Charles University
Faculty of Science

Special Chemical and Biological Programmes
Molecular Biology and Biochemistry of Organisms
Department of Genetics and Microbiology

Jan Eliáš

CRISPR/Cas9-based genome editing in mice: state of the art and future perspectives

Editace myšího genomu pomocí technologie CRISPR/Cas9: současný stav poznání a možnosti budoucího vývoje

Bachelor´s thesis

Supervisor: Mgr. Petr Kašpárek, Ph.D
Consultant: Mgr. Radmila Hanečková

Prague, 2018

Poděkování : Především bych chtěl poděkovat svému školiteli Petru Kašpárkovi za jeho vedení, cenné rady a trpělivost s vypracováváním mojí práce. Dále bych rád poděkoval své konzultantce Radmile Hanečkové za cenné tipy ohledně formulací a anglického jazyka. A v neposlední řadě bych také chtěl poděkovat své rodině a kamarádům za jejich neutuchající podporu.

Prohlášení: Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 10. 5. 2018 Jan Eliáš

Abstrakt

Mutantní myši jsou klíčové pro odhalení funkce genů *in vivo*. V posledních letech prošla jejich příprava revolucí díky rychlému rozvoji programovatelných nukleáz, především systému CRISPR/Cas9. Editace genomu založená na vnesení komponentů systému CRISPR/Cas9 do časných vývojových stádií myších embryí umožnila rychlou a levnou přípravu genově deficientních myších modelů, hlavně ve srovnání s tradičními metodami, založenými na modifikacích embryonálních kmenových buněk (ESCs). Schopnost systému CRISPR/Cas9 indukovat na daném místě v genomické DNA dvojvláknový zlom (DSB) umožňuje efektivní narušení správného fungování daného genu díky náhodným mutacím dané sekvence, nebo kompletním vyštěpením daného genu. Přesné modifikace, jako inkorporace fragmentu DNA do daného lokusu, nicméně pořád zůstávají obtížně proveditelné. V této práci shrnuji systém CRISPR/Cas9, jeho použití v produkci mutantních myší a jeho možné modifikace, které by vedly ke zvýšení účinnosti přesných modifikací.

Klíčová slova:

CRISPR/Cas9, myš, transgeneze, homologní rekombinace

Abstract

Mutant mice are crucial tools for understanding gene functions *in vivo*. Recently, generation of mouse mutants was revolutionized by rapid developement of programmable nucleases, predominantly by the CRISPR/Cas9 system. Genome editing based on introduction of CRISPR/Cas9 components into early stage mouse embyros allows fast and inexpensive generation of gene-deficient animal models, especially when compared to the traditional techniques based on modification of embryonic stem cells (ESCs). The ability of CRISPR/Cas9 to induce double-strand break (DSB) at a given location of genomic DNA enables effective gene-ablation by random modification of the coding sequences or by complete ablation of the gene. However, precise modification of the gene sequences, such as incorporation of a DNA fragment into specific loci, are still difficult to make. In this work, I present a review of CRISPR/Cas9 system, its use in production of mutant mice and possible modifications of the system to increase the efficiency of precise gene-targeting.

Keywords:

CRISPR/Cas9, mouse, transgenesis, homologous recombination

Glossary

Alt-NHEJ – alternative-NHEJ

ABE – adenine base editor

BE – base editor

BIR – break-induced repair

Cas – CRISPR-associated

CRISPR – clustered regularly interspaced short palindromic repeats

crRNA – CRISPR RNA

DSB – double-strand break

DSBR – double-strand break repair

dsDNA – double-strand DNA

ESC – embryonic stem cell

gRNA – guide RNA

HDR – homology-directed repair

HR – homologous recombination

LDI – low dose of gamma ray irradiation

MMEJ – microhomology-mediated end joining

NHEJ – non-homologous end joining

PAM – protospacer-adjacent motif

PN – programmable nuclease

PS – phosphorothioate

RGEN – RNA-guided endonuclease

RNP – ribonucleoprotein

SDSA – synthesis-dependent strand annealing

sgRNA – single guide RNA

shRNA – short-hairpin RNA

SSA – single-strand annealing

ssDNA – single-strand DNA

TALE – transcription activator-like effector

TALEN – transcription activator-like effector nuclease

tracrRNA – trans-activating CRISPR RNA

UDG – Uracil DNA glycosylase

UGI – Uracil glycosylase inhibitor

ZFN – zinc-finger nuclease

Table of contents

Introduction

Mouse models of various genetic backgrounds proved to be one of the most important model animals, for example for its physiological similarity to humans, quick maturation to reproduction age and ease of breeding, cost effectivness in their maintenance and more[1]. This has led them to be the model animal of choice for studying various areas of interest in a complex environment that can´t be achieved *in vitro*. Those areas include studying diseases, function of genes, testing novel medication and more by creating a model with specific genetic modification relevant for the area of interest[1–3].

Until recently, production of mutant mice was accomplished by usage of modified ESCs. In this approach, ESCs are modified via homologous recombination (HR), selected through an antibiotic selection for positive integrations and then injected into embryos in blastocyst stage. Those embryos are then delivered into the uterus of a pseudopregnant mouse. However, this approach is far from perfect, due to various disadvantages like tedious selection of modified cells, low efficiency of integrations and finally, the uncertainty of successful transfer into cells of the germ line[4–8].

The most promising approach that is gradually replacing the old method of mutant mice production is the use of programmable nucleases (PNs). This is because of their ability to induce a DSB at a locus of interest, which vastly increases the amount of HR events.[9] Various PNs exist and each of them has its advantages as well as disadvantages.

The first PNs to be used in genome editing were the zinc-finger nucleases (ZFNs). These nucleases, which saw their prime in genome editing in 2000s, were the result of a merge between designed arrays of zinc-finger DNA-binding domains and a nuclease domain of type IIs restriction endonuclease, the FokI[10]. Although they proved to be effective, prediction of on-target binding specificity of designed arrays, given by the complexity of DNA-binding domains of respective zinc fingers and their interactions, is still problematic and time-consuming compared to more recent approaches[11].

More promising approach was described in 2010, in which the authors used transcription activator-like effector (TALE) DNA-binding domain from plant pathogen of *Xanthomonas* genus and merged it with FokI nuclease domain[12]. This gave birth to the transcription activator-like effector nucleases (TALENs). Their main advantage over ZFNs is far more easier design of the DNA-binding domain, which made them the PN of choice for the time being[13].

However in 2012 and 2013, the most promising approach yet was presented by several groups, the Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated protein 9 (CRISPR/Cas9 system)[14–16]. Derived from the adaptive immune system present in almost all archae and a substantial amount of bacteria[17,18], the CRISPR/Cas9 system showed unprecedented ease of design and provided a substantial reduction in time needed to proceed with ones experiment. Main difference between CRISPR/Cas9 system and PNs mentioned above is that for both ZFNs and TALENs, it is necessary to design a whole protein to bind a specific sequence, whereas for CRISPR/Cas9 there is only the necessity for designing a short RNA sequence, called guide RNA (gRNA, sometimes abbreviated sgRNA as single guide RNA), which serves as a „homing device“ for the Cas9 protein[15,16].

The CRISPR/Cas9 system has enabled very efficient production of knock-out mouse models by its ability to introduce indels[19,20], mediated by non-homologous end joining (NHEJ), into sequences of interest or by deletion of a larger gene segment by using 2 gRNAs which are flanking that segment[21]. Production of knock-ins is still plagued by relatively low efficiency, but it is still substantially improved by the use of CRISPR/Cas9 when compared to the classical ESC method.

The process of knocking-in a sequence of interest is mediated by a mechanism called briefly mentioned above, HR, sometimes also called homology directed repair (HDR). HR is a set of DNA repair pathways that is most widely used for repair of DSBs that threaten the genomic integrity of a cell[22]. Without the DSBs, HR still can take place and as mentioned above, it was widely used in the ESC method of mutant mice production. However, by induction of a DSB at a locus of interest, the rate of HR significantly increases[9], thus enabling a broader use of this repair pathway in genome editing. Nevertheless, HR is not the only DSB repair mechanism as DSBs are more predominantly repaired by NHEJ which is, unlike HR, active throughout the whole cell cycle and actively competes with HR for DSB repair[23] and also MMEJ which is the least explored pathway and its effect on HR activity is not yet determined.

In the first chapter of this thesis, I aim to review the history of CRISPR and its mechanism relevant for the genome editing, various modifications of the system that vastly broaden the applications of the system and its current usage in genome editing and preparation of transgenic mice. In the second chapter, I focus on the possibilities of increasing the efficacy of HR through various means.

1. CRISPR/Cas9 system

For a long period of time, elucidation of the exact function of CRISPR-based adaptive immune systems remained elusive. However, the discovery of CRISPR, segments of conserved prokaryotic DNA comprised of arrays of short palindromic repeats and seemingly random segments called spacers, was made nearly 20 years before. CRISPR arrays were first discovered during the study of the *iap* gene which codes an enzyme involved in isozyme conversion of alkaline phosphatase in *E. Coli* in 1987[24]. Over the following years, the CRISPR arrays were described in many other organisms and eventually, it was found that this type of genetic element is present in over 40% of bacteria and 90% of archaea[17].

In 2002, clusters of *cas* (CRISPR-associated) genes were discovered and identified. They were found to be well conserved and usually in close proximity of CRISPR arrays[25], which foreshadowed their crucial role in the system.

Its function as an adaptive defense mechanism of some bacteria and archaea was first proposed in 2005, when a systematic analysis of spacer sequences indicated extrachromosomal or phage origins[26], but it wasn´t until 2007, when Barangou *et al*. provided first experimental evidence for a CRISPRs role in immunity[18]. After this proposal, a barrage of studies was published regarding CRISPRs mechanism of action. In 2008, it was described that CRISPR systems are targeting DNA and are inducing DSB at given loci in an endonuclease-like fashion[27] , although a year later, it was discovered that type III-B Cmr CRISPR system targets RNA rather than DNA[28]. And again in 2008, it was discovered that spacers are transcribed and processed into CRISPR RNAs (crRNAs) that act as a guide for the protein complex[29]. Finally, in 2013, the possibility of its usage in genome editing as an RNA-guided DNA endonuclease was demonstrated by engineering type II CRISPR systems from *Streptococcus pyogenes*[15,16] and *Streptococcus thermopilus*[15] to succesfully edit genome in mammalian cells by using either gRNA[15,16] or crRNA-trans activating RNA(tracrRNA) hybrid[15] as a guide. Differentiation of various CRISPR system types is based on different set of *Cas* genes, that are associated with each type[30] (Figure 1.).

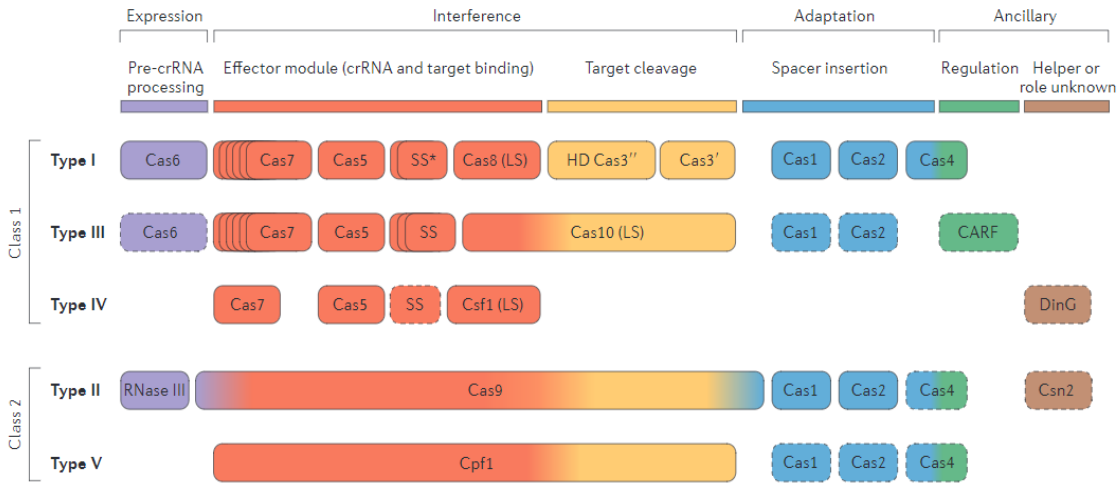

Figure 1. Classification of CRISPR systems
CRISPR systems can be divided into two classes, 1 and 2. The difference between those two classes is based on the composition of the effector module. Whereas class 1 systems have more than one effector module, class 2 systems have only one. This makes class 2 the systems of choice for genome editing due to their simplicity. Class 1 systems are organized into type I and type III, and additional putative type IV. Class 2 two has only one regular type, type II, and a putative type V. Differentiation of these types is based on different set of *cas* genes that each type possess. Putative types are usually not very simillar to other types in its respective class, but they were assigned to the class because of the effector module characteristic. Adapted from An updated evolutionary classification of CRISPR-Cas systems; Makarova *et al*., 2015.

1.1 Mechanism of CRISPR/Cas9

As was said before, the original function of CRISPR/Cas systems is to serve as an adaptive immune system against viral infection in various bacteria and archaea. Its mechanism of action in its original role can be divided into three stages : adaptation (acquisition of novel spacers), expression (crRNA and tracrRNA production and maturation) and interference (DSB induction)[31].(REVIEW) However, since this is largely irrelevant for the purpose of genome editing, I will cover only information that is relevant for this thesis from these three stages. For purpose of genome editing, the use of type II CRISPR/Cas9 system is the most prevalent, so any reference to the system is a reference to the type II, if not said otherwise. The reason for that is because the type II systems have the least components necessary for target DNA cleavage, only requiring Cas9 and crRNA:tracrRNA duplex, making it the easiest type to repurpose for genome editing. This statement is supported by the fact that almost every article cited in this thesis used a type II CRISPR/Cas9 system, with system from *Streptococcus pyogenes* being the most prevalent.

In its native form, homing of Cas9 to its DNA target is enabled by a tracrRNA:crRNA duplex, which forms a complex with Cas9[32] (Figure 2.a). Those RNAs are transribed from

CRISPR arrays mentioned above and matured by RNAseIII[33] before they can form a ribonucleoprotein complex with Cas9. Although this form was also briefly used in pioneering studies, it was soon found that this complex can be substituted by artificialy created gRNAs (Figure 2.b), which are the result of a merge between important domains of respective RNAs[15,16]. While keeping the editing efficiencies nearly unchanged, this circumvention enabled researchers to skip the necessary RNAseIII processing needed for crRNA and tracrRNA maturation[34].

1.2 CRISPR/Cas9 in genome editing

CRISPR/Cas9 system has revolutionized the field of genome editing by unprecedented ease of design compared to other previoulsy desribed PNs, as mentioned above. Further experimentation resulted in modifications of some components of the system and additional improvements of parameters of CRISPR/Cas9 system[34].

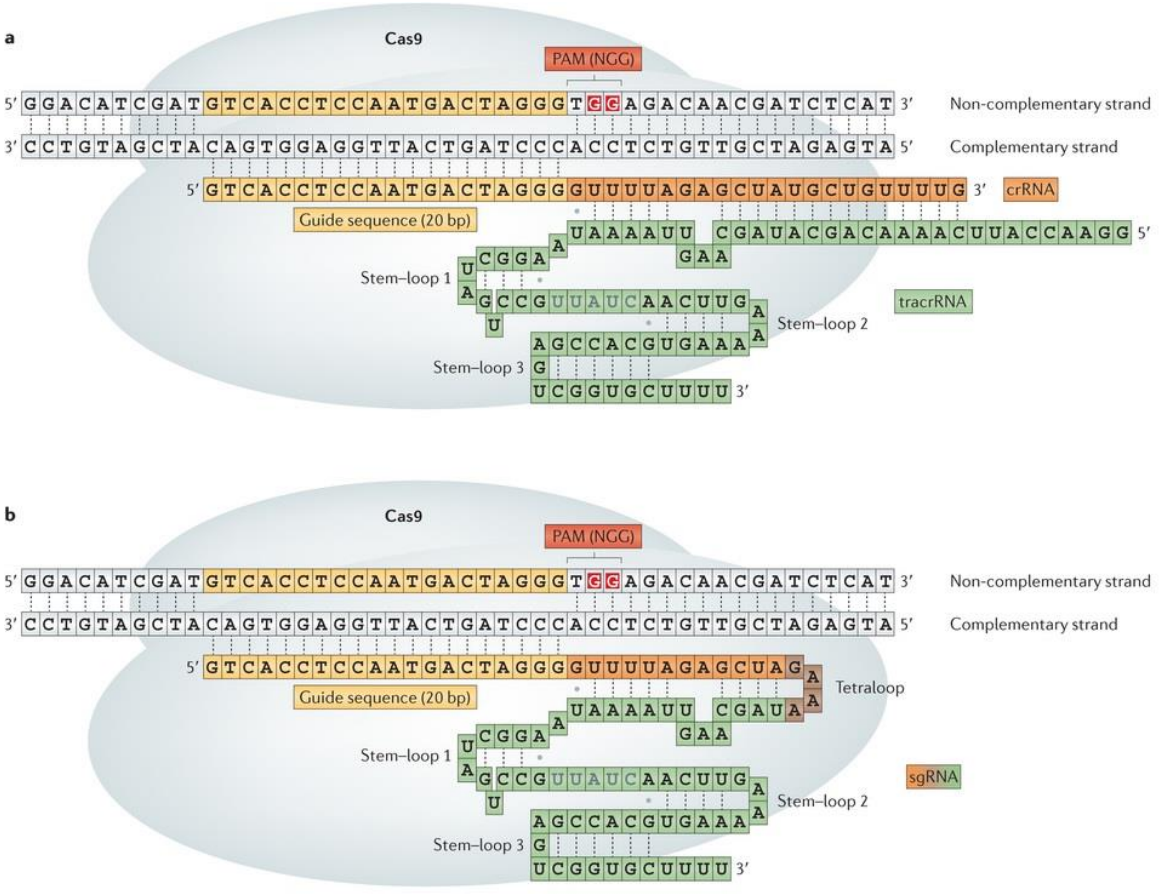

Figure 2. Cas9 in complex with crRNA:tracrRNA duplex and engineered gRNA.
a) Cas9 complex in its naturally occuring formation pairing a target sequence b) Cas9 complex with artificially engineered gRNA pairing a target sequence. Adapted from A guide to genome engineering with programmable nucleases; Kim and Kim, 2014.

Such modifications of CRISPR/Cas9 enabled its easy use in genome editing, since the only components required are now the gRNA and Cas9 protein. Once delivered into the nucleus, the Cas9:gRNA complex binds its target and induces a DSB[27]. DSBs are generally repaired by two major DNA repair mechanisms, NHEJ and HR[23], although a third one, microhomology-mediated end joining (MMEJ) also called alternative NHEJ (alt-NHEJ) exists[35]. However, due to the unpredictable results of MMEJ, it is not employed for the purpose of genome editing. Mechanism of NHEJ is largely used for creation of knock-out alleles, by either introduction of indels into the coding sequence, resulting in a frameshift mutations[19,20] (Figure 3.a), or by deleting a larger DNA fragment, like an exon (Figure 3.b)[21]. However, deleting an exon could still result in transcription and subsequent translation of now altered DNA which could result in unpredictable outcomes on the cellular context, so the first option is largely preferred. On the other hand, HR is used for precise mutations by adding a donor template containing elements of interest and subsequent incorporation into the DNA sequence (Figure 3.c). This mechanism was exploited before in the old ESC method, since HR can take place even without the presence of a DSB, although in much lower frequencies. It it obvious, that the most sought after repair pathway is the HR, since the repertoire of editing possibilities vastly increases, but its effective exploitation remains an unsolved problem for the field of genome editing.

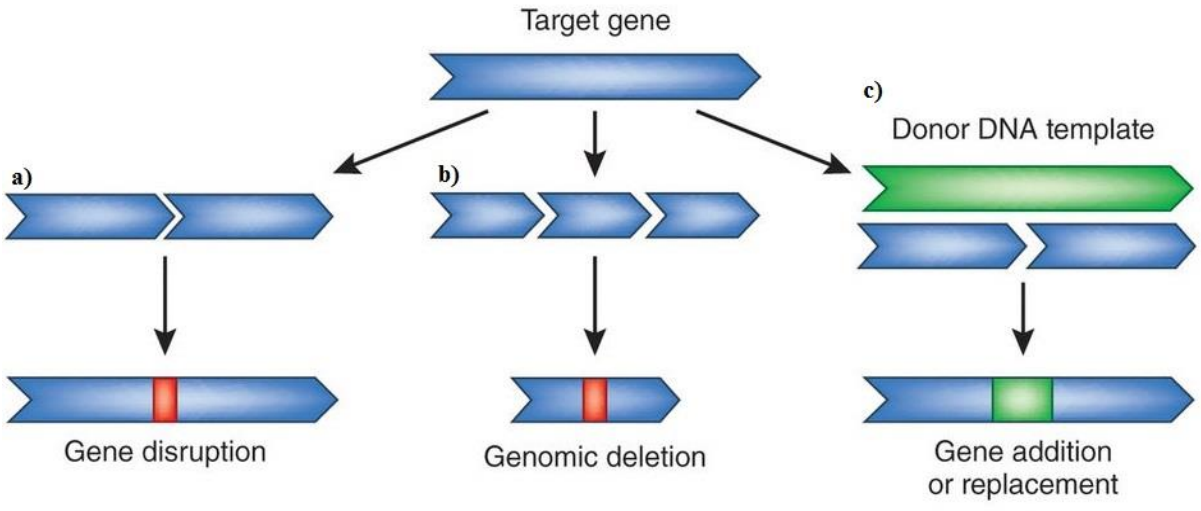

Figure 3. Most common utilizations of CRISPR/Cas9 based genome editing
a) Given sequence is targeted by single gRNA for disruption of target gene by introducing indels due to NHEJ-based DSB resolution. **b)** Given sequence is targeted by two gRNAs flanking the sequence to be deleted. Integrity is then reconstituted by NHEJ. **c)** Given sequence is targeted by single gRNA with adition of a donor DNA template. That template is then integrated into this sequence via HR. Adapted fromGenome engineering: the next genomic revolution; Gersbach, 2014.

All and all, CRISPR/Cas9 is a powerful addition to the pool of tools for genome editing, enabling easy and quick generation of new mutants by reducing the time neccessary for preparation and the amount of know-how needed compared to other mentioned PNs. It can subsitute either TALENs or ZFNs, since it has a much easier design of its guiding mechanisms, and circumvent the need to use ESC in creation of genetically modified organisms, although it has some problems that need to be addressed to perfect the system.

The most limiting factor of CRISPR/Cas9 is the necessity of a so called protospacer-adjacent motif (PAM), which is a short sequence on the 3´end of the protospacer. In its native function, Cas9 uses the absence of PAM to discriminate from the native CRISPR arrays present in the genome of the organism harboring said system[36], but for the usage in genome editing, it presents researchers with a limitation in target selection. For example, for *Streptococcus pyogenes* Cas9 PAM, which is 5´-NGG-3´, it is only present at a frequency circa once per 8 base pairs[14]. Solution for this problem, at least partially, was presented with CRISPR/Cpf1 system mentioned later in this thesis, however much more attractive way to circumvent this obstacle was published recently by Hu *et al*., in which the Cas9 protein was subjected to controlled evolution of PAM-interacting (PI) domain. This resulted in creation of proteins dubbed xCas9s that have their targeting capabilities vastly broadened by their capability to recognize and bind to multiple PAM sequences, seemingly without compromising their off-target activity[37]. However, more studies are necessary for definitive confirmation of this result.

Another problem is optimalization of gRNA design, since it was found that some nucleotides are more favored than others at certain positions. Guanine at 5´ end is very important for trancription from U6 promoter that is widely used because it does not need a poly-A signal, and guanine at 3´ end is helpful for loading onto Cas9 protein[38]. Furthemore, it was found that off-target activity could present a serious task to solve, especially in potentional therapeutic applications, where accurate edits are a requirement. It was reported, that Cas9 can tolerate up to several mismatches in PAM distal region, but does not tolerate any mismatch in so called „seed“ region, which is a sequence of 10-12 nucleotides proximal to PAM sequence[39–43]. This obstacle could be addressed by usage of various modifications of the system like dCas9-FokI or Cas9 nickases described later in the thesis.

For the versatility of use in vast amounts of applications, it is also important to consider how can the genomic localization and genomic context affect the activity of Cas9 in a particular set of circumstances. For example, it was found that nucleosomes impede Cas9

access to the target sequence[44], because they affect the accesibility of the target site. And last but not least, the issue this thesis is largely adressing is the issue of choosing the HR repair mechanism over NHEJ, which is of paramount importance if one is aiming for a tool that enables precise genome editing at reasonable frequencies and as little time as possible. The issue of increasing the HR efficiency and various means to improve it will be thouroughly discussed later in this thesis.

CRISPR/Cas9 system offers new perspectives in production of transgenic mice by enabling their production at a yet unrivalled rate, both when compared to other PNs and the classical method of using ESCs. Three methods have been utilized so far, and that is the microinjection[45–47] (Figure 4.a), electroporation of CRISPR/Cas9 components into zygotes, which is the main advantage of PNs usage in transgenic mice production[48], or by transfection of CRISPR/Cas9 components into ESCs[49] (Figure 4.b). The third approach is pretty similar to the classical way of transgenic mice production and gives reasearchers higher rates of integrations into ESCs, however it still requires a tedious process of selection mentioned above, so the advantage of shorter production time is not present. Since the first and the second approaches are largely preferred due to obvious reasons, I will solely focus on those approaches in my thesis.

Microinjections are routinely used for production of various transgenic animals, mice included. They are still a method of choice for many transgenic facilities for their reliability and ability to efficiently deliver the desired payload into the zygote[45–47]. The most notable modification that can be made in the process is injecting the payload either directly to the pronucleus or to the cytoplasm which, in theory, could affect the editing efficiency. However, a comprehensive study shows that editing efficiency is affected only slightly or not at all, resulting only in decreased embryo survival rate in the case of pronuclear injection[50]. Although this approach is very effective, it presents several problems. Those problems are relatively high labor-intensity with high requirements on skilled personel as inexperience in this approach may severely affect the outcome, very high time demands as zygotes need to be injected individually and necessity for sophisticated equipment like Piezo-electric micromanipulator.

Electroporations present a more perspective approach. Since the inception of its utilization in transgenic mice production, many improvements resulted in a stable electroporation protocols that circumvent earlier problems, like the disruption of zona pellucida[48,51]. Electroporation is also much less labor intensive when compared to

microinjections, by enabling multiple zygotes to be electroporated at once and by not requiring that high experise and skill from laboratory personnel. Nevertheless, electroporation may have limited efficiency in larger payload delivery, so it may be unsuitable for certain uses.

Albeit the use of CRISPR/Cas9 system in mice is far from perfect, it enables easy and fast production of deletions and indel mutations, resulting in production of knock-out mice. It still requires possible cross-breeding, depending on achievement of biallelic mutations or mosaicism, but is substantially faster than previously used methods. The main problem resides in insertion of various specific elements, like loxP sequences for exploitation of the Cre/lox system which enables creation of conditional alleles for study of lethal phenotypes or various reporter alleles for maping expression of targeted genes, localization of said genes by *in vivo* imaging and more. This could be solved by tipping the scales from NHEJ to HR, which is the focus of several groups[52–55], albeit there is not a single publication that addresses this issue directly in mouse zygotes.

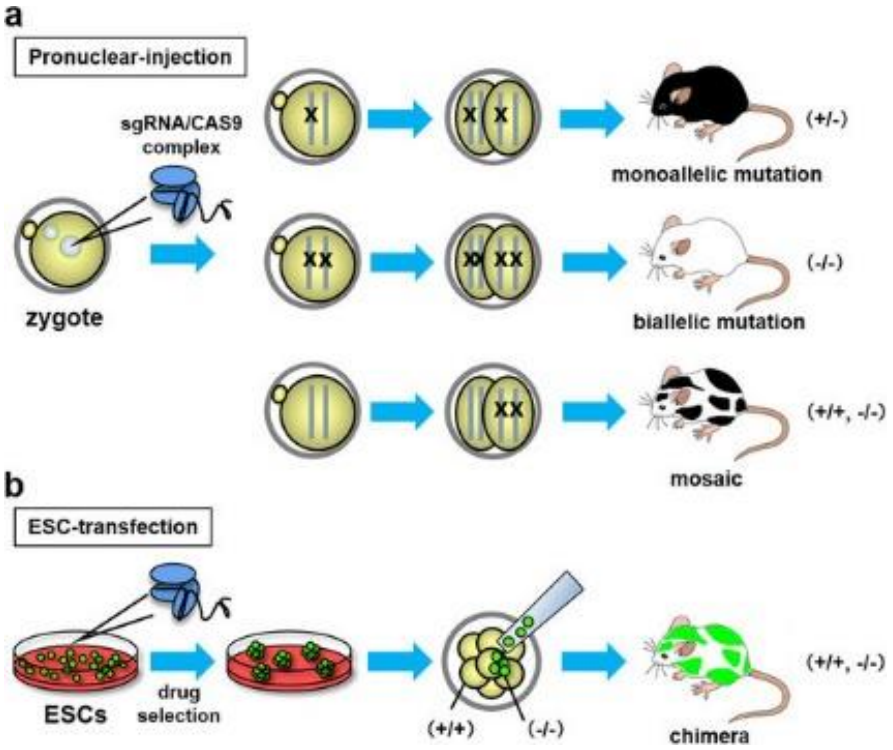

Figure 4. Methods of CRISPR/Cas9 induced modifications in mice.
a) Injection of components necessary for editing into a fertilized egg and subsequent impantation into a foster mother. Resulting mice are then selected and bred to remove possible mosaic mice. Electroporation have similar results in terms of induced mutations but are not pictured here. **b)** In ESC-transfection, CRISPR components are transfected into ESCs, which are then selected on a selection medium. Correctly modified ESCs are, after a few cell division, transferred into a foster mother. Adapted from CRISPR/Cas9 mediated genome editing in ES cells and its application for chimeric analysis in mice; Oji *et al*., 2016.

1.3. Modifications of CRISPR/Cas9 system and novel alternatives

The boom of CRISPR/Cas technology quickly led to development of modified versions of the system, which vastly increased the toolkit of not only genome editing, but also in genome mapping[56] or as a potential transcriptional regulator[57].

Among the most notable of CRISPR/Cas9 modifications are the Cas9 nickases (Cas9n), which are Cas9 proteins with a mutation at either D10A within the RuvC domain or H840A within the HNH domain (Figure 5.). Mutations in these amino acids leads to inactivation of the catalytic site, rendering one of the nuclease domains inactive[58]. This approach was primarily devised for increasing the specifity of the system by increasing the number of nucleotides that are necessary to be recognized by Cas9, by pairing two nickases together, which resulted in 50 - 10 000 fold increase in target efficiency, depending on its target. However it was reavealed, that usage of nickases is influencing not only the targetig specifity, but also the NHEJ:HR ratio. It was described, that usage of Cas9 nickases or double nickases increases the HR efficiency and the NHEJ pathway efficiency is decreased, however this alteration is strongly connected and affected by the cell type and the current stage of the cell cycle. The nickases also show generally lower efficacy of modifications compared to the wild type Cas9[58].

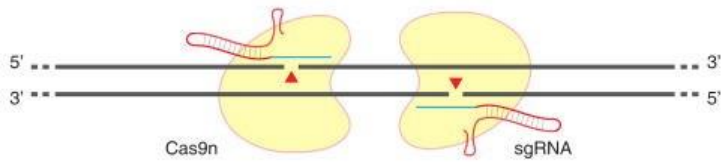

Figure 5. Mechanism of double-nicking approach.
Cas9n binds its target sites by utilization of a gRNA pair targeting opposite strands. Target DNA is then cleaved by its active nuclease domain, resulting in a staggered cut. Adapted from Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity; Ran *et al*., 2013.

Another example of a modification are various dCas variations. This Cas9 has both of its catalytic residues mutated, resulting in catallyticaly dead Cas9 – dCas9. dCas9 retain its DNA binding activity and is further altered by aditions of various elements to either the C or N terminal regions, like various fluorescence tags for genome mapping[56], artificial effectors

like KRAB or VP64[57] for trancsriptional regulation, or merging it with FokI nuclease domain[59].

The most interesting modification of those three is a merger of catalytically inactivated Cas9 and FokI nuclease domain which aims mainly at the reduction of potential off-target effects (Figure 6.)[59]. This is acomplished by fusing Cas9 with FokI nuclease domain on the C-terminus of Cas9. FokI is a type IIS restriction endonuclease found in *Flavobacterium okeanokoites* that needs to create a dimer to cleave[60], thus increasing the neccessary amount of correctly targeted nucleotides twofold[61]. Compared to the paired nickases approach, the dCas9-FokI shows generally lower amount of various unwanted mutations like deletions and substitutions. Also, since there is a necessity of correct guide orientation to dimerize with dCas9-FokI, but not with the paired nickases strategy, the potential off-target sites are further reduced. Furthemore, the paired nickases can bind far away from each other and still induce a cut, but dCas9-FokI is much more restrictive. The sequence between two respective gRNAs must be 14-17 bp for correct positioning of FokI to dimerize. Also, it was shown that there is almost no mutagenesis induced when only one gRNA pair is present, so it is very unprobable that there is a possible danger of half target cuts. All and all, dCas9-FokI presents a viable improvement of specifity and reduction of off-target cleavage.[61] However, it is obvious that the PAM restriction is also doubled, so the targeting options for this approach are more limited.

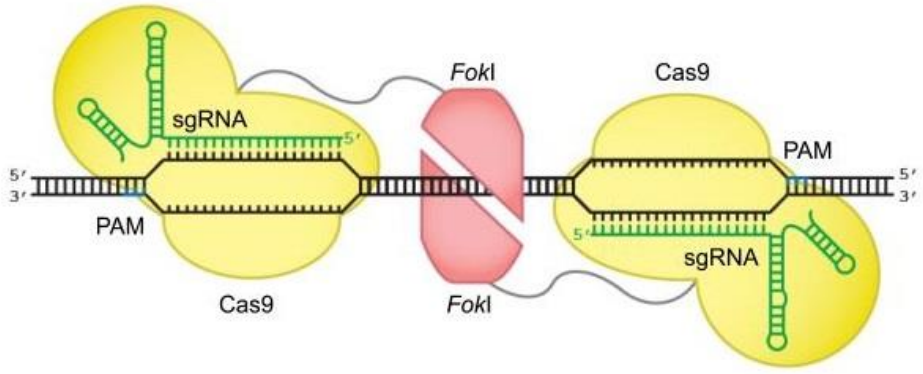

Figure 6. Mechanism of dCas9-FokI approach
dCas9 binds to the target DNA, whereas FokI domain searches for another FokI domain to form a dimer with. After dimerization, FokI homodimer cleaves DNA at a set distance from dCas9 binding sites. However, if dimerization doesn´t take place as FokI monomer alone is unable to cut DNA, which reduces possible off-target effects(not shown). Adapted from Fusion of catalytically inactive Cas9 to *Fok*I nuclease improves the specificity of genome modification; Guilinger et al.; 2014.

Another possible alterative is CRISPR/Cpf1 system[62] (Figure 7.). The most striking feature is a T-rich PAM sequence, which enables broader targeting range of T-rich genomes and gives researchers much more variability than CRISPR/Cas9, which has G-rich PAM

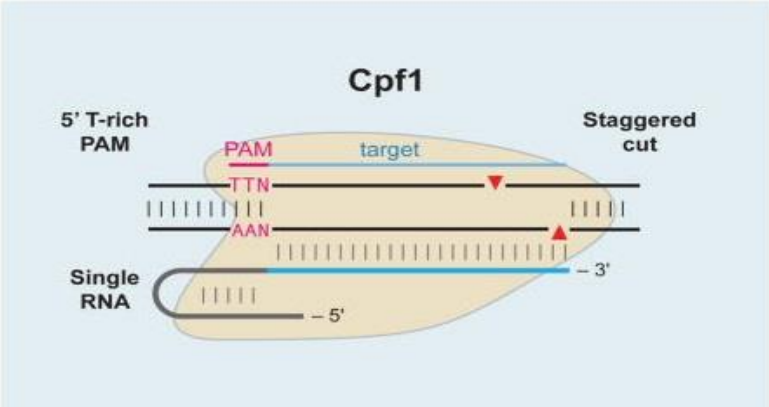

Figure 7. Cpf1 in complex with its guide RNA and target DNA.
Unlike the Cas9 system, Cpf1 has shorter sgRNAs and also different PAM requirements.Cpf1 also produces staggered cuts, unlike wild type Cas9, which produces blunt ends. This presents new opportunities for RGEN based genome editing. Adapted from Cpf1 is a single RNA-guided endonuclease of a Class 2 CRISPR-Cas system; Zetsche *et al*., 2015.

sequences. Also, it was theoreticized that since Cpf1 cuts far away from its seed region, it could cut the target until the recognition would be abolished by mutating the target site beyond recognition, which could improve precise HR-based editing efficiencies. [62] However, unpublished data from Berndt Zetsche suggest that this is not the case for the editing in mouse zygotes but further confirmation is neccessary.

A novel approach for editing single nucleotides was developed recently by Komor *et al*.[63] and Gaudelli *et al*.[64], which utilizes a fusion of various versions of Cas9 and enzymes that catalyze DNA editing. Altogether, 2 distinct systems were developed. The first one, base editors (BEs), are capable of C→T conversion and are 3 in total. The most developed is the BE3, which is composed of Cas9n, rat APOBEC1 cytidine deaminase fused to the N-terminus of Cas9n and uracil DNA glycosylase inhibitor (UGI) from *Bacillus subtilis* bacteriophage PBS1 fused to the C-terminus of Cas9n. The rationale behind picking those enzymes is simple, APOBEC1 would deaminate cytosine into uracil that is recognized as thymine during DNA replication and UGI would block the endogenous uracil DNA glycosylases (UDGs) which catalyze uracil removal. This is important because normally, uracil is removed by UDGs which results in base-excision repair of given base and in most cases, reversal back to C:G. They tested this approach on several cell lines, for example on immortalized mouse astrocytes, in which they tried to correct human *APOE4* gene that replaced endogenous *Apoe* gene by conversion of Arg158 to Cys158 (*APOE3r*). In comparison, base editors proved to be

several hundred times more effective at this particular locus than wtCas9 and ssDNA donor template[63]. This substantial increment in efficiency seemed consistent, although more data is needed for complete and definitive evaluation. The second one, adenine base editors (ABEs), are capable of A→G conversion and are much more numerous in total. The most crucial problem encountered during the development of ABE was the fact that there was no adenine deaminase that would accept DNA as a substrate. This was solved by fusing chosen candidate adenine deaminase, *E. coli* TadA, to N-terminus of Cas9n and subjecting it to several rounds of directed evolution in order to change its substrate specificity and further increase its editing efficiency and specificity[64]. Adenine deamination results in the formation of inosine, which pairs with cytosine and is therefore recognized as guanine during DNA replication. The most developed is ABE7.10, altered and upgraded version of original ABEs which has broad application window and high rate of both efficiency and specificity, although it may be advised to use earlier version of ABEs for certain applications. However, ABE7.10. shows highest overall performance compared to previous versions and presents researchers with versatile tool for A→G conversion, reaching efficiencies that surpass regular HDR method substantially. The amount of efficiency increment is variable and may depend on several factors like target genomic locus, cellular context and overall time during which cells are exposed to ABE[64]. However, detailed description and evaluation of both BEs and ABEs is beyond the scope of this thesis and it will be necessary to test these approaches in various cellular environments to evaluate their versatility. This is especially true for mouse zygotes, since there is no study yet which tries this approach in mutant mice production.

2. Increasing the efficiency of HR-mediated genome editing

As was mentioned before, CRISPR/Cas9 serves the genome editing as a tool to induce DSBs at given loci, which are then repaired by endogenous repair mechanisms. Those machanisms include NHEJ, HR and alternative NHEJ (alt-NHEJ) also called microhomology-mediated end-joining (MMEJ) (Figure 8.). The activity of these mechanisms is concurrent to each other, i. e. they compete, but most of the DSBs are repaired via NHEJ, which is active throughout the whole cell cycle unlike HR and alt-NHEJ, that are mostly active in G and S2 phase[65].

NHEJ is one of the major repair pathways of DNA DSB damage. It is a fast and crucial part of keeping the integrity of the genome and plays a key role in V(D)J recombination[66], but its usage in genome editing is limited due to the nature of the alterations

that are made in the process i.e. insertions or deletions of several nucleotides[67]. It is useful in creation of KO alleles, but it lacks the possibility to introduce precise modifications like HR. NHEJ machinery is composed of a set of proteins, the most important proteins are KU70/ KU80[68], DNA ligase IV[69] and DNA-PKs[70]. The pathway starts by recognition of a DSB and subsequent binding by KU70/KU80 heterodimer. This prevents further resection that can lead to alternative repair pathway[68]. The whole complex of proteins then assembles on the DSB and the strands are ligated together by DNA Ligase IV after processing[69].

Alt-NHEJ, or MMEJ, is usually employed when neither NHEJ nor HR can be utilized for the damage, or more likely by inactivation or depletion of DNA Ligase IV[71], since it was shown that NHEJ activity and HR activity do not have a direct correlation between each other[72]. It is mostly active is G2 and S phase of the cell cycle[35]. The alt-NHEJ employs a set of proteins that facilitates the joining based on microhomology on the strands that are about to be ligated. Precise mechanism of this pathway remains to be elucidated, but so far it was discovered that the most important proteins involved are PARP-1[73] and DNA Ligase III[74].

HR is a set of repair mechanisms relying on a homologous DNA sequence from which the damaged template is reconstructed. HR is the pathway utilized in precise genome editing. HR can be generally divided into three steps : presynapsis (DSB recognition), synapsis (D-loop creation) and postsynapsis (DSB resolution and repair), but there are several models for HR-mediated repair, each with a distinc mechanism[75]. Those models and their relations are far from being completely elucidated. There are several important features for HR-based DNA repair, the most important being the 5´-3´resection of DSBs in order to enable template annealing and inhibit Ku70/80 heterodimer from binding and tiping the scales in favor of NHEJ[76–78]. The most important components of each system are noted in Table 1.

DNA repair pathway	Important components
NHEJ	Ku70, Ku80, DNA-PKcs, DNA ligase I XRCC4, Artemis
MMEJ	PARP-1, DNA ligase III
HR	Rad51, RPA, BRCA2, ATM, Nbs1, Rad Mre11, Rad52, CtIP, Exo1

Table 1. Key proteins in DSB DNA repair pathways

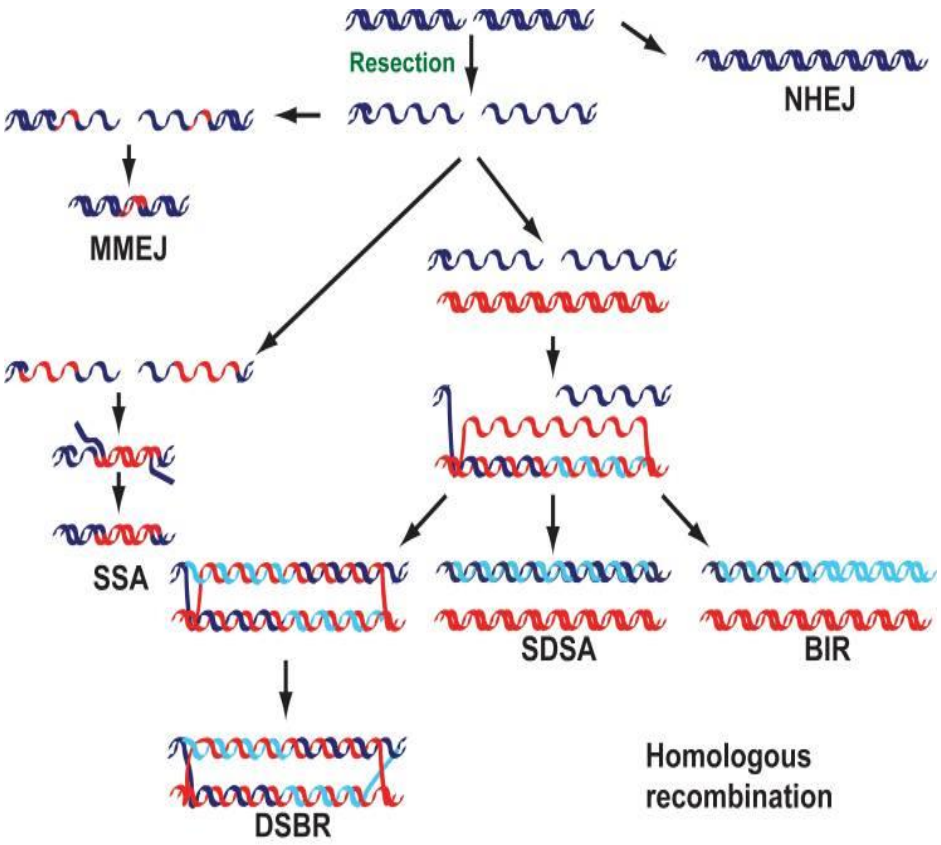

Figure 8. Possible repair pathways of DNA DSBs.
Repair of DSBs can be undertaken by several repair pathways. The most prevalent in cells is the non-homologous end-joining(NHEJ), showed in upper right, which, after none or minimal end resection, ligates ends of DNA strands together, which leads to indels. The least explored pathway is the microhomology-mediated end-joining(MMEJ), showed in upper left. It resolves DSBs by annealing to small homology regions on paralell strands, resulting in deletion of sequence between those regions and loss of information. And finally, the set of most preciseDSB repair mechanism, homologous recombination(HR). HR consists of several distinct repair pathways, all relying on long homology regions on which they base their repair. Those pathways are double-strand break repair(DSBR), synthesis-dependent strand annealing(SDSA), break-induced repair(BIR) and single-strand annealing(SSA). All of them are fairly reliable, with SSA being the most deletirious due to potential deletions of sequences between homologies(as seen in picture). Adapted from DNA resection in eukaryotes: deciding how to fix the break; Huertas, 2010.

2.1. NHEJ inhibition

NHEJ inhibition was demonstrated as a viable way to increase the frequency of HR in various organisms. Possible ways of silencing the NHEJ pathway are several, all targeting key components of the NHEJ machinery. Among those targets are KU70[79], and mostly DNA Ligase IV[52–54,80], which plays key part in this repair pathway. Another possible inhibitor is against DNA-PKs, NU7026[81,82], but that inhibitor was never tried as a mean to increase HR efficiency. The inhibition can be mediated by various means, for example by gene silencing or by short hairpin RNAs (shRNAs) or by coexpression of adenovirus 4 proteins E1B55K

and E4orf6[79]. Although the NHEJ repair pathway activity does not directly correlate with HR pathway activity, it indeed affects HR[72]. In a recent study, it was shown that in the presence of either KU70 or DNA Ligase IV shRNAs, the activity increased up to fivefold and in the presence of Ad4 protein pair, it increased up to sevenfold[79]. Another potential inhibitor was found, DNA ligase IV inhibitor SCR7[52–54,79]. This molecule showed very promising results in several studies, but different studies had come to a conclusion that Scr7 doesn´t affect the HR efficiency[80] nor it is a selective DNA ligase IV inhibitor[83]. This may be due to various cell types and organisms used during those studies, further implicating strong importance of how given cellular context affects actions and results of DNA repair machineries. Further investigation regarding those discrepancies is necessary for correct evaluation of current data.

2.2. Overproduction of HR repair pathway components

Theoretically, it could be possible to further increase the HR efficiency by increasing the amount of HR pathway components like Rad51, analogically to the NHEJ inhibition described above. This should present DSBs with more frequent HR machinery components, which will in turn lead to increased HR efficacy. Surprisingly, not many studies were yet published regading this possibility.

It was shown by Song *et al*., that increasing HR efficacy via stimulation of its component production (Rad51) or by stimulation of Rad51 with RS-1 enhancer is a viable option to increase the HR efficiency[80]. However, this experiment was performed on a rabbit model which has different cellular context than a mouse model. This may affect the viability of such method on mice, so further testing will be necessary. RS-1 showed 2-4 fold increase in HR efficiency (tested on 2 different targets) and Rad51 mRNA showed simillar results. Further testing is needed, like testing stimulation of other HR components, like RPA, BRCA2 or MRN complex [80].

Another study published by Paulsen *et al*. in 2017 utilized a more complex approach by transient ectopic expression of various genes, which products participate in DNA repair pathways like Rad51, Rad52 or Exo1. They achieved circa 2,7 fold increment in correct insertions when coexpressing Rad52 and dn53BP1 in HEK293 cell line while using a ssDNA template. dn53BP1 is a modified version of naturally occuring 53BP1 protein, which contains tandem Tudor domains and acts as an inhibitor of its native form which seems to be important for NHEJ and also Rad51-based HR. Introduction of dsDNA template did not result in any substantial increment in HR efficiency, suggesting that this approach utilizes the Rad52-dependent SSA for integration od donor DNA[55].

Rather unual approach compared to those described above was taken by Charpentier *et al.*, that utilizes a fusion of minimal N-terminal fragment of CtIP protein necessary for increasing the efficiency of HR, dubbed HDR enhancer (HE), and Cas9[84]. The N-terminal fragment contains several essential domains for CtIP function in HR, like CDK phosphorylation domains[85] and the multimerization domain[86,87]. In this study, said approach was evaluated in several environments, for example HEK293T cells and rat oocytes. In HEK293T cells, Cas9-HE increased the knock-in efficiency roughly by 2-3 fold, depending on targeted genomic locus[84]. It is also important to note, that the design of gRNA played a significant role in overall efficiencies and must be taken into account. In rat oocytes, the results were inconclusive at best since the amount of embryos evaluated was not significant and only 2 genomic loci were tested. Furthemore, the results seem to contradict themselves, since in one case the fusion indeed achieved higher efficiencies than regular Cas9, but in other instance it was the exact opposite. Further evaluation of its ability to increase HR efficiency is definitely needed, especially in relation to production of rodent animal models.

So far, this seems to be a way to consistently improve HR-based integration efficiencies, however further studies are needed to confirm the validity of such approach. Lack of more studies regarding application of this method prevents us from correctly and decisively evaluating presented data. Also, it has been shown that its efficiency is strongly dependent of target locus of Cas9 and the cell type used for given study. Unfortunately, not a single study applying this approach to transgenic mice production or its usage in mouse cells has been published.

2.3. Alternative ways of increasing HR efficiency

It was also shown by Hatada *et al*, that low dose of of gamma ray irradiation(LDI) vastly increases the frequency of HR in human pluripotent stem cells, which is in contrary to the popular belief that irradiation generally kills cells either by induction of DSBs beyond repair and by starting the apoptotic pathway.[88] This method was so far tested in human pluripotent stem cells, but it may be viable in other applications. The correct dosage of irradiation is absolutely crucial for this method, which was shown to be effective at 0,4 Gy, but will likely differ for various species and cell types. More irradiation led to induction of apoptotic pathway and death of the cells within 24 hours[88]. Altogether, LDI presents an interesting alternative to other HR improving methods, but it is necessary to further analyze its possible non-desirable effects, mostly possible developmental impairment due to unexpected mutations and possibilities of its utilization in mouse zygotes.

Another potential HR enhancers, L755507 and Brefeldin A, were identified by Yu *et al*[53]. L755507 is a β3-adrenergic receptor agonist[89], whereas Brefeldin A is an inhibitor of protein transport from ER to GA[90]. These molecules were succesfully used in increasing the efficiency of HR, although the efficiency increase was highly dependent on the donor template length. For larger inserts of several kilobases, the L755507 showed 3x larger efficiency than the control without it, and Brefeldin A showed 2x increased efficiency. In this study, a plasmid donor template was used and it showed consistent HR increments in various cell types. For smaller inserts of several hundred bp, only L755507 was evaluated with almost 9x increased efficiency. Both molecules showed little to no toxicity when used in their optimized concentrations.[53]

It was shown by Liang *et al*., that sequential delivery of donor DNA and Cas9 RNP substantially increases HR efficiency by rougly 2x, regardless of the donor DNA form (ssDNA or dsDNA)[91]. However, to my knowledge, there is no other study yet published which focus is on increasing HR efficacy through this method. The necessity of further investigation is necessary, since the increase in HR efficacy was substantial.

2.4. Donor template design

Another crucial factor for HR efficiency and specificity is the design of donor DNA template, which greatly affects overall performance. Modifications of donor templates reside in several alterations. Those modifications resides in employment of either ssDNA or dsDNA template, in the case of ssDNA if it is complementary to PAM strand or non-PAM strand and in the case of dsDNA templates if it is a circular or linearized plasmid or a PCR product. Additional factors are the length of homology arms and chemical modifications of said template. Major problem in evaluation of published results dwell in either the fact that it is a pioneering study in its approach or that the groups responsible for said results are not using standardized approaches directly comparable with previously published data. A study encompassing all these factors simultaneously would provide a more definitive point than a series of studies that approach this issue separately. However, such a study would require massive effort and colaboration of several research groups to be finished at a reasonable time. And finally, since a direct comparison focusing on evaluation of said variables in preparation of transgenic mice is still unanavailable, I will present here only cell culture based results[92,93], with only one exception[94] and one novel method of donor design used in mutant mice production that has no direct comparison of said variables[95].

A study published in 2017 by Li *et al*. has provided the most comprehensive comparison of the question of ssDNA vs dsDNA template.[92] They evaluated the use of ssDNA, circular plasmid and PCR product templates with various homology arms lengths and also the effect of ssDNA template orientation, i.e. if the donor template is complementary or non-complementary to target strand. Based on the information from their experiment in HEK293T cell line, PCR products seemed as the most efficient, however the issue of off-target integration proved to be the most crucial issue for this type of donor template. By ddPCR assay evaluation, PCR products showed only 20-40% integration specificity, circular plasmids 40-80% integration specificity and ssDNA templates were almost identical to control, showing the highest specificity of all compared templates. Further evaluation was focused on ssDNA templates, their ideal homology arms length and the effect of template orientation. 400-700 base pairs, dependent on genomic locus, tend to show maximum knock-in efficiency and any further increments does not lead to further improvements. Those results were consistent in both HEK293T and K562 cell lines and both loci tested. Next, they tested the effect of complementary and non-complementary donor template on three different loci. In one exception, non-complementary proved to be twice as effective as the complementary donor, but showed similar efficiencies on other two loci. This could mean that it is dependent on the targeted locus, but further evaluation is needed to eliminate the possibility of an error. Li *et al*. presented that the ssDNA donor templated provide researchers with high on-target efficiency for donors ranging from several hundred base pairs to several kilobases[92]. Nevertheless, some evidence, which points out the fact that ssDNA templates are probably integrated via different DNA repair pathway than dsDNA templates[55] further amplifies the necessity to elucidate exact mechanisms behind DNA repair machinery and their relations.

Another approach was taken in a study by Richardson *et al*., in which the group designed ssDNA donors with assymetrical homology arms either complementary or non-complementary to the target strand of various lengths.[93] Non-complementary strands have fared far better in this experiment. In their optimal achieved design with 36 base pairs homology arm on the PAM-distal side and 91 base pairs on the PAM-proximal side, non-complementary strand donor template had achieved almost two times higher efficiency that complementary donor template of similar homology arm design. Result of non-complementary strand preference was consistent in HEK293 and K562 cell lines and at multiple genomic loci.[93] However, it should be noted that in this experiment, Richardson *et al*. only attempted to change several bases and integration of a larger DNA fragment was not

attempted. Validation of its usefulness for larger integrations is needed, but it nevertheless presents an interesting way of increasing efficiencies for introduction of point mutations.

Renaud *et al*. showed, that modification of ssDNA ends with phosphorothioate (PS) increased the HR efficiency substantially, by inserting a short loxP sequence at ROSA safe harbor locus. Whereas donor without the phosphorothioate modification resulted in precise edit rates at only 9,5%, donor with phosphorothiorate modification resulted in 62,5% knock-in rate[94]. However, the study was carried at a statistically irelevant amount of zygotes, 21 for non-PS and 8 for PS, so further evaluation by a statistically relevant amount of zygotes is necessary to provide a definitive result.

Quadros *et al*. described a method dubbed „Easi-CRISPR" for mutant mice production by using long ssDNA donors with homology arms significantly shorter than previously described for inserts o such length. Homology arms of just 50-100 nucleotides were long enough to enable researchers to achieve knock-in efficiencies ranging usually from 30% to 60%, but for some loci it may be even 100% or, on the other side, just 8,5%.[95] This further proves the importance of cellular and genomic context in regard to genome editing. However, this method was proved to work only on inserts up to 1,5 kb in length, so this approach may not be feasible for larger inserts and requires further evaluation.

Summary

CRISPR/Cas9 has shown a tremendous potential in various fields, transforming the field of genome editing as we know it by its quick and easy preparation and use. Although the PNs, like ZFNs and TALENs, were used before, their disadvantages had kept them from everyday use in laboratories for the most part. Tedious evaluation of activity and correct design, in case of ZFNs, and laborous cloning, in case of TALENs, posed a serious impediment in implementation of these attractive technologies for use on a massive scale, a feat that CRISPR/Cas9 system achieved with ease. Just several years after first publication desribing its use as a PN was enough for the scientific community to adapt this system for a broad range of applications like animal model production, immunotherapy[96,97], crop engineering[98] and more.

During a short period of time, CRISPR/Cas9 has been successfully modified to address several issues that the system presented, like possible off-target effects, specificity and targeting possibilities (dCas9-FokI, Cas9n, xCas9) and even reengineered with customized

effector modules for novel applications (VP64 and KRAB fusions, BEs, ABEs,...). Although the solution or improvement of said problems is indeed very helpful, especially in the case of BEs, ABEs and xCas9 addition to the repertoire, those tools are still very new and their effects on various genomic loci and cellular context needs to be further evaluated.

Even though many problems that the system has presented have been solved or improved substantially, the most crucial problem still resides. Competition of NHEJ and HR pathways presents a serious issue for researchers attempting a precise insertion of a DNA fragment into a locus of interest. Several groups have been addressing this issue by different methods with variable levels of success. Albeit some of the presented results seem to be very promising, especially, donor template design and utilization of HR pathway components, further experimental data are needed to complete our picture of DSB repair pathways. That would enable us to predict the effect of any induced modulation with more confidence and success.

Production of transgenic animals, and in this case mice, has been taken by a storm by enabling researchers to produce them at record rates and times, with efficiencies comparable or better to already available PNs. However, CRISPR/Cas9 is still a relatively new technology and as such, it needs to be further studied for s complete elucitadion of both its potential but also its limitations. Such elucidation will require further reasearch into the exact workings of DNA repair mechanisms, as they are one of the driving forces in genome editing. Due to our incomplete understanding of how they work, their exact relations and how may our interventions change the cellular context, manipulation of DNA repair mechanisms to tip the scales in favor of HR is hampered and precise insertions of DNA fragments into living organisms still remain an issue for contemporary researchers. And last but not least, further evaluation of all previously mentioned method of CRISPR/Cas9 system enhancements in the environment of mouse zygotes is required. Given the novelty of the system, it is understandable that our knowledge so far is still growing, but that prevents us from correctly evaluating all potential factors that can affect genome editing. So far, only minimum amount of studies attempted to increase knock-ins in the environment of mouse zygotes.

Reviews are marked with an asterisk (*)

Literature

* 1. Perlman, R. L. Mouse models of human disease. *Evol. Med. Public Health* **2016,** 170–176 (2016).

*2. Cheon, D.-J. & Orsulic, S. Mouse models of cancer. *Annu. Rev. Pathol.* **6,** 95–119 (2011).

*3. Ly, J. P., Onay, T. & Quaggin, S. E. Mouse models to study kidney development, function and disease. *Curr. Opin. Nephrol. Hypertens.* **20,** 382–390 (2011).

*4. Capecchi, M. R. Altering the genome by homologous recombination. *Science* **244,** 1288–1292 (1989).

5. Hansen, G. M. *et al.* Large-scale gene trapping in C57BL/6N mouse embryonic stem cells. *Genome Res.* **18,** 1670–1679 (2008).

6. Keskintepe, L., Norris, K., Pacholczyk, G., Dederscheck, S. M. & Eroglu, A. Derivation and comparison of C57BL/6 embryonic stem cells to a widely used 129 embryonic stem cell line. *Transgenic Res.* **16,** 751–758 (2007).

*7. Tompers Dennie M. & Labosky Patricia A. Electroporation of Murine Embryonic Stem Cells: A Step-by-Step Guide. *STEM CELLS* **22,** 243–249 (2008).

8. Jasin, M. & Berg, P. Homologous integration in mammalian cells without target gene selection. *Genes Dev.* **2,** 1353–1363 (1988).

9. Liang, F., Han, M., Romanienko, P. J. & Jasin, M. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 5172–5177 (1998).

10. 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 (1996).

11. Mani, M., Kandavelou, K., Dy, F. J., Durai, S. & Chandrasegaran, S. Design, engineering, and characterization of zinc finger nucleases. *Biochem. Biophys. Res. Commun.* **335,** 447–457 (2005).

12. Christian, M. *et al.* Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics* **186,** 757–761 (2010).

13. Cermak, T. *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **39,** e82 (2011).

14. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337,** 816–821 (2012).

15. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339,** 819–823 (2013).

16. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339,** 823–826 (2013).

17. Mojica, F. J. M., Díez-Villaseñor, 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 (2000).

18. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315,** 1709–1712 (2007).

19. Wang, H. *et al.* One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. *Cell* **153,** 910–918 (2013).

20. Zhou, J. *et al.* One-step generation of different immunodeficient mice with multiple gene modifications by CRISPR/Cas9 mediated genome engineering. *Int. J. Biochem. Cell Biol.* **46,** 49–55 (2014).

21. Wang, L. *et al.* Large genomic fragment deletion and functional gene cassette knock-in via Cas9 protein mediated genome editing in one-cell rodent embryos. *Sci. Rep.* **5,** 17517 (2015).

22. Serrano, L. *et al.* Homologous recombination conserves DNA sequence integrity throughout the cell cycle in embryonic stem cells. *Stem Cells Dev.* **20,** 363–374 (2011).

23. Sonoda, E., Hochegger, H., Saberi, A., Taniguchi, Y. & Takeda, S. Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. *DNA Repair* **5,** 1021–1029 (2006).

24. 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,** 5429–5433 (1987).

25. Jansen, R., Embden, J. D. A. van, Gaastra, W. & Schouls, L. M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **43,** 1565–1575 (2002).

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

27. Marraffini, L. A. & Sontheimer, E. J. CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. *Science* **322,** 1843–1845 (2008).

28. Hale, C. R. *et al.* RNA-Guided RNA Cleavage by a CRISPR RNA-Cas Protein Complex. *Cell* **139,** 945–956 (2009).

29. Brouns, S. J. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321,** 960–964 (2008).

*30. Makarova, K. S. *et al.* An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* **13,** 722–736 (2015).

*31. Rath, D., Amlinger, L., Rath, A. & Lundgren, M. The CRISPR-Cas immune system: Biology, mechanisms and applications. *Biochimie* **117,** 119–128 (2015).

32. Karvelis, T. *et al.* crRNA and tracrRNA guide Cas9-mediated DNA interference in Streptococcus thermophilus. *RNA Biol.* **10,** 841–851 (2013).

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

34. 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 (2014).

35. 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 (2013).

36. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155,** 733–740 (2009).

37. Hu, J. H. *et al.* Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556,** 57 (2018).

38. 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 (2014).

39. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31,** 827–832 (2013).

40. Fu, Y. *et al.* High frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31,** 822–826 (2013).

41. Mali, P. *et al.* CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* **31,** 833–838 (2013).

42. Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* **31,** 839–843 (2013).

43. Sternberg, S. H., LaFrance, B., Kaplan, M. & Doudna, J. A. Conformational control of DNA target cleavage by CRISPR–Cas9. *Nature* **527,** 110 (2015).

44. Horlbeck, M. A. *et al.* Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *eLife* **5,** e12677 (2016).

45. Nakagawa, Y. *et al.* Production of knockout mice by DNA microinjection of various CRISPR/Cas9 vectors into freeze-thawed fertilized oocytes. *BMC Biotechnol.* **15,** (2015).

46. Terao, M. *et al.* Utilization of the CRISPR/Cas9 system for the efficient production of mutant mice using crRNA/tracrRNA with Cas9 nickase and FokI-dCas9. *Exp. Anim.* **65,** 275–283 (2016).

47. Hara, S. *et al.* Microinjection-based generation of mutant mice with a double mutation and a 0.5 Mb deletion in their genome by the CRISPR/Cas9 system. *J. Reprod. Dev.* **62,** 531–536 (2016).

48. Hashimoto, M. & Takemoto, T. Electroporation enables the efficient mRNA delivery into the mouse zygotes and facilitates CRISPR/Cas9-based genome editing. *Sci. Rep.* **5,** 11315 (2015).

49. Oji, A. *et al.* CRISPR/Cas9 mediated genome editing in ES cells and its application for chimeric analysis in mice. *Sci. Rep.* **6,** 31666 (2016).

50. Horii, T. *et al.* Validation of microinjection methods for generating knockout mice by CRISPR/Cas-mediated genome engineering. *Sci. Rep.* **4,** 4513 (2014).

51. Grabarek, J. B., Plusa, B., Glover, D. M. & Zernicka-Goetz, M. Efficient delivery of dsRNA into zona-enclosed mouse oocytes and preimplantation embryos by electroporation. *Genes. N. Y. N 2000* **32,** 269–276 (2002).

52. Maruyama, T. *et al.* Inhibition of non-homologous end joining increases the efficiency of CRISPR/Cas9-mediated precise [TM: inserted] genome editing. *Nat. Biotechnol.* **33,** 538–542 (2015).

53. Yu, C. *et al.* Small Molecules Enhance CRISPR Genome Editing in Pluripotent Stem Cells. *Cell Stem Cell* **16,** 142–147 (2015).

54. Hu, Z. *et al.* Ligase IV inhibitor SCR7 enhances gene editing directed by CRISPR-Cas9 and ssODN in human cancer cells. *Cell Biosci.* **8,** 12 (2018).

55. 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 (2017).

56. Chen, B. *et al.* Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. *Cell* **155,** 1479–1491 (2013).

57. Kearns, N. A. *et al.* Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. *Dev. Camb. Engl.* **141,** 219–223 (2014).

58. Ran, F. A. *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154,** 1380–1389 (2013).

59. Guilinger, J. P., Thompson, D. B. & Liu, D. R. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* **32,** 577–582 (2014).

60. Wah, D. A., Bitinaite, J., Schildkraut, I. & Aggarwal, A. K. Structure of FokI has implications for DNA cleavage. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 10564–10569 (1998).

61. Tsai, S. Q. *et al.* Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* **32,** 569–576 (2014).

62. Zetsche, B. *et al.* Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* **163,** 759–771 (2015).

63. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533,** 420–424 (2016).

64. Gaudelli, N. M. *et al.* Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551,** 464–471 (2017).

65. Mathiasen, D. P. & Lisby, M. Cell cycle regulation of homologous recombination in Saccharomyces cerevisiae. *FEMS Microbiol. Rev.* **38,** 172–184 (2014).

66. Dai, Y. *et al.* Nonhomologous end joining and V(D)J recombination require an additional factor. *Proc. Natl. Acad. Sci.* **100,** 2462–2467 (2003).

67. Zaboikin, M., Zaboikina, T., Freter, C. & Srinivasakumar, N. Non-Homologous End Joining and Homology Directed DNA Repair Frequency of Double-Stranded Breaks Introduced by Genome Editing Reagents. *PloS One* **12,** e0169931 (2017).

68. Fell, V. L. & Schild-Poulter, C. Ku Regulates Signaling to DNA Damage Response Pathways through the Ku70 von Willebrand A Domain. *Mol. Cell. Biol.* **32,** 76–87 (2012).

69. Wilson, T. E., Grawunder, U. & Lieber, M. R. Yeast DNA ligase IV mediates non-homologous DNA end joining. *Nature* **388,** 495–498 (1997).

70. Uematsu, N. *et al.* Autophosphorylation of DNA-PKCS regulates its dynamics at DNA double-strand breaks. *J. Cell Biol.* **177,** 219–229 (2007).

71. Wang, H. *et al.* Genetic evidence for the involvement of DNA ligase IV in the DNA-PK-dependent pathway of non-homologous end joining in mammalian cells. *Nucleic Acids Res.* **29,** 1653–1660 (2001).

72. Miyaoka, Y. *et al.* Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. *Sci. Rep.* **6,** 23549 (2016).

73. Robert, I., Dantzer, F. & Reina-San-Martin, B. Parp1 facilitates alternative NHEJ, whereas Parp2 suppresses IgH/c-myc translocations during immunoglobulin class switch recombination. *J. Exp. Med.* **206,** 1047–1056 (2009).

74. He, M.-D. *et al.* Efficient ligase 3-dependent microhomology-mediated end joining repair of DNA double-strand breaks in zebrafish embryos. *Mutat. Res.* **780,** 86–96 (2015).

*75. Krejci, L., Altmannova, V., Spirek, M. & Zhao, X. Homologous recombination and its regulation. *Nucleic Acids Res.* **40,** 5795–5818 (2012).

76. Limbo, O. *et al.* Ctp1 is a cell-cycle-regulated protein that functions with Mre11 complex to control double-strand break repair by homologous recombination. *Mol. Cell* **28,** 134–146 (2007).

77. Sartori, A. A. *et al.* Human CtIP promotes DNA end resection. *Nature* **450,** 509–514 (2007).

78. Buis, J. *et al.* Mre11 nuclease activity has essential roles in DNA repair and genomic stability distinct from ATM activation. *Cell* **135,** 85–96 (2008).

79. 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 (2015).

80. Song, J. *et al.* RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat. Commun.* **7,** 10548 (2016).

81. Willmore, E. *et al.* A novel DNA-dependent protein kinase inhibitor, NU7026, potentiates the cytotoxicity of topoisomerase II poisons used in the treatment of leukemia. *Blood* **103,** 4659–4665 (2004).

82. Li, Y.-H. *et al.* Inhibition of Non-Homologous End Joining Repair Impairs Pancreatic Cancer Growth and Enhances Radiation Response. *PLOS ONE* **7,** e39588 (2012).

83. Greco, G. E. *et al.* SCR7 is neither a selective nor a potent inhibitor of human DNA ligase IV. *DNA Repair* **43,** 18–23 (2016).

84. Charpentier, M. *et al.* CtIP fusion to Cas9 enhances transgene integration by homology-dependent repair. *Nat. Commun.* **9,** 1133 (2018).

85. Wang, H. *et al.* The interaction of CtIP and Nbs1 connects CDK and ATM to regulate HR-mediated double-strand break repair. *PLoS Genet.* **9,** e1003277 (2013).

86. Wang, H. *et al.* CtIP protein dimerization is critical for its recruitment to chromosomal DNA double-stranded breaks. *J. Biol. Chem.* **287,** 21471–21480 (2012).

87. Davies, O. R. *et al.* CtIP tetramer assembly is required for DNA-end resection and repair. *Nat. Struct. Mol. Biol.* **22,** 150–157 (2015).

88. Hatada, S. *et al.* Low-Dose Irradiation Enhances Gene Targeting in Human Pluripotent Stem Cells. *Stem Cells Transl. Med.* **4,** 998–1010 (2015).

89. Parmee, E. R. *et al.* Discovery of L-755,507: a subnanomolar human beta 3 adrenergic receptor agonist. *Bioorg. Med. Chem. Lett.* **8,** 1107–1112 (1998).

90. Ktistakis, N. T., Linden, M. E. & Roth, M. G. Action of brefeldin A blocked by activation of a pertussis-toxin-sensitive G protein. *Nature* **356,** 344–346 (1992).

91. Liang, X., Potter, J., Kumar, S., Ravinder, N. & Chesnut, J. D. Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA. *J. Biotechnol.* **241,** 136–146 (2017).

92. Li, H. *et al.* Design and specificity of long ssDNA donors for CRISPR-based knock-in. *bioRxiv* (2017). doi:10.1101/178905

93. 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 (2016).

94. 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 (2016).

95. 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 (2017).

96. Su, S. *et al.* CRISPR-Cas9 mediated efficient PD-1 disruption on human primary T cells from cancer patients. *Sci. Rep.* **6,** 20070 (2016).

97. Su, S. *et al.* CRISPR-Cas9-mediated disruption of PD-1 on human T cells for adoptive cellular therapies of EBV positive gastric cancer. *Oncoimmunology* **6,** e1249558 (2017).

98. Rodríguez-Leal, D., Lemmon, Z. H., Man, J., Bartlett, M. E. & Lippman, Z. B. Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing. *Cell* **171,** 470–480.e8 (2017).