes in mice were traceable to specific genotypes and definable alterations in particular genes opened up the possibility of researchers themselves making directed changes in genes to study mammalian developmental biology. A fairly common general approach early on was the use of broad-spectrum random mutation introduction into mice by radiation or mutagenesis with N-ethyl-N-nitrosourea. Following exposure of mice to N-ethyl-N-nitrosourea, either as mESCs or animals, the mice would be functionally screened for phenotypes, which could then in many cases be traced back to a defined causal mutation.

Groundbreaking, relatively more specific but still fairly imprecise approaches to the creation of useful genetically modified mice were also implemented. In the mid-1970s, Rudolf Jaenisch produced the first genetically modified or 'transgenic' mice by infecting embryos with viruses [4], but in the beginning such mice were effectively one-time experiments because the genetic change did not enter the germline. By 1980, Gordon and Ruddle reported production of mice with germline mutations that could pass such changes stably onto offspring in additional generations [5]. Such mutations were termed to have 'gone germline', meaning that the mutations had been stably introduced into the genomes of germ cells and hence were heritable. In another major milestone, mESC were derived [6,7], subsequently such mESC were targeted for mutations *in vitro*, and the mutant mESC could then be used to make genetically modified mice through incorporation into embryos, which were then implanted *in utero*. This new mESC-based genetic technology enabled production of mice on a large scale with any of a number of genetic changes allowing for a leap in scale for the developmental biology field. The modern field of mammalian genome modification was born via the collaboration of the stem cell and genetic approaches, leading to the 2007 Nobel Prize in Physiology or Medicine being awarded to Martin Evans, Mario Capecchi and Oliver Smithies.

In certain other early genetic applications including screening approaches, a gene-trapping strategy was uti-

lized. In gene trapping screens, a trapping cassette was introduced into mESC and randomly into genes (usually actively expressed genes), which were then identified by subsequent functional screening [3]. The gene trap was designed to 'trap' expression by, for instance, splicing in an artificial exon containing a poly A signal near the 5′-end of a gene so as to block normal, full-length RNA production. While some gene trapped alleles ended up being true nulls, others were only hypomorphic and still expressed some level of product.

Genetic changes mediated by homologous recombination

For most of the past three decades, the field of mouse genetics has produced specific mutant mice via a variety of methods that nonetheless in each case rely upon the process of homologous recombination. In a typical scenario, researchers would modify a specific genetic strain of mESC and then create chimeric mice most often by blastocyst injection of the targeted mESCs. The hope was that one or more of the resulting chimeric mice would have an introduced mutation that had gone germline so it could be transmitted to future generations in every cell in the body. These last steps have not changed greatly even up until today (some groups use other embryo-based approaches such as tetraploid complementation), but the methodology used at the crucial beginning step of introducing a specific genetic change into mESC has undergone substantial improvement.

In the early days of producing mutant mice the first step involved obtaining a genomic clone of your target gene of interest. This clone was then used to make a targeting vector that included a foreign sequence or sequences (e.g., an expression cassette such as *lacZ* or green fluorescent protein [*GFP*], or a selectable marker) flanked by homology 'arms' that via homologous recombination had the potential to integrate into the target locus and introduce a genetic change into the mESC, which would then subsequently be used for making the new mice. The targeting vector would also contain a selectable marker that conferred specific antibiotic resistance such as to neomycin. Along the way, there were many potential roadblocks and time sinks so that most often the process was slow and consumed a great deal of resources.

A common sticking point was the production of the genetically targeted mESC cells, which was typically an inefficient process, especially early on despite the robustness of mESC. The linearized targeting vector was generally introduced into the mESC via electroporation. In the 1980s starting with 1×10^6–5×10^8 electroporated mESCs, it was not unusual to only observe a few antibiotic resistant colonies after electroporation representing a poor initial efficiency rate, and only the subset of those with the best morphology had a reasonable chance of producing high-quality chimeras with targeted rather than random integration. In addition, before proceeding to the blastocyst injection step to try to generate chimeras, it was necessary to first validate that the mESC specifically had the desired genetic change. In most cases some candidate mutant mESC sublines exhibited abnormal banding patterns on Southern blots or other indications of integration of the targeting vector into the wrong locus or spurious events at the correct locus such as the targeting vector introducing both the desired mutation and other alterations too. In one of the earliest reports of a targeted gene disruption or knockout, just a single mESC colony made the research possible. Thomas and Capecchi described a striking brain phenotype with disruption of the *int-1* gene, but this mouse was produced from the one mESC clone that survived and passed validation out of a starting population of 5×10^8 cells that had initially survived electroporation, illustrating the challenges with efficiency [8]. The production of this mouse strain was such a groundbreaking scientific milestone that for first time it enabled the targeted manipulation of the mammalian genome. In the end, sometimes in attempts at gene targeting none of the candidate mESC colonies passed validation or even if they did, the mutation did not go germline in mice representing an experimental dead end at least at the organismal level. Fortunately, efficiencies have substantially improved in recent years [9].

Sometimes genetic changes were introduced into other species of cells such as rat or even human cells, but not embryonic stem cells (ESCs) of these species. Although, these kinds of efforts were relatively much less common than projects intending to generate mutant organisms such as mice, they produced some very useful cells and research. For instance, John Sedivy's laboratory reported production of rat fibroblasts via homologous recombination in 1997 that were null for c-Myc and these cells have been of great utility [10].

Despite the challenges, many research teams successfully created mutant strains of mice bearing defined modifications. The range of types of genetic changes was relatively limited, however. In many cases a replacement approach was utilized whereby a selectable marker cassette such as one coding for strong promoter-driven expression of the neomycin phosphotransferase or GFP gene was inserted in place of an essential part of a gene such as critical exon(s). This strategy often but not always resulted in a full null allele.

Conditional knockout technology based upon recombinases such as Cre or FLP later grew in use and allowed for the production of genetic modifica-

tions that were more elegant. The conditional genetic changes were not immediately intended to be disruptive of expression of the targeted gene initially, but could later be induced by expression of a particular recombinase either in mESC or in mice via breeding with mice bearing a recombinase transgene [11]. In the case of Cre-mediated conditional knockouts, LoxP sites (relatively small insertions) would first be introduced flanking an essential part of a gene in mESC (and then in mice) and in the presence of Cre the intervening sequence would be deleted often controllably in a spatial and temporal manner. This spatial and temporal control was mediated through the use of specific gene regulatory elements. For example, the use of Nestin–Cre allowed for a neural stem cell specific gene disruption, while the addition of estrogen receptor fusions to CRE protein or elements of the tetracycline-controlled gene expression system mediated temporal control. The production of conditional knockouts still relied mostly in the past upon the relatively inefficient early production step of homologous recombination based genetic alterations of mESC.

A need for new gene targeting approaches

Collectively for all these reasons, the traditional process of targeting vector-mediated homologous recombination-based gene targeting in mESC to make mutant mice was historically a relatively challenging and expensive, even if powerful and state-of-the-art technology of its time. In addition, the creation of combinatorial mutant stem cells and mice (often necessary due to redundancy and compensation in gene families) involved the lengthy process of repeating this procedure in sequential order and in many cases also conducting time-consuming breeding of mice to for instance generate double knockouts. This problematic *status quo* lasted for many years up until very recently. The creation of new technological innovations in genome modification discussed next, focusing on ZFNs, TALENs and CRISPR–Cas9, has opened the door to new avenues for genetics.

Zinc finger nucleases

Following the discovery of DNA repair and recombination mechanisms in bacteria and yeast that respond to DNA double-strand breaks (DSBs) [12,13], it was recognized that methods introducing precise DSBs could serve as an avenue for targeted genome modification. In addition to having the potential to disrupt genes via nonhomologous end joining (NHEJ) mediated nonsense deletions and insertions, repair of DSBs has been shown to stimulate homology directed repair (HDR) when a homologous DNA template is provided [14,15], resulting in potential gene insertion or precise substitution. One of the pioneering methods for targeted, site-specific modification of the genome is a class of engineered proteins called ZFNs. These proteins consist of zinc finger DNA binding domains fused to FokI, a restriction endonuclease that when brought together as a dimer will cleave DNA to generate a DSB in proximity to a specific bound sequence in the genome (Figure 1A). The endogenous DNA repair mechanisms in the cell responding to the DSB result in modification of the genome near the cleavage site.

ZFN technology was made possible by the first description of the DNA binding transcription factor TFIIIA in 1985 [16], which contains the Cys2–His2 zinc finger motif shown to be the most common DNA-binding motif in humans [17]. Coupled with the discovery and characterization of the FokI endonuclease [18–22], zinc finger–FokI fusion proteins were developed to form the first ZFNs in 1996 [23]. Contributing to the specificity with which ZFNs cleave DNA is the fact that FokI requires dimerization for cleavage activity and has weak native self-interaction [24,25]. Highly specific genome editing is thus facilitated by the requirement of two ZFNs binding nearby stretches of nucleotides to induce FokI dimerization (Figure 1A).

ZFNs have been utilized to modify genes in numerous organisms, including targeted gene disruption studies in *Drosophila* [26,27], zebrafish (*Danio rerio*) [28,29], rat [30,31] and mouse [30,32]. In human cells ZFNs have been used to target embryos, ESCs or somatic cells for heritable modification [33–36]. These include the addition of genes via plasmid delivery in human ESC and induced pluripotent stem (iPS) cells to insert a drug resistant marker [34] and generate new alleles of several genes [35]. Until recently, ZFNs were the only class of gene modification nucleases being used in clinical studies, initially for disrupting the chemokine CCR5 receptor that contributes to the HIV infection pathway (ClinicalTrials.gov identifier numbers NCT02388594, NCT00842634, NCT01044654, NCT01252641, NCT02225665, NCT02500849). As part of the initial exploration of targeting CCR5, ZFNs were employed to tag the *CCR5* locus in human ESC lines with GFP for CCR5 protein tracking [33]. In addition to studies targeting CCR5, other current clinical trials utilizing ZFNs include studies for treatment of cervical cancer (NCT02800369), hemophilia B (NCT02695160) and mucopolysaccharidoses (NCT02702115).

Despite numerous successes of ZFN mediated gene editing, several challenges prevent it from gaining widespread use in basic science and clinical studies. The difficult and nuanced design and validation of zinc finger proteins for specific nucleotide sequence binding remain perhaps the foremost roadblocks to broad adaptation of ZFN for gene targeting [37]. Several strategies

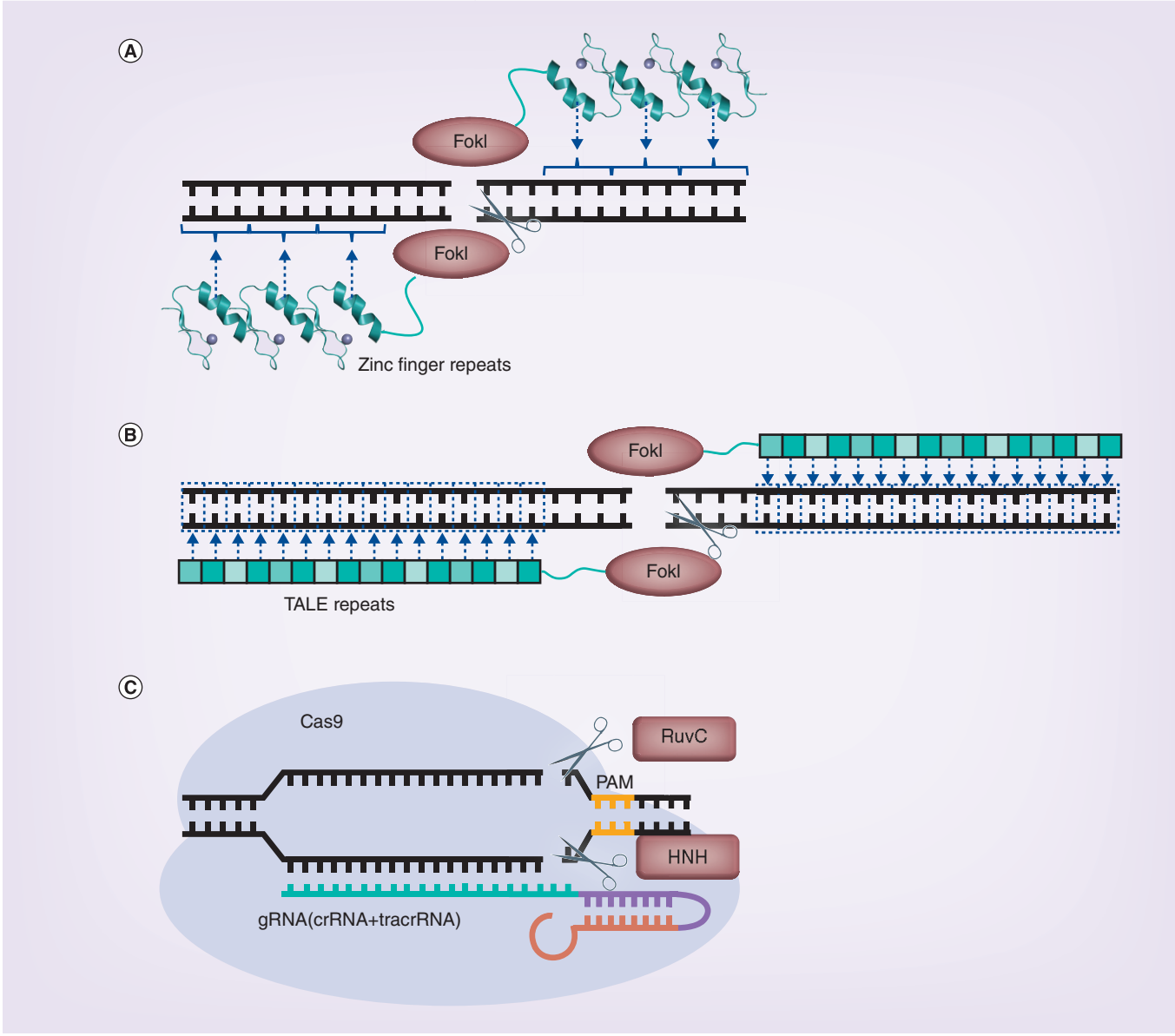

Figure 1. Targeted DNA nuclease technologies. (A) Zinc finger nucleases consist of zinc finger protein repeats (cyan helical structures) fused to FokI nuclease (magenta oval). Each zinc finger domain recognizes three nucleotides (dotted blue arrows) and can be linked to increase DNA binding specificity. FokI nuclease when dimerized induces a double-strand cut. Coordinated zinc ions are shown as gray spheres in this cartoon representation. **(B)** TALE nucleases incorporate TALE repeats (cyan squares) that have 1:1 recognition of DNA base pairs across both strands (dotted blue arrows and boxes). Typically, 16–22 TALEs are linked along with FokI nuclease (magenta oval) to induce a targeted double-strand cut. **(C)** CRISPR–Cas9 gene editing uses guide RNAs (gRNA) for recognition of the target DNA. Each gRNA consists of RNA (cyan) that is complementary to the DNA target and linked to segments mimicking bacterial CRISPR RNA (purple) and trans-activating CRISPR RNA (magenta). These features allow Cas9 (blue shading) to bind to the target DNA and gRNA duplex, provided that there is an 'NGG' protospacer adjacent motif site (orange) on the target DNA. Cas9 contains two nucleases domains (red labels) – RuvC-like domain and HNH domain – which each cut one strand of DNA.

for zinc finger protein engineering have been developed over the past decade, including phage selection [38], modular assembly [39] and computational-assisted design [40–42]. However, ZFNs containing *in vitro* validated zinc finger proteins often fail to induce genome modification in cells and *in vivo* [43]. Additionally, general concerns inherent to genome modification, such as off-target effects, cytotoxicity and editing efficacy present baseline challenges for ZFNs. While more accessible genome editing techniques such as TALENs and CRISPR–Cas9 have gained more traction over the years (Figure 2) studies utilizing ZFNs to re-engineer stem cells continue to make important contributions to development and disease studies [44–47].

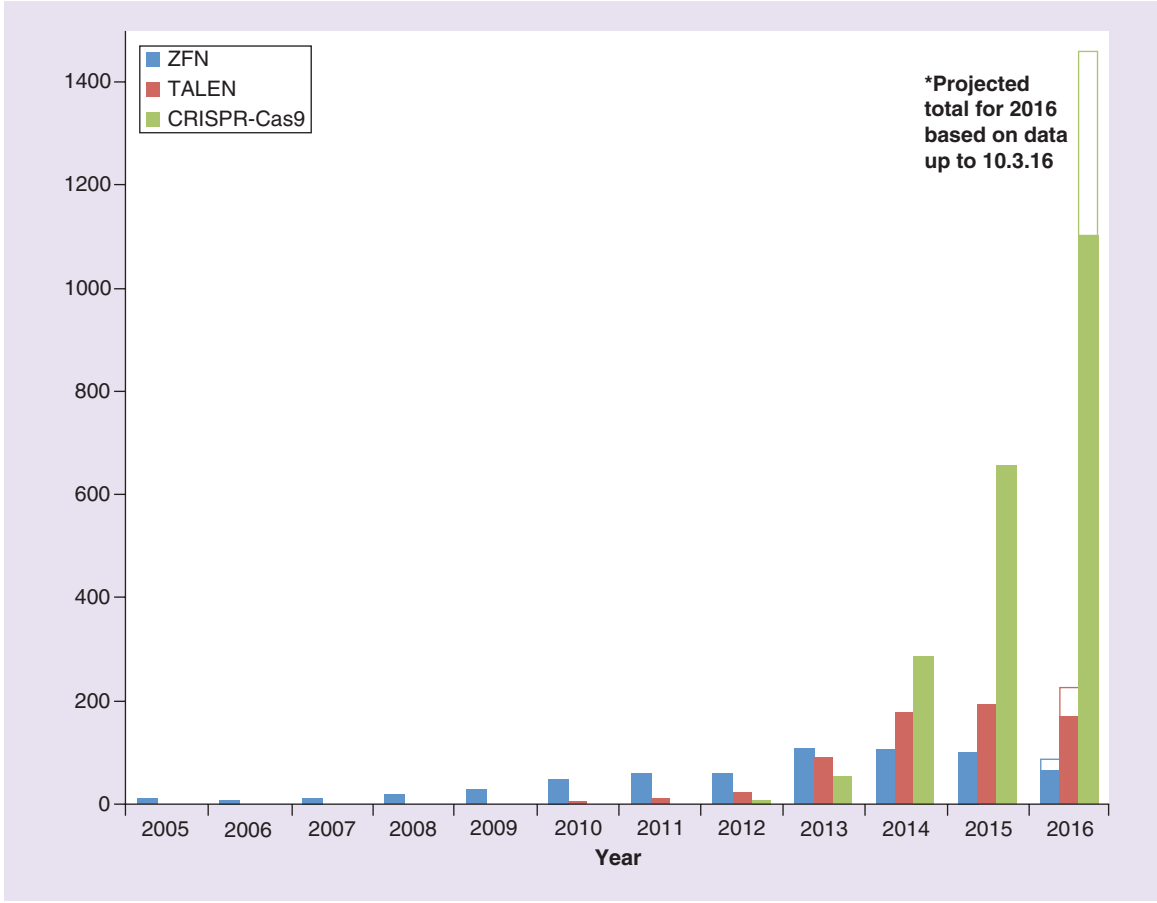

Figure 2. Publications per year for major gene editing nuclease methodologies. The number of publications per year relating to each of the major targeted nuclease methods plotted up to 2016. Projected publications for 2016 are calculated by assuming the same publication rate for each method for the remainder of 2016 (3 October–31 December) as the elapsed part of the year (1 January–3 October). Since this comparison is specific to the gene editing functions of these technologies, publications not mentioning nuclease activity or names of their associated nucleases in the Title/Abstract were excluded. The exact PubMed search parameters are as follows: ZFNs: zfn[Title/Abstract] OR zinc finger nuclease[Title/Abstract] OR zfp nuclease[Title/Abstract] OR zinc finger fokI[Title/Abstract]. TALENs: TAL effector nuclease[Title/Abstract] OR TALE nuclease[Title/Abstract] OR TALEN[Title/Abstract] OR transcription activator-like effector nuclease[Title/Abstract] NOT triplesalen[Title/Abstract]. CRISPR–Cas9: CRISPR Cas9[Title/Abstract] OR CRISPR associated Cas9[Title/Abstract].

Transcription activator-like effector nucleases

Transcription activator-like effectors (TALEs) are another family of DNA-binding proteins in this case isolated from the phytopathogenic bacteria of genus *Xanthomonas* whose DNA recognition code was deciphered and reported in 2009 [48,49]. Like zinc finger proteins, TALE proteins contain motifs that bind specific stretches of DNA to affect gene expression. However, TALEs uniquely contain highly conserved central regions of 33–35 amino acid tandem repeats that are only variable at two amino acid positions termed 'repeat variable diresidues,' which are responsible for nucleotide binding specificity [50]. The single base recognition of a TALE repeat for DNA binding contrasts the triplet base recognition of a zinc finger domain, providing designed TALEs more flexibility for binding a wider variety of nucleotide sequences (Figure 1) [51]. Additionally, because TALE repeats are largely invariable, designing TALEs requires less empirical screening than zinc finger proteins [51].

With the genome modification field primed by interest in ZFNs, it was not long before TALEs were fused with various nucleases [50,52] and transcriptional regulators [53,54], for the same intended gene targeting purposes as ZFNs. Building on the development and improvement of ZFNs, TALEs were fused to FokI for targeted double-strand cleavage of DNA to induce gene alterations via DNA repair mechanisms (Figure 1B). TALENs were rapidly adopted by the field for genome modifications, and within 5 years of deciphering the recognition code of TALEs, they had overtaken ZFNs

in popularity (based on number annual of publications, Figure 2). TALENs were used to engineer mouse and human stem cells for targeted developmental studies and disease modeling [52,55–60], on one-cell mouse embryos to generate genomic deletions for knockout studies [61], and for reprogramming of mESCs into iPS cells [62]. TALENs-mediated gene correction has been demonstrated in fibroblasts derived from epidermolysis bullosa patients [63]. Additionally, numerous studies have used TALENs on human pluripotent stem cells for development of targeted gene therapy [64–67].

TALENs have been shown to have encouragingly lower toxicity than ZFNs [68,69], and incorporation of engineered FokI obligate heterodimers [40,70] along with optimization of TALE repeat lengths have reduced off-target binding effects [69,71]. However, due to inherent barriers to protein design and synthesis and more unexpectedly the ascension of the CRISPR–Cas9 RNA-guided nuclease technology [72] (Figure 2), TALENs were not as widely adopted by the scientific community as initially predicted [51,68,73–74].

CRISPR–Cas9

Gene modification via engineered nucleases experienced a remarkable expansion in usage with the functional characterization of CRISPR and the adaptation of the Cas9 nuclease protein for targeted DNA cleavage beginning in 2012 [72]. Unlike ZFNs and TALENs that require engineering of protein domains to bind and cleave DNA for each project, the Cas9 system natively consists of the single common Cas9 protein containing two nuclease domains and is targeted to DNA via complementary RNA (Figure 1C). Once Cas9 was shown to be targeted to DNA by engineered single guide RNAs (gRNAs) [72], gene editing became highly accessible as generating gRNAs is decidedly simpler than designing and validating numerous DNA-binding proteins [73]. Publications on the continued development of CRISPR–Cas technology and applications of CRISPR–Cas9 in basic science and translational interrogations have dominated the gene editing field since 2014 (Figure 2).

Though it was only recently that CRISPR–Cas has been engineered and adapted for genome editing use, the discovery of CRISPRs dates back to 1987 when they were first found in *Escherichia coli* as short repeating sequences separated by unique spacers [75]. Along with the elucidation of *cas* genes encoding DNA binding and nuclease proteins [76–79], CRISPR–Cas was determined to be a type of immunological adaptive defense system that uses RNA signatures of past infections to target future invasive DNA for destruction [80–84]. Furthermore, CRISPR–Cas systems were initially classified into three types (I, II and III) with differing mechanisms of DNA binding and cleavage [85–87], and later reclassified into six types (I–VI) based on analysis of the gene composition and architecture of the CRISRP–Cas loci across numerous bacterial species [88–90]. Of these, Type II only requires a single Cas protein for DNA binding and cleavage [72,91]. The type II Cas9 from *Streptococcus pyogenes* was the first of these Cas proteins to be characterized as a dual RNA-guided endonuclease [72,87]. Bacterial CRISPR–Cas9 adaptive immunity was shown to require Cas9 to bind an RNA duplex consisting of a CRISPR RNA (crRNA) generated by insertion of invasive DNA segments into the CRISPR array and a trans-activating crRNA (tracrRNA) that is part of the array and functions to bind the repeating spacers of the crRNA [72,92]. Additionally, short 'NGG' motifs termed protospacer adjacent motifs (PAMs) on the invading DNA adjacent to the sequence targeted by the crRNA were determined to be essential for Cas9 function [93,94] (Figure 1C). A single gRNA consisting of a DNA recognition segment and a double strand-forming segment that binds Cas9 was engineered to mimic the crRNA:tracrRNA duplex [72] (Figure 1C). With this, gene modification could be achieved by simply delivering Cas9 and gRNA into cells via isolated protein/RNA, transfection of plasmids encoding these elements or viral transduction of said plasmids. Soon after, Cas9 was successfully deployed for genome editing of human cells [95–97]. A recent review of CRISPR Cas systems and functional mechanisms of Cas9 gene modification/regulation provides additional helpful background [98].

CRISPR–Cas9 was rapidly embraced by the scientific community and within a few years became the gene modifying technology of choice that continues to be further engineered for improvements in efficacy, specificity and expanded use for gene regulation in additional to gene editing. An example of Cas9 re-engineering stems from manipulation of its two nuclease domains, the HNH and RuvC-like domains. It was shown that the HNH domain cleaves the strand complementary to the crRNA or gRNA, and the RuvC-like domain cleaves the opposite strand, forming a DSB that stimulates DNA repair [72,91]. Mutational inactivation of either nuclease domain results in a Cas9 nickase that only cleaves a single strand, and inactivation of both nuclease domains results in a deactivated Cas9 (dCas9) incapable of cleaving DNA [72]. dCas9 has been utilized in various forms for many gene regulation studies, whether acting as transcriptional repressors through steric inhibition by dCas9-gRNA binding to targeted gene promoters (which has been dubbed CRISPRi) [73,99–101], or as gene activators when dCas9 is tethered to activating transcription factors [73,102–106].

In regard to CRISPR–Cas9 DNA targeting specificity, numerous studies have probed the off-target effects of Cas9 variants, and while off-target cleavage has been shown to occur in varying degrees [71,102,107], methods have been developed to predict and assess the level of off-target modification by designed gRNAs [108–110]. Researchers can predict gRNA binding by scanning genome databases of their target organism and probe for on-target and off-target editing events by sequencing these sites post-CRISPR'ing or they may undertake a more general unbiased whole-genome sequencing approach to screen for off-target effects [71,102,111]. One of the major challenges remaining with CRISPR–based gene targeting in stem cells is the production of clonal targeted lines that specifically only have the desired mutations. Additionally, efficient delivery of the targeting machinery to the desired cells remains a largely unaddressed hurdle.

Another challenge facing nuclease-mediated gene editing is that HDR has low rate of occurrence compared with NHEJ even when HDR donor templates are provided, which makes precise gene replacement or site-directed point mutations difficult to achieve with high efficiency [73,112–113]. Several studies have examined HDR template parameters such as oligonucleotide length, plasmid versus linear DNA and double-stranded versus single-stranded [114–116]. The general conclusions from these studies suggest that either long (>1000 bp) duplex DNA plasmid or short (~200 bp) single-stranded DNA (ssDNA) facilitated levels of HDR up to 20–35% [115–118]. Because it is thought that HDR and NHEJ repair mechanisms are in direct competition with one another [73], efforts have been made to shift the equilibrium in favor of HDR. One approach that has increased the rate of HDR is pharmacological disruption of NHEJ enzymes, such as DNA ligase IV [113,119]. However, adding pharmacological compounds to inhibit NHEJ may lead to unforeseen and undesirable cellular effects due to DNA repair deficiency, making optimization for gene therapy applications potentially even more difficult. A recent study characterizing in detail the dissociation of Cas9 from its target DNA found that Cas9 asymmetrically and consistently releases the 3′ end of the cleaved DNA that is not complementary to the gRNA prior to complete dissociation [112]. Following this observation, ssDNA donor templates were rationally designed to be complementary to the strand of cleaved target DNA that is released first and along with optimization of the donor length, yielded HDR rates in human cells up to 60% [112]. Compared with the previously more common homologous recombination-based approaches to gene targeting, optimized CRISPR–Cas9 targeting is generally much more efficient.

Yet another factor that constrains the design and coverage of gRNAs for Cas9 targeting is the aforementioned requirement of PAM sites for Cas9 cleavage and function. The most commonly used Cas9s are derived from *S. pyogenes* (SpCas9) and require 'NGG' PAM sites in the target sequence, around which gRNA can be designed. Several studies have explored Cas9 homologs from different species [120–125], including that of *S. aureus* (SaCas9), which is smaller than SpCas9, and more amenable for packaging into viral transduction vectors such as the adeno-associated virus commonly used for *in vivo* delivery of DNA [120]. SaCas9 was determined to function best with an 'NNGRRT' PAM motif (where 'R' denotes purine) [120]. Further characterization of Cas9 homologs and other type II Cas proteins have the potential to expand the range of PAMs and thereby the DNA regions targetable by CRISPR–Cas9 systems.

Despite these limitations, CRISPR–Cas9 has proven to be a versatile tool for gene regulation and genome editing, and has been employed in various stem cell applications. CRISPR–mediated gene disruption for functional study of developmental genes has been performed on mouse embryonic brain cells and neural stem cells [126]. Fertility and spermatogenesis studies have benefitted from CRISPR–Cas9 SNP interrogation in mouse spermatogonial stem cells [127–129]. Additionally, one-cell mouse and rat embryos have been injected with CRISPR–Cas9 components to generate organism-wide genome edits [61,130]. Human stem cells have also been a focus of CRISPR studies, resulting in methodologies facilitating engineering of gene knockouts in human ESC lines and numerous applications in human iPS cells for disease modeling and probing of gene function [47,58,131–133], including the transcriptional programming of neuronal differentiation in iPS cells [134]. Additionally, inducible genome editing in stem cells has been achieved using a combination of TALE-mediated gene targeting and drug-induced Cas9 expression (dubbed iCRISPR [135]), which has been successfully used to perform gene editing and regulation in human pluripotent stem cells [135], mESCs [136] and hESCs [137]. CRISPR–Cas9 gene therapy applications being explored in human stem cells include the repair of the *CFTR* gene in the intestinal stem cells of cystic fibrosis patients [138], improved hematopoietic differentiation of gene-corrected stem cells from β-thalassemia patients [139], interrogation of sequential gene mutations in colorectal cancer in human intestinal stem cells [140] and corrective mutation of the *FANCC* gene in Fanconi Anemia patient-derived fibroblasts [141] to name a few.

Current gene therapy studies in clinical trials

To date there are 12 clinical trials that make use of targeted nucleases for gene editing (Table 1). The majority

of these use ZFNs, but a clinical trial using TALENs (ClinicalTrials.gov identifier number NCT02808442) and one using CRISPR–Cas9 (NCT02793856) were received by the NIH in 2016. While clinical trial NCT02246491 started in 2014 incorporates the use of TALENs for gene editing of cells collected from ataxia–telangiectasia (A–T) patients, it does not use the TALENs-edited cells for any therapies, only for *ex vivo* interrogation of the disease. It is included in this review as an example of iPS cell gene editing. The larger number of clinical trials using ZFNs rather than TALENs or CRISPR–Cas9 is perhaps due to the rapid development of genome-modifying technologies that researchers and regulatory entities must keep abreast of and is more reflective of the delay getting newly developed methodologies from basic science research safely into the clinic. Additionally, a large number of these clinical trials are immunotherapies that employ targeted nucleases to modify genes in T cells to alter their susceptibility to autoimmune viruses (e.g., HIV) or enhance their recognition and binding to diseased cells (e.g., for cancer treatment, Table 1). One of the approaches to minimize risk of off-target gene editing events is to perform the gene editing and selection *ex vivo*, and infuse the selected cells back into patients. T cells are easy to extract, culture and reintroduce

Table 1. Summary of gene therapy clinical trials using targeted nucleases for gene editing.

	Application	Gene editing method	Clinical trial ID#	Year received/ completed
1	Autologous T cells genetically modified at the *CCR5* gene by zinc finger nucleases SB-728 for HIV (Zinc-Finger)	ZFN	NCT00842634	2009/2014
2	Phase I dose escalation study of autologous T cells genetically modified at the *CCR5* gene by zinc finger nucleases in HIV-infected patients	ZFN	NCT01044654	2010/2015
3	Dose escalation study of cyclophosphamide in HIV-infected subjects on HAART receiving SB-728-T	ZFN	NCT01543152	2012
4	Repeat doses of SB-728mR-T after cyclophosphamide conditioning in HIV-infected subjects on HAART	ZFN	NCT02225665	2014
5	A Phase I study of T cells genetically modified at the *CCR5* gene by zinc finger nucleases SB-728mR in HIV-infected patients	ZFN	NCT02388594	2015
6	**Safety study of zinc finger nuclease CCR5-modified hematopoietic stem/progenitor cells in HIV-1 infected patients**	**ZFN**	**NCT02500849**	**2015**
7	Ascending dose study of genome editing by the ZFP therapeutic SB-FIX in subjects with severe hemophilia B	ZFN	NCT02695160	2016
8	Study of molecular-targeted therapy using zinc finger nuclease in cervical precancerous lesions	ZFN	NCT02800369	2016
9	Ascending dose study of genome editing by the ZFP therapeutic SB-318 in subjects with MPS I	ZFN	NCT02702115	2016
10	**Cell-based approaches for modeling and treating ataxia–telangiectasia**	**TALEN**	**NCT02246491**	**2014**
11	Study of UCART19 in pediatric patients with relapsed/refractory B acute lymphoblastic leukemia (PALL)	TALEN	NCT02808442	2016
12	PD-1 knockout engineered T cells for metastatic non-small-cell lung cancer	CRISPR	NCT02793856	2016

Clinical trials utilizing targeted nuclease technologies are tabulated. The two trials incorporating gene editing of stem cells are shown in bold. Note that while TALENs are utilized downstream of clinical applications in NCT02246491 for mutation correction of ataxia telangiectasia mutated gene in ataxia–telangiectasia iPS cells from patients, the corrected cells are not being utilized for any form of patient therapy.
HIV: Human immunodeficiency virus; HAART: highly active antiretroviral therapy; PALL: Pediatric acute lymphoblastic leukemia.

into patients, making immunotherapies the first applications of gene editing in clinical trials.

However, in correlation with the aforementioned successful gene modifications of stem cells in numerous model organisms and in human iPS and ESCs, there are now two clinical trials underway that edit genes in stem cells (Table 1). One study is evaluating the use of CCR5-modified hematopoietic stem and progenitor cells as an alternative to modifying T cells for HIV treatment (NCT02500849). By targeting the precursor stem cells for modification via ZFNs, the study hopes to address patients that have suboptimal $CD4^+$ T cells, which makes T-cell *ex vivo* culture, editing and selection more difficult. The other study is the one mentioned earlier in this section, which is collecting blood and skin sample from patients with A–T, a neurodegenerative disease causing severe disability, in effort to explore cell-based modeling and treatment of A–T (NCT02246491). The researchers are generating iPS cells from the collected samples for gene correction of the A–T mutated gene (*ATM*), a kinase that is found to be mutated in all A–T patients. While the study does not have plans for using the stem cells for treatment of patients, the knowledge gained from probing the role of *ATM* mutation in iPS cells from A–T patients can potentially lead to improved therapeutics for disease treatment. With the intensification of gene editing interrogations of stem cells both *ex vivo* and in model organisms, applications of stem cell gene engineering for gene therapies are likely to expand.

Future directions: other nucleases & synthetic genomes

Additional novel and specific nucleases may be developed in the future as the basis for gene editing technologies. For instance recently, a DNA-guided nuclease from the Argonaute protein family was reported to be effective for gene editing applications, requiring only 5′-phosphorylated ssDNA of roughly 24 bases for specific targeting of DNA for cleavage [142]. Argonaute proteins have been studied since the late 1990s, and were characterized to play key roles in plant and *Drosophila* development [143–145] as well as regulation of RNA interference (RNAi) and microRNA (miRNA) pathways in numerous eukaryotic and prokaryotic organisms (reviewed in Hutvagner and Simard, 2008 [146]). An Argonaute from *Natronobacterium gregori* (NgAgo) was recently reported to function as a DNA-targeted DNA nuclease at physiological temperature that performed gene editing *in vitro* and in human cells with efficacies comparable to Cas9 [142]. While NgAgo and perhaps other as yet unreported DNA-guided Argonautes hold promise to become versatile targeted nucleases at disposal for synthetic biology and gene therapy, further validation and characterization by the wider scientific community is warranted and currently underway. In fact at the time of this review, extensive initial replication efforts of the findings reported for NgAgo have not validated its purported gene editing functions so far, reducing enthusiasm for the potential of NgAgo as a basis for gene targeting [147].

Another potential hypothetical future method for achieving specific gene editing outcomes in stem cells that has been discussed and that is more holistic is synthesizing whole genomes [148]. In this approach, rather than editing an existing genome, researchers could in a custom manner produce entire new genomes containing one or more genetic changes. This entirely hypothetical approach would be extremely powerful if realized, but faces serious obstacles including cost, the currently time-consuming nature of large-scale DNA synthesis, chromatinization of the resulting synthetic DNA and the need for introduction of the new genome into stem or germ cells.

Conclusion

Since the discovery of heritable traits followed by the discovery of DNA, humans have sought to decipher and manipulate the genetic code for understanding biology and treatment of disease. Studies of genes *in vivo* have evolved from being limited to examinations of exogenous expression by plasmid vectors inserted into cells to much more precise and efficient approaches. Gene manipulation in stem cells is of particular interest for generation of genetically-modified organisms, developmental studies and disease modeling. Taking advantage of homologous recombination, genomic modifications were made possible in stem cells of mice and other model organisms, albeit using laborious, time-consuming processes with low rates of success. Within the last decade the rapid development of targeted nucleases for gene editing has stimulated an exponential increase in gene modifications of somatic, stem and even germline cells with continued development of novel therapeutics and synthetic biology applications. New frontiers are already evident in this field including human germline gene editing, which are likely to continue to stimulate debate for decades to come.

Acknowledgements

The authors would like to thank K Lloyd and D Segal for critically reading and providing expert feedback on the manuscript.

Financial & competing interests disclosure

This work was support by NIH grant 1R01GM116919-01 to PS Knoepfler and by NCI training fellowship T32CA108459 to KY Chen. The authors have no other relevant affiliations or

financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Executive summary

Building a better mouse gene trap & other early mutation strategies

- Early efforts in the field of mouse gene editing were laborious and imprecise, relying on mutagenesis induced through radiation, chemical treatment and viral infection of mouse embryos.
- Improvements were made by creating heritable, germline mutations and by incorporating *in vitro* mutated mESCs and gene traps to generate gene-modified mouse lines.

Genetic changes mediated by homologous recombination

- Homologous recombination, while typically inefficient, was used to mediate genetic changes in mice and other species by introducing transgenes flanked by homology arms that allow for recombination with genomic DNA.
- Though inefficient, homologous recombination strategies yielded key studies on the disruption of specific genes in mice and other organisms.

Zinc finger nucleases

- Representing one of the first nucleases for site-specific gene editing, zinc finger nucleases (ZFNs) were heralded as a breakthrough in the field, allowing for gene disruption as well as gene insertion and precise mutation (via homology directed repair).
- While ZFNs have been and continue to be utilized in important studies probing gene function and gene correction in human diseases (e.g., HIV), the difficulty in designing and validating these proteins for DNA recognition and their modest success in cells and *in vivo* have prevented ZFNs from gaining widespread use.

Transcription activator-like effector nucleases

- Like zinc finger proteins, the transcription activator-like effector (TALE)-DNA binding code was deciphered, and subsequently TALEs were fused to nucleases (e.g., FokI) to form targeting nucleases for gene editing (TALENs).
- TALENs exhibited encouragingly lower toxicity than ZFNs, and optimization of TALE repeats along with FokI engineering have reduced off-target binding leading to rapid adoption of TALENs for genome modifications in embryos, embryonic stem cells and induced pluripotent stem cells.

CRISPR–Cas9

- Widely regarded as a breakthrough gene editing technology, the CRISPR-Cas9 nuclease system made targeted DNA cleavage and manipulation of transcriptional regulation relatively easy and cost effective.
- While technical limitations and challenges remain (motif requirements limiting targetable DNA regions, off-target binding, low rate of homology directed repair, etc.), CRISPR-Cas9 has rapidly become a versatile tool for gene editing and regulation studies across many model organisms and systems, including human stem cells.

Current gene therapy studies in clinical trials

- Twelve clinical trials up to now have made use of targeted nuclease technology for gene editing.
- While the majority of these trials are immunotherapies making use of gene editing to re-engineer T cells, two of these trials target stem cells for editing (NCT02500849 and NCT02246491).

Future directions: other nucleases & synthetic genomes

- Moving forward, the search for alternative nucleases for gene editing is ongoing with the hopes of yielding higher efficiencies with less toxicity and off-target binding.
- Additionally, the prospect of synthesizing whole genomes as a means to create model organisms for study is currently being discussed.

Conclusion

- In seeking to understand biology and disease through gene function, scientists have come a long way in probing and manipulating the genes of cells and organisms.
- Discovery and utilization of targeted nuclease technology has recently resulted in an exponential growth in the field of gene editing, and by extension, gene manipulation in stems cells holds promise to yield novel insights into biological processes as well as facilitate the development of therapeutics for human diseases.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest.

1 Valancius V, Smithies O. Testing an "in-out" targeting procedure for making subtle genomic modifications in mouse embryonic stem cells. *Mol. Cell Biol.* 11(3), 1402–1408 (1991).

2 CRISPR/Cas9 Guide. www.addgene.org/CRISPR/guide/

3 Soriano P. Gene targeting in ES cells. *Annu. Rev. Neurosci.* 18(1983), 1–18 (1995).

4 Jaenisch R. Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. *Proc. Natl Acad. Sci. USA* 73(4), 1260–1264 (1976).

5 Gordon JW, Scangos GA, Plotkin DJ, Barbosa JA, Ruddle FH. Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc. Natl Acad. Sci. USA* 77(12), 7380–7384 (1980).

•• **Breakthrough study on inducing heritable germline mutations in mice.**

6 Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292(5819), 154–156 (1981).

•• **Along with** [7]**, represents the first derivation of mouse embryonic stem cells in culture.**

7 Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl Acad. Sci. USA* 78(12), 7634–7638 (1981).

•• **Along with** [6]**, represents the first derivation of mouse embryonic stem cells in culture.**

8 Thomas KR, Capecchi MR. Targeted disruption of the murine int-1 proto-oncogene resulting in severe abnormalities in midbrain and cerebellar development. *Nature* 346, 847–850 (1990).

•• **A landmark paper on targeted gene disruption in mice, leading to the generation of a knockout mouse line from a single embryonic cell.**

9 Pettitt SJ, Liang Q, Rairdan XY *et al.* Agouti C57BL/6N embryonic stem cells for mouse genetic resources. *Nat. Methods* 6(7), 493–495 (2009).

10 Mateyak MK, Obaya AJ, Adachi S, Sedivy JM. Phenotypes of c-Myc-deficient rat fibroblasts isolated by targeted homologous recombination. *Cell Growth Differ.* 8(10), 1039–1048 (1997).

11 Kilby NJ, Snaith MR, Murray JAH. Site-specific recombinases: tools for genome engineering. *Trends Genet.* 9(12), 413–421 (1993).

12 Plessis A, Perrin A, Haber JE, Dujon B. Site-specific recombination determined by I-SceI, a mitochondrial group I intron-encoded endonuclease expressed in the yeast nucleus. *Genetics* 130(3), 451–460 (1992).

13 Rudin N, Sugarman E, Haber JE. Genetic and physical analysis of double-strand break repair and recombination in Saccharomyces cerevisiae. *Genetics* 122(3), 519–534 (1989).

14 Rouet P, Smih F, Jasin M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc. Natl Acad. Sci. USA* 91(13), 6064–6068 (1994).

•• **One of the pioneering studies on the stimulation of homologous recombination by double-strand breaks.**

15 Smih F, Rouet P, Romanienko PJ, Jasin M. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Res.* 23(24), 5012–5019 (1995).

16 Miller J, McLachlan AD, Klug A. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *J. Trace Elem. Exp. Med.* 14(2), 157–169 (2001).

17 Pabo CO, Peisach E, Grant RA. Design and selection of novel cys 2 his 2 zinc finger proteins. *Annu. Rev. Biochem.* 70(1), 313–340 (2001).

18 Hiroyuki S, Susumu K. New restriction endonucleases from *Flavobacterium okeanokoites* (FokI) and *Micrococcus luteus* (MluI). *Gene* 16(1-3), 73–78 (1981).

19 Szybalski W, Kim SC, Hasan N, Podhajska AJ. Class-IIS restriction enzymes – a review. *Gene* 100, 13–26 (1991).

20 Li L, Chandrasegaran S. Alteration of the cleavage distance of Fok I restriction endonuclease by insertion mutagenesis. *Proc. Natl Acad. Sci. USA* 90(7), 2764–2768 (1993).

21 Li L, Wu LP, Chandrasegaran S. Functional domains in Fok I restriction endonuclease. *Proc. Natl Acad. Sci. USA* 89(10), 4275–4279 (1992).

22 Li L, Wu LP, Clarke R, Chandrasegaran S. C-terminal deletion mutants of the Fok I restriction endonuclease. *Gene* 133(1), 79–84 (1993).

23 Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl Acad. Sci. USA* 93(3), 1156–1160 (1996).

• **One of the first reports of zinc finger protein-FokI nuclease fusions for targeted DNA cleavage.**

24 Vanamee ES, Santagata S, Aggarwal AK. FokI requires two specific DNA sites for cleavage. *J. Mol. Biol.* 309(1), 69–78 (2001).

25 Bitinaite J, Wah DA, Aggarwal AK, Schildkraut I. FokI dimerization is required for DNA cleavage. *Proc. Natl Acad. Sci. USA* 95(18), 10570–10575 (1998).

26 Bibikova M, Golic M, Golic KG, Carroll D. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics* 161(3), 1169–1175 (2002).

27 Bibikova M, Beumer K, Trautman JK, Carroll D. Enhancing gene targeting with designed zinc finger nucleases. *Science* 300(5620), 764 (2003).

28 Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat. Biotechnol.* 26(6), 695–701 (2008).

29 Doyon Y, McCammon JM, Miller JC *et al.* Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat. Biotechnol.* 26(6), 702–708 (2008).

30 Cui X, Ji D, Fisher DA, Wu Y, Briner DM, Weinstein EJ. Targeted integration in rat and mouse embryos with zinc-finger nucleases. *Nat. Biotechnol.* 29(1), 64–67 (2011).

31 Geurts AM, Cost GJ, Freyvert Y *et al.* Knockout rats via embryo microinjection of zinc-finger nucleases. *Science* 325(5939), 433 (2009).

32 Carbery ID, Ji D, Harrington A *et al.* Targeted genome modification in mice using zinc-finger nucleases. *Genetics* 186(2), 451–459 (2010).

33 Lombardo A, Genovese P, Beausejour CM *et al.* Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat. Biotechnol.* 25(11), 1298–1306 (2007).

34 Zou J, Maeder ML, Mali P *et al.* Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell Stem Cell* 5(1), 97–110 (2009).

35 Hockemeyer D, Soldner F, Beard C *et al.* Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat. Biotechnol.* 27(9), 851–857 (2009).

36 DeKelver RC, Choi VM, Moehle EA *et al.* Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Res.* 20(8), 1133–1142 (2010).

37 Perez-Pinera P, Ousterout DG, Gersbach CA. Advances in targeted genome editing. *Curr. Opin. Chem. Biol.* 16(3-4), 268–277 (2012).

38 Rebar EJ, Pabo CO. Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. *Science* 263(5147), 671–673 (1994).

39 Hye JK, Lee HJ, Kim H, Cho SW, Kim JS. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res.* 19(7), 1279–1288 (2009).

40 Szczepek M, Brondani V, Büchel J, Serrano L, Segal DJ, Cathomen T. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat. Biotechnol.* 25(7), 786–793 (2007).

41 Ashworth J, Havranek JJ, Duarte CM *et al.* Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature* 441(7093), 656–659 (2006).

42 Sander JD, Zaback P, Joung JK, Voytas DF, Dobbs D. Zinc finger targeter (ZiFiT): an engineered zinc finger/target site design tool. *Nucleic Acids Res.* 35(Suppl. 2), W599–W605 (2007).

43 Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* 11(9), 636–646 (2010).

44 Cai Y, Laustsen A, Zhou Y *et al.* Targeted, homology-driven gene insertion in stem cells by ZFN-loaded "all-in-one" lentiviral vectors. *Elife* 5, 1–15 (2016).

45 Booth C, Gaspar HB, Thrasher AJ. Treating immunodeficiency through HSC gene therapy. *Trends Mol. Med.* 22(4), 317–327 (2016).

46 Ott de Bruin LM, Volpi S, Musunuru K. Novel genome-editing tools to model and correct primary immunodeficiencies. *Front. Immunol.* 6, 250 (2015).

47 Hockemeyer D, Jaenisch R. Induced pluripotent stem cells meet genome editing. *Cell Stem Cell* 18(5), 573–586 (2016).

48 Boch J, Scholze H, Schornack S *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326(5959), 1509–1512 (2009).

• **Along with Ref. [49], represents the first deciphering of transcription activator-like effector (TALE)-DNA recognition code.**

49 Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science* 326(5959), 1501 (2009).

• **Along with [48], represents the first deciphering of TALE-DNA recognition code.**

50 Christian M, Cermak T, Doyle EL *et al.* Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186(2), 756–761 (2010).

51 Gaj T, Gersbach CA, Barbas CF. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* 31(7), 397–405 (2013).

52 Hockemeyer D, Wang H, Kiani S *et al.* Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* 29(8), 731–734 (2011).

•• **A key contribution to gene editing of pluripotent cells, demonstrating targeted genetic manipulation of human embryonic stem cells and induced pluripotent stem cells with TALENs.**

53 Miller JC, Tan S, Qiao G *et al.* A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* 29(2), 143–148 (2011).

54 Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol.* 29(2), 149–154 (2011).

55 Ru R, Yao Y, Yu S *et al.* Targeted genome engineering in human induced pluripotent stem cells by penetrating TALENs. *Cell Regen. (Lond.)* 1, 1–8 (2013).

56 Ding Q, Lee YK, Schaefer EAK *et al.* A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell* 12(2), 238–251 (2013).

57 Xia G, Gao Y, Jin S *et al.* Genome modification leads to phenotype reversal in human myotonic dystrophy type 1 induced pluripotent stem cell-derived neural stem cells. *Stem Cells.* 33(6), 1829–1838 (2015).

58 Martinez RA, Stein JL, Krostag ARF *et al.* Genome engineering of isogenic human ES cells to model autism disorders. *Nucleic Acids Res.* 43(10), e65 (2015).

59 Pu J, Frescas D, Zhang B, Feng J. Utilization of TALEN and CRISPR/Cas9 technologies for gene targeting and modification. *Exp. Biol. Med.* 240(8), 1065–1070 (2015).

60 Buechele C, Breese EH, Schneidawind D *et al.* MLL leukemia induction by genome editing of human $CD34^+$ hematopoietic cells. *Blood* 126(14), 1683–1694 (2015).

61 Brandl C, Ortiz O, Röttig B, Wefers B, Wurst W, Kühn R. Creation of targeted genomic deletions using TALEN or CRISPR/Cas nuclease pairs in one-cell mouse embryos. *FEBS Open Bio* 5(1), 26–35 (2015).

62 Gao X, Yang J, Tsang JCH, Ooi J, Wu D, Liu P. Reprogramming to pluripotency using designer TALE transcription factors targeting enhancers. *Stem Cell Rep.* 1(2), 183–197 (2013).

63 Osborn MJ, Starker CG, McElroy AN *et al.* TALEN-based gene correction for epidermolysis bullosa. *Mol. Ther.* 21(6), 1151–1159 (2013).

64 Menger L, Gouble A, Marzolini MAV *et al.* TALEN-mediated genetic inactivation of the glucocorticoid receptor in cytomegalovirus-specific T cells. *Blood* 126(26), 2781–2790 (2015).

65 Benjamin R, Berges BK, Solis A, Igbinedion O, Strong CL, Schiller MR. *TALEN* gene editing takes aim on HIV. *Hum. Genet.* 135(9), 1059–1070 (2016).

66 Niu J, Zhang B, Chen H. Applications of TALENs and CRISPR/Cas9 in human cells and their potentials for gene therapy. *Mol. Biotechnol.* 56(8), 681–688 (2014).

67 Biffi A. Clinical translation of TALENS: treating *SCID-X1* by gene editing in iPSCs. *Cell Stem Cell* 16(4), 348–349 (2015).

68 Mussolino C, Cathomen T. TALE nucleases: tailored genome engineering made easy. *Curr. Opin. Biotechnol.* 23(5), 644–650 (2012).

69 Mussolino C, Alzubi J, Fine EJ *et al.* TALENs facilitate targeted genome editing in human cells with high specificity and low cytotoxicity. *Nucleic Acids Res.* 42(10), 6762–6773 (2014).

70 Miller JC, Holmes MC, Wang J *et al.* An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.* 25(7), 778–785 (2007).

71 Yee J-K. Off-target effects of engineered nucleases. *FEBS J.* 283(17), 3239–3248 (2016).

72 Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096), 816–821 (2012).

•• **A landmark CRISPR–Cas9 gene editing paper, demonstrating targeted, site-specific Cas9 nuclease activity when directed by an engineered single RNA guide.**

73 Doudna JA, Charpentier E. The new frontier of genome engineering with CRISPR–Cas9. *Science* 346(6213), 1258096–1258096 (2014).

74 Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.* 14(1), 49–55 (2013).

75 Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* 169(12), 5429–5433 (1987).

76 Bolotin A, Quinquis B, Sorokin A, Dusko Ehrlich S. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151(8), 2551–2561 (2005).

77 Pourcel C, Salvignol G, Vergnaud G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* 151(3), 653–663 (2005).

78 Jansen R, Embden JD, Gaastra W, Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 43(6), 1565–1575 (2002).

79 Haft DH, Selengut J, Mongodin EF, Nelson KE. A guild of 45 CRISPR–associated (Cas) protein families and multiple CRISPR/cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* 1(6), 0474–0483 (2005).

80 Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. 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(1), 7 (2006).

81 Barrangou R, Fremaux C, Deveau H *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315(5819), 1709–1712 (2007).

82 Barrangou R, Marraffini LA. CRISPR–cas systems: prokaryotes upgrade to adaptive immunity. *Mol. Cell* 54(2), 234–244 (2014).

83 van der Oost J, Westra ER, Jackson RN, Wiedenheft B. Unravelling the structural and mechanistic basis of CRISPR–Cas systems. *Nat. Rev. Microbiol.* 12(7), 479–492 (2014).

84 Bondy-Denomy J, Davidson AR. To acquire or resist: the complex biological effects of CRISPR–Cas systems. *Trends Microbiol.* 22(4), 218–225 (2014).

85 Makarova KS, Aravind L, Wolf YI, Koonin EV. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR–Cas systems. *Biol. Direct.* 6(1), 38 (2011).

86 Makarova KS, Haft DH, Barrangou R *et al.* Evolution and classification of the CRISPR–Cas systems. *Nat. Rev. Microbiol.* 9(6), 467–477 (2011).

87 Chylinski K, Le Rhun A, Charpentier E. The tracrRNA and Cas9 families of type II CRISPR–Cas immunity systems. *RNA Biol.* 10(5), 726–737 (2013).

88 Shmakov S, Abudayyeh OO, Makarova KS *et al.* Discovery and functional characterization of diverse class 2 CRISPR–Cas systems. *Mol. Cell* 60(3), 385–397 (2015).

89 Makarova KS, Wolf YI, Alkhnbashi OS *et al.* An updated evolutionary classification of CRISPR–Cas systems. *Nat. Rev. Microbiol.* 13(11), 722–736 (2015).

90 Mohanraju P, Makarova KS, Zetsche B, Zhang F, Koonin EV, van der Oost J. Diverse evolutionary roots and mechanistic variations of the CRISPR–Cas systems. *Science* 353(6299), aad5147 (2016).

91 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. USA* 109(39), E2579–E2586 (2012).

92 Deltcheva E, Chylinski K, Sharma CM *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471(7340), 602–607 (2011).

93 Mojica FJM, Díez-Villaseñor C, García-Martínez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60(2), 174–182 (2005).

94 Deveau H, Barrangou R, Garneau JE *et al.* Phage response to CRISPR–encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190(4), 1390–1400 (2008).

95 Cong L, Ran FA, Cox D *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121), 819–823 (2013).

• **Along with** [96,97], **comprises the first studies describing genome engineering of human cells with Cas9.**

96 Mali P, Yang L, Esvelt KM *et al.* RNA-guided human genome engineering via Cas9. *Science* 339(6121), 823–826 (2013).

• **Along with** [95,97], **comprises the first studies describing genome engineering of human cells with Cas9.**

97 Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. *eLife* 2(2), e00471 (2013).

• **Along with** [95,96], **comprises the first studies describing genome engineering of human cells with Cas9.**

98 Wright AV, Nuñez JK, Doudna JA. Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* 164(1-2), 29–44 (2016).

99 Dominguez AA, Lim WA, Qi LS. Beyond editing: repurposing CRISPR–Cas9 for precision genome regulation and interrogation. *Nat. Rev. Mol. Cell Biol.* 17(1), 5–15 (2015).

100 Bassett AR, Liu J-L. CRISPR/Cas9 and genome editing in *Drosophila*. *J. Genet. Genomics* 41(1), 7–19 (2014).

101 Ma D, Liu F. Genome editing and its applications in model organisms. *Genomics Proteomics Bioinform.* 13(6), 336–344 (2015).

102 Wu X, Scott DA, Kriz AJ *et al.* Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.* 32(7), 670–676 (2014).

103 Nuñez JK, Harrington LB, Doudna JA. Chemical and biophysical modulation of Cas9 for tunable genome engineering. *ACS Chem. Biol.* 11(3), 681–688 (2016).

104 Zhang D, Li Z, Li J-F. Targeted gene manipulation in plants using the CRISPR/Cas technology. *J. Genet. Genomics* 43(5), 251–262 (2016).

105 Singh P, Schimenti JC, Bolcun-Filas E. A mouse geneticist's practical guide to CRISPR applications. *Genetics* 199(1), 1–15 (2015).

106 Shalem O, Sanjana NE, Zhang F. High-throughput functional genomics using CRISPR–Cas9. *Nat. Rev. Genet.* 16(5), 299–311 (2015).

107 Hendel A, Fine EJ, Bao G, Porteus MH. Quantifying on- and off-target genome editing. *Trends Biotechnol.* 33(2), 132–140 (2015).

108 Doench JG, Fusi N, Sullender M *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR–Cas9. *Nat. Biotechnol.* 34(2), 184–191 (2016).

109 Port F, Chen H-M, Lee T, Bullock SL. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in *Drosophila*. *Proc. Natl Acad. Sci. USA* 111(29), E2967–E2976 (2014).

110 Naito Y, Hino K, Bono H, Ui-Tei K. CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics* 31(7), 1120–1123 (2015).

111 Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. *Science* 351(6268), 84–88 (2016).

112 Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR–Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* 34(3), 339–344 (2016).

113 Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. Increasing the efficiency of precise genome editing with CRISPR–Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* 33(5), 538–542 (2015).

114 Ran FA, Hsu PPD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR–Cas9 system. *Nat. Protoc.* 8(11), 2281–2308 (2013).

115 Chen F, Pruett-Miller SM, Huang Y *et al.* High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat. Methods* 8(9), 753–755 (2011).

116 Davis L, Maizels N. Homology-directed repair of DNA nicks via pathways distinct from canonical double-strand break repair. *Proc. Natl Acad. Sci. USA* 111(10), E924–E932 (2014).

117 Ding Q, Regan S, Xia Y, Oostrom L, Cowan C, Musunuru K. Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem* 12(4), 393–394 (2013).

118 Yang L, Guell M, Byrne S *et al.* Optimization of scarless human stem cell genome editing. *Nucleic Acids Res.* 41(19), 9049–9061 (2013).

119 Chu VT, Weber T, Wefers B *et al.* Increasing the efficiency of homology-directed repair for CRISPR–Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* 33(5), 543–548 (2015).

120 Ran FA, Cong L, Yan WX *et al. In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature* 520(7546), 186–190 (2015).

121 Nishimasu H, Cong L, Yan WX *et al.* Crystal structure of *Staphylococcus aureus* Cas9. *Cell* 162(5), 1113–1126 (2015).

122 Hirano H, Gootenberg JS, Horii T *et al.* Structure and engineering of *Francisella novicida* Cas9. *Cell* 164(5), 950–961 (2016).

123 Hou Z, Zhang Y, Propson NE *et al.* Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc. Natl Acad. Sci. USA* 110(39), 15644–15649 (2013).

124 Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods* 10(11), 1116–1121 (2013).

125 Braff JL, Yaung SJ, Esvelt KM, Church GM. Characterization of Cas9–Guide RNA orthologs: TABLE 1. *Cold Spring Harb. Protoc.* 2016 (5), pdb.top086793 (2016).

126 Kalebic N, Taverna E, Tavano S *et al.* CRISPR/Cas9-induced disruption of gene expression in mouse embryonic brain and single neural stem cells *in vivo*. *EMBO Rep.* 17(3), 338–348 (2016).

127 Chapman KM, Medrano GA, Jaichander P *et al.* Targeted germline modifications in rats using CRISPR/Cas9 and spermatogonial stem cells. *Cell Rep.* 10(11), 1828–1835 (2015).

128 Sato T, Sakuma T, Yokonishi T *et al.* Genome editing in mouse spermatogonial stem cell lines using TALEN and double-nicking CRISPR/Cas9. *Stem Cell Rep.* 5(1), 75–82 (2015).

129 Singh P, Schimenti JC. The genetics of human infertility by functional interrogation of SNPs in mice. *Proc. Natl Acad. Sci. USA* 112(33), 10431–10436 (2015).

130 Shao Y, Guan Y, Wang L *et al.* CRISPR/Cas-mediated genome editing in the rat via direct injection of one-cell embryos. *Nat. Protoc.* 9(10), 2493–2512 (2014).

131 Chen Y, Cao J, Xiong M *et al.* Engineering human stem cell lines with inducible gene knockout using CRISPR/Cas9. *Cell Stem Cell* 17(2), 233–244 (2015).

132 Shetty DK, Inamdar MS. Generation of a heterozygous knockout human embryonic stem cell line for the OCIAD1 locus using CRISPR/CAS9 mediated targeting: BJNhem20-OCIAD1-CRISPR–20. *Stem Cell Res.* 16(2), 207–209 (2016).

133 Fontes A, Lakshmipathy U. Advances in genetic modification of pluripotent stem cells. *Biotechnol. Adv.* 31(7), 994–1001 (2013).

134 Chavez A, Scheiman J, Vora S *et al.* Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* 12(4), 326–328 (2015).

135 González F, Zhu Z, Shi ZD *et al.* An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. *Cell Stem Cell* 15(2), 215–226 (2014).

136 Dow LE, Fisher J, O'Rourke KP *et al.* Inducible *in vivo* genome editing with CRISPR–Cas9. *Nat. Biotechnol.* 33(4), 390–394 (2015).

137 Zhu Z, Verma N, González F, Shi Z-D, Huangfu D. A CRISPR/Cas-mediated selection-free knockin strategy in human embryonic stem cells. *Stem Cell Rep.* 4(6), 1103–1111 (2015).

138 Schwank G, Koo BK, Sasselli V *et al.* Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* 13(6), 653–658 (2013).

139 Song B, Fan Y, He W *et al.* Improved hematopoietic differentiation efficiency of gene-corrected beta-thalassemia induced pluripotent stem cells by CRISPR/Cas9 system. *Stem Cells Dev.* 24(9), 1053–1065 (2015).

140 Drost J, van Jaarsveld RH, Ponsioen B *et al.* Sequential cancer mutations in cultured human intestinal stem cells. *Nature* 521(7550), 43–47 (2015).

141 Osborn MJ, Gabriel R, Webber BR *et al.* Fanconi anemia gene editing by the CRISPR/Cas9 system. *Hum. Gene Ther.* 26(2), 114–126 (2015).

142 Gao F, Shen XZ, Jiang F, Wu Y, Han C. DNA-guided genome editing using the Natronobacterium gregoryi Argonaute. *Nat. Biotechnol.* 34(7), 768–773 (2016).

143 Moussian B, Schoof H, Haecker A, Ju G, Laux T. Role of the *ZWILLE* gene in the regulation of central shoot meristem cell fate during *Arabidopsis embryogenesis*. *Curr. Opin. Plant Biol.* 1(3), 188 (1998).

144 Bohmert K, Camus I, Bellini C, Bouchez D, Caboche M, Banning C. *AGO1* defines a novel locus of Arabidopsis controlling leaf development. *EMBO J.* 17(1), 170–180 (1998).

145 Lin H, Spradling AC. A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the *Drosophila* ovary. *Development* 124(12), 2463–2476 (1997).

146 Hutvagner G, Simard MJ. Argonaute proteins: key players in RNA silencing. *Nat. Rev. Mol. Cell Biol.* 9(1), 22–32 (2008).

147 Cyranoski D. Replications, ridicule and a recluse: the controversy over NgAgo gene-editing intensifies. *Nature* 536(7615), 136–137 (2016).

148 Swetlitz I. Top scientists hold closed meeting to discuss building a human genome from scratch. www.statnews.com