Trinucleotide repeat instability during double-strand break repair: from mechanisms to gene therapy

Valentine Mosbach, Lucie Poggi, Guy-Franck Richard

Trinucleotide repeat instability during double-strand break repair: from mechanisms to gene therapy

Valentine Mosbach[1,3], Lucie Poggi[1,2,3,4] and Guy-Franck Richard[1, 3]

[1] Institut Pasteur, Department Genomes & Genetics, 25 rue du Dr Roux, F-75015 Paris, France

[2] Sorbonne Université, Collège Doctoral, 4 Place Jussieu, F-75005 Paris, France

[3] CNRS, UMR3525, F-75015 Paris, France

[4] Biologics Research, Sanofi R&D, 13 Quai Jules Guesde, 94403 Vitry sur Seine, France

Corresponding author: Guy-Franck Richard

Tel: +33-1-45-68-84-36

e-Mail: gfrichar@pasteur.fr

ORCID: 0000-0002-8233-3532

Acknowledgements

The authors wish to thank the continuous support of the Institut Pasteur and of the Centre National de la Recherche Scientifique (CNRS). L. P. is the recipient of a Cifre PhD fellowship from Sanofi. V. M. was the recipient of two post-doctoral fellowships from Fondation Guy Nicolas and from Fondation Hardy.

Abstract

Trinucleotide repeats are a particular class of microsatellites whose large expansions are responsible for at least two dozen human neurological and developmental disorders. Slippage of the two complementary DNA strands during replication, homologous recombination or DNA repair is generally accepted as a mechanism leading to repeat length changes, creating expansions and contractions of the repeat tract. The present review focuses on recent developments on double-strand break repair involving trinucleotide repeat tracts. Experimental evidences in model organisms show that gene conversion and break-induced replication may lead to large repeat tract expansions, while frequent contractions occur either by single-strand annealing between repeat ends or by gene conversion, triggering near-complete contraction of the repeat tract. In the second part of this review, different therapeutic approaches using highly specific single- or double-strand endonucleases targeted to trinucleotide repeat loci are compared. Relative efficacies and specificities of these nucleases will be discussed, as well as their potential strengths and weaknesses for possible future gene therapy of these dramatic disorders.

Keywords

Gene conversion, break-induced replication, single-strand annealing, ZFN, TALEN, CRISPR-Cas9

List of abbreviations used

BIR: break-induced replication

CRISPR: Clustered regularly interspaced short palindromic repeats

SSA: single-strand annealing

ZFN: zinc-finger nucleases

TALEN: transcription activator-like effector nuclease

DSB: double-strand break

SDSA: synthesis-dependent strand annealing

UAS: upstream activating sequence

MRX complex: Mre11-Rad50-Xrs2 complex

PAM: protospacer adjacent motif

iPSC: induced pluripotent stem cells

sgRNA (or gRNA): single-guide RNA

*Sp*Cas9: *Streptococcus pyogenes* Cas9

*Sa*Cas9: *Staphylococcus aureus* Cas9

HNH: homing endonuclease domain

HEK293: human embryonic kidney cell line 293

K562: human immortalized myelogenous leukemia cell line

AAV: adenovirus-associated vector

Introduction

Trinucleotide repeats are a particular class of microsatellites whose large expansions are responsible for at least two dozen human neurological and developmental disorders, discovered over the past 27 years (Fu et al. 1991). Molecular mechanisms responsible for these dramatic large expansions are not totally understood. Yet, experiments in model organisms (mainly bacteria, yeast and mouse) have been fruitful in unraveling some of the key processes underlying trinucleotide repeat instability. These mechanisms involve two features: the ability for these repeats to form stable secondary structures in a test tube (and most probably *in vivo* too; Liu et al. 2010) and the capacity to form DNA heteroduplex (or slipped-strand DNA) by slippage of the newly synthesized strand on the template strand, during DNA synthesis associated with replication, repair or recombination. These features have been extensively described and commented in a number of recent reviews on trinucleotide repeats (Richard et al. 2008; McMurray 2010; Kim and Mirkin 2013; Usdin et al. 2015; Neil Alexander J. et al. 2017; McGinty and Mirkin 2018). Here, we will specifically focus on recent developments involving double-strand breaks as a source of genetic variability for these unstable repeated sequences. The role of gene conversion, break-induced replication (BIR) and single-strand annealing (SSA) in trinucleotide repeat expansions and contractions will be discussed. In addition, several approaches using highly specific DNA endonucleases, such as zinc-finger nucleases (ZFN), TALE nucleases (TALEN) or CRISPR-Cas nucleases were undertaken as possible gene therapies for disorders associated to trinucleotide repeat expansions. Progresses as well as obstacles in each of these different approaches will be discussed.

Double-strand break repair triggers CAG/CTG repeat expansions and contractions by different mechanisms

Some trinucleotide repeats impair replication fork progression, leading to chromosomal fragility and double-strand breaks (DSB), like for example CGG repeats in the fragile X syndrome (Yudkin et al., 2014). Former experiments in yeast showed that some repeats exhibit a length-dependent propensity to break *in vivo* (Callahan et al., 2003; Freudenreich et al., 1998; Jankowski et al., 2000; Kim et al., 2008). In addition, the absence of either *MEC1*, *DDC2* or *RAD53*, which detect DNA damage during replication and transduce the checkpoint response, also led to an increase in chromosomal fragility. However, the strongest increase in fragility was observed when *RAD9*, a checkpoint gene signaling unprocessed DSBs, was deleted (Lahiri et al., 2004). These results suggest that both stalled forks and unrepaired DSBs occur in cells containing long CAG/CTG repeat tracts. Given all these observations, it was therefore legitimate to address the role of DSB-repair in trinucleotide repeat instability.

Gene conversion and BIR lead to CAG/CTG repeat expansions

Initial studies performed almost 20 years ago pointed out the role of gene conversion in CAG/CTG repeat expansions and contractions. The authors used the I-*Sce* I or HO endonucleases, to induce a single DSB into a yeast chromosome. Both nucleases were discovered in the yeast *Saccharomyces cerevisiae*. I-Sce I is a meganuclease encoded by a mitochondrial homing intron (Colleaux et al. 1986) and HO initiates mating type switching by making a double-strand break at the MAT locus (Kostriken et al. 1983). In experimental systems using these nucleases, the induced DSB was repaired using a CAG/CTG repeat-containing homologous template as the donor sequence (Richard et al., 1999, 2000, 2003). Frequent expansions and contractions were observed and suggested

that they occurred through a Synthesis-Dependent Strand Annealing (SDSA) mechanism, a particular type of gene conversion that is never associated to crossover (Figure 1; Richard and Pâques 2000).

Trinucleotide repeat instability may also occur by homologous recombination in the absence of an induced DSB. Such length changes arise from replication fork blocking and/or spontaneous breakage during S phase replication. It was shown that CAG/CTG repeat expansions occurred in a *srs2* yeast mutant, most probably by homologous recombination between sister chromatids (Kerrest et al. 2009). In the absence of the Srs2 helicase activity, recombination intermediates were increased, as visualized by 2D gel electrophoresis. They partly disappeared when *RAD51*, the main recombinase gene in yeast, was deleted, proving that they were *bona fide* recombining molecules (Nguyen et al. 2017).

Expansions were also studied in mice deficient for the RAD52 recombination gene, but no difference in the rate of instability of a $(CTG)_{300}$ repeat tract was found, as compared to control mice (Savouret et al. 2003). However, RAD52 does not play the same role in mammals as it is playing in *S. cerevisiae*. In yeast cells, it is the mediator of all homologous recombination events (SSA, BIR, gene conversion) whereas it is only an accessory recombination gene whose exact function is not totally understood in mammalian cells. Therefore, it would be interesting to address the effect of BRCA1 and/or BRCA2 mutants on CAG/CTG repeat expansions, since these two genes belong to the real recombination mediator complex in human cells (Moynahan et al. 1999, 2001).

Large CAG/CTG repeat expansions were also investigated in yeast using an experimental assay based on the insertion of a $(CTG)_{140}$ repeat tract between the *GAL1* UAS and its TATA box. Transcriptional activation of the downstream reporter no longer occurred if the repeat tract was too long. The average size of detected expansions ranged from 60 to more

than 150 triplets. Expansions decreased in the absence of *RAD51* and *RAD52*, proving that homologous recombination was the key mechanism (Kim et al., 2017). *POL32* (a non-essential DNA polymerase δ subunit) and the *PIF1* helicase were also involved, suggesting that expansions were controlled by BIR (Llorente et al., 2008; Lydeard et al., 2007). A one-ended DSB occurring within the repeat tract could invade the sister chromatid out-of-register, creating a D-loop. BIR would progress until colliding a converging fork or reaching the telomere, eventually resulting in an expansion (Figure 1). Altogether these data tend to show that homologous recombination (gene conversion and BIR) may become a major source of CAG/CTG triplet repeat expansion if not properly controlled.

Gene conversion and SSA lead to CAG/CTG repeat contractions

Initial studies with the I-*Sce* I endonuclease suggested that DSB repair occurred in 67% of the cases by annealing between two short CAG/CTG repeats flanking the I-*Sce* I restriction site (Richard et al., 1999). More recently, a TALE nuclease (TALEN) was used to specifically induce a DSB within a $(CTG)_{80}$ repeat tract integrated in a yeast chromosome. Expression of this nuclease promoted repeat contraction at a high frequency (Mosbach et al., 2018; Richard et al., 2014). Repair was dependent on *RAD50*, *SAE2* and *RAD52*, but did not require *RAD51*, *POL32* or *LIG4*. It was therefore concluded that neither gene conversion nor BIR were the prefered contraction mechanism. It was instead proposed that progressive repeat contractions occurred through iterative cycles of DSB formation followed by SSA (Mosbach et al., 2018). In hamster CHO cells, CAG/CTG repeat contractions were also found to be associated to gene conversion and SSA events, at a frequency (5%) more than 10-fold increased as compared to replicating cells (Meservy et al. 2003).

In conclusion, trinucleotide repeat expansions and contractions appear to occur through different recombination mechanisms (Figure 1). However, it is still unclear whether some of the spontaneous contractions observed during S phase replication in model systems may be triggered by a spontaneous DSB followed by SSA, or are mainly induced by gene conversion associated to DNA slippage.

Role of the SbcCD/MRX complex in CAG/CTG repeat instability

The Mre11-Rad50-Xrs2 (MRX) complex is one of the first players acting at a DSB. The complex triggers end trimming in such a way that resection enzymes -exonucleases and helicases- may be subsequently recruited to produce recombinogenic 3'-hydroxyl single-strand extremities. The Sae2 protein works with the MRX complex in resection initiation, but it is still debated whether Sae2 exhibits a nuclease activity by itself or stimulates Mre11 nuclease activity to initiate resection (Zhu et al. 2008; Mimitou and Symington 2008; Lengsfeld et al. 2007). The MRX complex as well as Sae2 are also required to resolve hairpin-capped natural DSBs in yeast (Lobachev et al., 2002).

Repeat instability following an induced double-strand break

Repair by gene conversion of an HO-induced DSB using a homologous template containing a long CAG/CTG repeat tract led to longer repeat expansions when *MRE11* or *RAD50* were overexpressed (Richard et al., 2000). In addition, it was recently discovered that resection of a TALEN-induced DSB in a $(CTG)_{80}$ tract was completely abolished in the absence of Rad50, and that Sae2 was required to resect the DSB end containing the longest part of the triplet repeat tract (Mosbach et al., 2018). So the MRX complex, along with Sae2, are essential to process a DSB within a CTG trinucleotide repeat, suggesting the presence of secondary structures that need to be removed by the nuclease complex. These results are

strengthened by previous evidences showing the accumulation of unrepaired natural chromosomal breaks within long CTG repeats in the absence of *RAD50* (Freudenreich et al., 1998).

Repeat instability following spontaneous DNA damage

Spontaneous $(CTG)_{70}$ repeat expansions of moderate lengths were increased during S phase in a *mre11*Δ mutant, these expansions being dependent on the *RAD52* gene (Sundararajan et al. 2010). These moderate expansions were very frequent, reaching 8.6% of colonies analyzed. In comparison, large scale $(CTG)_{140}$ repeat expansions were decreased in a *mre11*Δ mutant, from 10^{-5} to 10^{-6} per cell per division. Differences in stability, as well as in the role of Mre11 may reflect differences in mechanisms underlying moderate and large scale CTG repeat expansions: replication-triggered recombination versus BIR. Interestingly, it was recently shown that the MRX complex drove expansions of short $(CTG)_{20}$ trinucleotide repeats (which are not prone to spontaneous breakage) by a process independent of the nuclease function of Mre11 and of the Rad51 recombinase (Ye et al., 2016). This suggests that MRX may promote CTG repeat expansions by recombination-dependent and -independent mechanisms, the relative importance of each during cell life remaining to be determined.

In *Escherichia coli*, it was found that a CAG/CTG repeat tract stimulates the instability of a 275-bp tandem repeat located up to 6.3 kb away (Blackwood et al., 2010). Interestingly, this stimulation required neither DSB-repair nor the hairpin endonuclease SbcCD (homologue of Mre11-Rad50), suggesting that the primary lesion generated at the CAG/CTG repeat was not a DSB. Instead, the authors showed that the mismatch repair machinery triggered the instability observed, probably by recognizing loops of a single triplet formed during replication, leading to the production of single-strand DNA nicks. In

eukaryotes, although its precise role is not totally clear, the mismatch repair machinery appears to be an important player of repeat instability by its propensity to recognize mismatches in hairpins formed by trinucleotide repeats while being unable to repair them (Pearson et al. 1997; Owen et al. 2005; Tomé et al. 2009, 2013; Williams and Surtees 2015; Slean et al. 2016; Viterbo et al. 2016). It is reasonnable to assume that DNA nickases now available will help to study the possible involvement of single stranded DNA nicks on CAG/CTG trinucleotide repeat instability.

GAA/TTC repeat instability occurs by template switching

A genetic assay was designed in yeast to study large-scale expansions of a $(GAA)_{78\text{-}150}$ repeat tract inserted into an artificial intron of the *URA3* gene, larger repeat lengths inhibiting intron splicing, therefore inactivating the gene (Shishkin et al., 2009). Expansions reaching more than 300 triplets were observed, as well as small insertions/deletions or substitutions outside the repeat tract. Large chromosomal deletions including the *URA3* gene and its flanking sequences were also detected. *RAD50* or *RAD52* deletion had no effect on the expansion rate, ruling out the implication of homologous recombination in this process. On the contrary, the absence of replication fork-stabilizing proteins increased the expansion rate while it was decreased in the absence of postreplication DNA repair proteins or the Sgs1 DNA helicase. This strongly suggests that template switching during replication fork progression through GAA repeats was responsible for the observed GAA expansions (Shishkin et al., 2009). More recently, advances in long-read DNA sequencing technologies allowed to identify complex genomic rearrangements originating from improper repair of naturally occurring DSBs at GAA repeats. Various chromosomal rearrangements involving gene conversion between Ty retrotransposons and formation of neochromosomes by BIR were described. These

rearrangements apparently originated from DSBs into the GAA repeat tract (McGinty et al., 2017).

It is worth noting that recombination-independent recognition of DNA homology associated to mutation in *Neurospora crassa* (and probably in *Ascobolus immersus* too) is enhanced by GAC/GTC trinucleotides (Gladyshev and Kleckner 2017). It would be interesting to know if other triplets also interfer with homology recognition and whether such a mechanism could be involved in trinucleotide repeat instability.

In conclusion, although both CAG/CTG and GGA/TTC repeats are apparently able to trigger DSB formation in yeast, expansions involve different sets of genes, therefore different molecular pathways. These differences may be due to: i) distinct secondary structures formed by both types of triplet repeats, GAA tracts folding into triplex DNA whereas CTG repeats form imperfect hairpins; ii) the nature of DNA damage triggered by these structures, double- vs single-strand breaks or gaps; iii) the amount of single-stranded DNA exposed following such damage; iv) differences in chromatin conformation depending on the repeat tract sequence and structure. All these assumptions being not mutually exclusive, understanding the genetic complexity of trinucleotide repeat instability will probably require alternative methods to those applied so far.

Gene editing of trinucleotide repeat expansions

No cure is available for any triplet repeat disorder, although several preclinical and clinical trials have been attempted. Given that microsatellite disorders are always associated to an expansion of the repeat array, deleting or shortening the expanded array to non-pathological lengths should suppress symptoms of the pathology. Indeed, when a trinucleotide repeat contraction occurred during transmission from father to daughter of

an expanded myotonic dystrophy allele, clinical examination of the 17-year old daughter showed no sign of the symptoms (O'Hoy et al. 1993). In another study, a reversible model of DM1 transgenic mice, was relying on a mutant GFP gene under the control of the TetOn promoter, fused to the DMPK 3' UTR. After doxycycline treatment arrest, the GFP-DMPK transgene expression was stopped and sick mice reverted to normal (Mahadevan et al. 2006). Reversible mouse models of Huntington's disease (Yamamoto et al. 2000) and Spinocerebellar Ataxia Type 1 (Zu et al. 2004) showed that suppressing the expression of the toxic mutant protein led to a reversion of severe phenotypes associated to both disorders, including complex motor tasks, even at late disease stages. Hence, gene editing trinucleotide repeat tracts stands as an appealing approach to partially or totally cure these disorders.

Four families of highly specific nucleases may be used to edit trinucleotide repeats: meganucleases, Zinc-Finger Nucleases (ZFN), Transcription Activator Like Effector Nucleases (TALEN) and CRISPR-Cas9. Meganucleases are highly specific DNA endonucleases whose recognition site covers more than 12 bp, originally discovered in group I self-splicing introns in *S. cerevisiae* mitochondria (Dujon 1989). ZFNs were engineered from the fusion of a zinc-finger DNA binding domain to the FokI nuclease domain (Kim et al. 1996). ZFNs are active as heterodimers in which two arms need to dimerize in order to induce a DSB. TALENs are fusion proteins between a TAL effector derived from *Xanthomonas* bacteria and FokI, and also function as heterodimers (Cermak et al. 2011). The Cas9 protein is an RNA-guided nuclease belonging to the CRISPR system of bacterial acquired immune system. It needs the presence of a Protospacer Adjacent Motif (PAM) next to its guide sequence to induce one single-strand break on each DNA strand, resulting in a DSB (Doudna and Charpentier 2014). *Streptococcus pyogenes* Cas9 (*Sp*Cas9) was engineered by an aspartate-to-alanine substitution (D10A) in the RuvC

catalytic domain to convert the double-strand endonuclease into a single-strand nickase (Cong et al, 2013). The same approach was used at the HNH catalytic site to generate the symetrical nickase cutting the opposite DNA strand (N863A). Depending on their bacterial origin, Cas9 proteins recognize different PAM and exhibit different activities. ZFN, TALEN and Cas9 were used to delete or shorten trinucleotide repeats, using two different approaches: i) induce two DSBs upstream and downstream the repeat tract to completely delete it, or ii) induce a DSB inside the repeat tract in order to shorten it (Figure 2).

Huntington's disease

Huntington's disease is a dominant disorder caused by the expansion of a CAG repeat tract in the first exon of the *HTT* gene. In a first study, iPSCs (induced pluripotent stem cells) derived from Huntington patients harboring 72 CAG triplets were electroporated with a modified bacterial artificial chromosome containing 11.5 kb of the genomic region surrounding *HTT* first exon harboring 21 CAG triplets as well as an eGFP reporter cassette and a neomycin resistance gene. Out of 203 analyzed clones, only two showed the incorporation of the wild-type locus by homologous recombination. In these two clones, there was no detectable toxic huntingtin and modified cells retained the modifications when differentiated into neurons (An et al. 2012) (Table1).

In another study, patient derived fibroblasts of variable CAG length were transfected with the D10A nickase and two guide RNAs, each targeting upstream and downstream the CAG repeat tract. Excision of the CAG repeat in the transfected non-clonal population showed decreased levels of the *HTT* mRNA and protein, from 68% to 82% depending on the cell line, suggesting that at least one allele was efficiently deleted, on the average. Four out of 13 predicted exonic off-target sites were tested and no mutation was detected

(Dabrowska et al. 2018).

Alternative approaches exploited the presence of SNPs specific of the mutant CAG expanded allele. Two studies analyzing *HTT* haplotype were recently published, in which the authors took advantage of specific SNPs to remove the expanded allele in HD fibroblasts (Shin et al. 2016; Monteys et al. 2017). One of the studies also demonstrated that sgRNA/Cas9 complexes are also effective *in vivo* in an HD mouse model harboring the HD human allele. Viral delivery of sgRNA/*Sp*Cas9 complexes reduced human mutant HTT expression to 40% in the treated hemisphere as compared to the control untreated one (Monteys et al. 2017).

Myotonic dystrophy type I (DM1 or Steinert disease)

DM1 is an autosomal dominant disorder caused by an RNA-gain of function mutation: the expanded CTG repeat tract located at the 3'UTR of the DMPK gene is translated into a CUG-expanded RNA which accumulates into the nucleus and forms aggregates with splicing-effector proteins such as MBNL1 and CUG-BP1 (Miller et al. 2000). Deleting the CTG repeat tract should result in the suppression of the toxic RNA. The first work introducing the use of a highly specific nuclease to shorten a long CTG repeat from a DM1 patient, reported that a DSB made by a TALEN into the repeat tract induced a contraction of the repeat in 99% of cases, in yeast cells (Richard et al. 2014). In another study, a reporter assay was built in HEK293 cells to monitor contractions and expansions of a CTG repeat tract integrated into a synthetic intron interrupting a GFP gene. Efficacy of Cas9 D10A nickase, wild-type Cas9 and ZFNs cutting into the CTG repeat tract were compared. All induced contractions and expansions of the CTG repeat, but the nickase was the most efficient at inducing contractions (Cinesi et al. 2016).

Two proofs of concept of the removal of CTG repeats to cure DM1 were subsequently

established. The introduction of Cas9 and a pair of guide RNAs each targeting a specific locus upstream and downstream the DM1 repeats in patient cells resulted in the deletion of the CTG repeats, the suppression of RNA foci and splicing defects (Van Agtmaal et al. 2017; Provenzano et al. 2017). Those two studies used different cell types, respectively myogenic DM1 myoblast and DM1 fibroblasts and different target loci and achieved respectively 46% and 14% of successfully edited cells. Indels were found in both cases at cut sites and few loci were tested for off-target effects.

One last strategy consisted in inserting a polyA signal upstream the CTG tract to prevent its transcription. This was carried out by making a DSB between exon 9 and 10 of the DMPK gene, induced by a TALEN, while co-transfecting the polyA cassette (Xia et al. 2015). Successfully edited cells showed phenotype reversion including foci disappearance and normal splicing of MBNL1 and MBNL2.

Fragile X syndrome

The fragile X syndrome is caused by the expansion of a CGG repeat tract in the 5' UTR of the *FMR1* gene which leads through an undetermined mechanism to the methylation of the *FMR1* promoter. FXS iPSCs (more than 450 CGG) were transfected with *Sp*Cas9 and a guideRNA targeting the region upstream the repeat tract (Park et al. 2015). Four potential off target sites were tested and no mutation was detected. Two successfully edited clones over 100 tested were obtained. In these two clones, promoter hypermethylation was abolished and *FMR1* expression was reactivated. A similar study was conducted by cutting upstream and downstream the CGG repeats using *Sp*Cas9. The authors observed a decrease in the methylation profile of the *FMR1* promoter in one of their analysed clones along with partial restoration of the FMR1 protein (Xie et al. 2016).

Friedreich's ataxia (FRDA)

FRDA is a recessive disorder caused by an expanded GAA (up to 2000 triplets) located in intron 1 of the frataxin gene, inducing a heterochromatization of the *FXN* locus leading to low frataxin levels (Campuzano et al. 1996). Heterozygous carriers are asymptomatic. Two ZFNs were designed to specifically cut upstream and downstream the GAA repeat tract. FRDA lymphoblasts and fibroblasts were transfected with both ZFN arms. Successful edition was achieved for 7 out of 305 lymphoblasts (2,3% efficiency) and 23 out of 344 fibroblasts (6.7% efficiency). Heterozygous modifications were observed as well as large deletions at ZFN cut sites. Edited cells exhibited increased expression of frataxin. When differentiated into neurons the cells retained the corrections. Ten top off-target sites were studied in established cell line K562 cells and no mutation was detected (Li et al. 2015). *Sp*Cas9 was targeted in transgenic mice fibroblasts and whole animal muscles, upstream and downstream GAA repeats in order to remove them (Ouellet et al. 2017). Successful *in vitro* edition ranged from 4% to 15% depending on the couple of gRNA used. Indels were found at sequenced junctions in successfully edited clones. Gene editing events were observed by PCR in fibroblasts, as well as *in vivo*. *Sa*Cas9 was also transfected in mice fibroblasts but its expression level was much lower than *Sp*Cas9 and editing was not very efficient.

Limitations of nuclease approaches: off-target effects

One major concern about specific nucleases is the potential effect of off-target mutations due to a lack of specificity. *In silico* programs are poor predictors of real off-target sites and there is no simple rule so far to accurately predict off-targets. The first genome-wide assessment of Cas9 off-target sites was carried out using the GUIDE-seq method. Briefly, double-stranded modified oligonucleotides are transfected alongside the nuclease and

integrate in the genome at all DSB sites generated by the nuclease. They can subsequently be amplified and serve as primers for genome-wide sequencing of their insertion sites. This analysis revealed that off-targets are difficult to predict, ranging from little cleavage outside the target to as many off-target as on-target DSBs, depending on the gRNA chosen. Cleavage can occur on sites bearing up to seven mismatches and no canonical PAM (Tsai et al. 2015). CIRCLE-seq is a simpler and more sensitive method to detect off-target sites *in vitro*, but requires the purified nuclease (Tsai et al. 2017). Using this approach, genomic DNA that was cleaved by the nuclease in a test tube was amplified and sequenced. This method is very sensitive but may not be relevant for *in vivo* assays and may depend on each cell type and chromatin state. Recently, the VIVO method was set up for *in vivo* validation of off-target sites found by CIRCLE-seq, demonstrating that careful choice of the gRNA may strongly reduce off-target effects, while keeping a good on-target efficacy (Akcakaya et al. 2018). The same team engineered a more specific version of *Sp*Cas9, called HF1, by mutating residues involved in the binding to the target DNA strand. Cas9-HF1 retains on-target activity comparable to wild-type on 85% of gRNAs tested and rendered all or nearly all off-target events not detectable by GUIDE-seq (Kleinstiver et al. 2016). No such extensive off-target study was carried out in any of the aforementioned articles. Such approaches must be encouraged in future assessments of gene therapy strategies for trinucleotide repeat disorders.

Limitations of nuclease approaches: vectorization

Nuclease vectorization is clearly a problem that also needs to be addressed. Adenovirus-associated vectors (AAV) are popular in gene therapy because they exhibit low integration frequency, but they have a limited cargo capacity making it impossible to deliver a full length *Sp*Cas9 with its cognate guide, or a TALEN. In this case, each of the

two TALEN arms must be delivered by two different vectors, lowering the efficacy of the transduction. Alternative non-viral delivery systems such as cationic lipid transfection particles was efficient to deliver a Cas9-gRNA complex as well as a TALEN both *in vitro* and *in vivo*, achieving 20% efficacy in genome modification in mice (Zuris et al. 2015). AAV-based delivery could also potentially increase the rate of off-target site cleavage due to prolonged expression of the nuclease. To circumvent this problem, a self-limiting CRISPR-Cas9 system was implemented *in vivo* by inserting the sequence recognized by the nuclease on the plasmid encoding it such that the expression plasmid would be cut and eliminated following *Sp*Cas9 expression (Ruan et al. 2017).

An alternative approach would solve the vectorization as well as the immune response issues: *in vitro* modification of patient induced pluripotent stem cells, followed by reprogrammation of nuclease-treated iPSC into the desired cell type (neuron, myoblast, etc.). However, such an advance in regenerative medicine is still hampered by the need for expressing four transcription factors from retroviral vectors in order to induce pluripotency, with all the risks associated to retrovirus integration into human cells (Takahashi et al. 2007; Yu et al. 2007).

Conclusion

Little is known yet about the immune response toward these nucleases. A very recent work identified pre-existing immunity against Cas9 from *Streptococcus pyogenes* and *Staphylococcus aureus* (Charlesworth *et al,* 2018). The authors showed that 70% of healthy adults have antibodies directed to the nuclease and that *Sa*Cas9 induced a T-cell response in adult blood. A strong immune response may be a potential drawback to the use of Cas9 in future gene therapy.

An additional difficulty is raised by checkpoint effectors, such as p53, controlling the

cellular response to double-strand breaks. Two studies have recently shown that during gene editing, cells with a functional p53 pathway were counterselected, due to cell arrest triggered by p53 upon DSB formation. Therefore, checkpoint activity should be tightly controlled when developing cell-based therapies utilizing CRISPR–Cas9 (Haapaniemi et al. 2018; Ihry et al. 2018).

These first reports of gene therapy attempts of trinucleotide repeat disorders are certainly promising and already give us insights into crucial factors to be considered when evaluating the success of a gene therapy approach: off-target sites number and frequency, nuclease efficacy, cell type to be targeted and vectorization method. Successful gene editing was achieved in a mouse model for Duchenne muscular dystrophy, by three independent teams. Using AAV delivery of Cas9, they obtained partial restoration of dystrophin levels that were sufficient to allow partial muscle strength recovery (Long et al. 2016; Nelson et al. 2016; Tabebordbar et al. 2016). Forthcoming experiments in a mouse model for trinucleotide repeat disorders will establish if a similar success may be achieved.

Reference list

Agtmaal EL van, André LM, Willemse M, et al (2017) CRISPR/Cas9-Induced (CTG·CAG)n Repeat Instability in the Myotonic Dystrophy Type 1 Locus: Implications for Therapeutic Genome Editing. Molecular Therapy 25:24–43. doi: 10.1016/j.ymthe.2016.10.014

Akcakaya P, Bobbin ML, Guo JA, et al (2018) In vivo CRISPR-Cas gene editing with no detectable genome-wide off-target mutations. bioRxiv 272724. doi: 10.1101/272724

An MC, Zhang N, Scott G, et al (2012) Genetic Correction of Huntington's Disease Phenotypes in Induced Pluripotent Stem Cells. Cell Stem Cell 11:253–263. doi: 10.1016/j.stem.2012.04.026

Blackwood JK, Okely EA, Zahra R, et al (2010) DNA tandem repeat instability in the Escherichia coli chromosome is stimulated by mismatch repair at an adjacent CAG·CTG trinucleotide repeat. Proc Natl Acad Sci USA 107:22582–22586. doi: 10.1073/pnas.1012906108

Callahan JL, Andrews KJ, Zakian VA, Freudenreich CH (2003) Mutations in yeast replication proteins that increase CAG/CTG expansions also increase repeat fragility. Mol Cell Biol 23:7849–7860

Campuzano V, Montermini L, Molto MD, et al (1996) Friedreich's Ataxia: autosomal

recessive disease caused by an intronic GAA triplet repeat expansion. Science 271:1423–1427
Cermak T, Doyle EL, Christian M, et al (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39:e82. doi: 10.1093/nar/gkr218
Cinesi C, Aeschbach L, Yang B, Dion V (2016) Contracting CAG/CTG repeats using the CRISPR-Cas9 nickase. Nature Communications 7:13272. doi: 10.1038/ncomms13272
Colleaux L, d'Auriol L, Betermier M, et al (1986) Universal code equivalent of a yeast mitochondrial intron reading frame is expressed into E. Coli as a specific double strand break endonuclease. Cell 44:521–533
Dabrowska M, Juzwa W, Krzyzosiak WJ, Olejniczak M (2018) Precise Excision of the CAG Tract from the Huntingtin Gene by Cas9 Nickases. Front Neurosci 12:. doi: 10.3389/fnins.2018.00075
Doudna JA, Charpentier E (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 346:1258096. doi: 10.1126/science.1258096
Dujon B (1989) Group I introns as mobile genetic elements: facts and mechanistic speculations--a review. Gene 82:91–114
Freudenreich CH, Kantrow SM, Zakian VA (1998) Expansion and length-dependent fragility of CTG repeats in yeast. Science 279:853–856
Fu Y-H, Kuhl DPA, Pizzuti A, et al (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell 67:1047–1058
Gladyshev E, Kleckner N (2017) Recombination-independent recognition of DNA homology for repeat-induced point mutation. Curr Genet 63:389–400. doi: 10.1007/s00294-016-0649-4
Haapaniemi E, Botla S, Persson J, et al (2018) CRISPR–Cas9 genome editing induces a p53-mediated DNA damage response. Nature Medicine 1. doi: 10.1038/s41591-018-0049-z
Ihry RJ, Worringer KA, Salick MR, et al (2018) p53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells. Nature Medicine 1. doi: 10.1038/s41591-018-0050-6
Jankowski C, Nasar F, Nag DK (2000) Meiotic instability of CAG repeat tracts occurs by double-strand break repair in yeast. Proc Natl Acad Sci USA 97:2134–2139. doi: 10.1073/pnas.040460297
Kerrest A, Anand R, Sundararajan R, et al (2009) SRS2 and SGS1 prevent chromosomal breaks and stabilize triplet repeats by restraining recombination. Nature Structural and Molecular Biology 16:159–167
Kim H-M, Narayanan V, Mieczkowski PA, et al (2008) Chromosome fragility at GAA tracts in yeast depends on repeat orientation and requires mismatch repair. EMBO J 27:2896–2906. doi: 10.1038/emboj.2008.205
Kim JC, Harris ST, Dinter T, et al (2017) The role of break-induced replication in large-scale expansions of (CAG)n/(CTG)nrepeats. Nat Struct Mol Biol 24:55–60. doi: 10.1038/nsmb.3334
Kim JC, Mirkin SM (2013) The balancing act of DNA repeat expansions. Current Opinion in Genetics & Development 23:280–288. doi: 10.1016/j.gde.2013.04.009
Kim YG, Cha J, Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A 93:1156–60
Kleinstiver BP, Pattanayak V, Prew MS, et al (2016) High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. Nature 529:490–495. doi: 10.1038/nature16526
Kostriken R, Strathern JN, Klar AJ, et al (1983) A site-specific endonuclease essential for mating-type switching in Saccharomyces cerevisiae. Cell 35:167–74. doi: 0092-8674(83)90219-2 [pii]

Lahiri M, Gustafson TL, Majors ER, Freudenreich CH (2004) Expanded CAG repeats activate the DNA damage checkpoint pathway. Mol Cell 15:287–293. doi: 10.1016/j.molcel.2004.06.034

Lengsfeld BM, Rattray AJ, Bhaskara V, et al (2007) Sae2 Is an Endonuclease that Processes Hairpin DNA Cooperatively with the Mre11/Rad50/Xrs2 Complex. Molecular Cell 28:638–651. doi: 10.1016/j.molcel.2007.11.001

Li Y, Polak U, Bhalla AD, et al (2015) Excision of Expanded GAA Repeats Alleviates the Molecular Phenotype of Friedreich's Ataxia. Mol Ther 23:1055–1065. doi: 10.1038/mt.2015.41

Liu G, Chen X, Bissler JJ, et al (2010) Replication-dependent instability at (CTG) x (CAG) repeat hairpins in human cells. Nat Chem Biol 6:652–9. doi: 10.1038/nchembio.416

Llorente B, Smith CE, Symington LS (2008) Break-induced replication: what is it and what is it for? Cell Cycle 7:859–864. doi: 10.4161/cc.7.7.5613

Long C, Amoasii L, Mireault AA, et al (2016) Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. Science 351:400–403. doi: 10.1126/science.aad5725

Lydeard JR, Jain S, Yamaguchi M, Haber JE (2007) Break-induced replication and telomerase-independent telomere maintenance require Pol32. Nature 448:820–823. doi: 10.1038/nature06047

Mahadevan MS, Yadava RS, Yu Q, et al (2006) Reversible model of RNA toxicity and cardiac conduction defects in myotonic dystrophy. Nat Genet 38:1066–1070. doi: 10.1038/ng1857

McGinty RJ, Mirkin SM (2018) Cis- and Trans-Modifiers of Repeat Expansions: Blending Model Systems with Human Genetics. Trends Genet 34:448–465. doi: 10.1016/j.tig.2018.02.005

McGinty RJ, Rubinstein RG, Neil AJ, et al (2017) Nanopore sequencing of complex genomic rearrangements in yeast reveals mechanisms of repeat-mediated double-strand break repair. Genome Res 27:2072–2082. doi: 10.1101/gr.228148.117

McMurray CT (2010) Mechanisms of trinucleotide repeat instability during human development. Nat Rev Genet 11:786–99. doi: 10.1038/nrg2828

Meservy JL, Sargent RG, Iyer RR, et al (2003) Long CTG Tracts from the Myotonic Dystrophy Gene Induce Deletions and Rearrangements during Recombination at the APRT Locus in CHO Cells. Mol Cell Biol 23:3152–3162. doi: 10.1128/MCB.23.9.3152-3162.2003

Miller JW, Urbinati CR, Teng-umnuay P, et al (2000) Recruitment of human musclebind proteins to (CUG)n expansions associated with myotonic dystrophy. EMBO J 19:4439–4448

Monteys AM, Ebanks SA, Keiser MS, Davidson BL (2017) CRISPR/Cas9 Editing of the Mutant Huntingtin Allele In Vitro and In Vivo. Mol Ther 25:12–23. doi: 10.1016/j.ymthe.2016.11.010

Mosbach V, Poggi L, Viterbo D, et al (2018) TALEN-Induced Double-Strand Break Repair of CTG Trinucleotide Repeats. Cell Rep 22:2146–2159. doi: 10.1016/j.celrep.2018.01.083

Moynahan ME, Chiu JW, Koller BH, Jasin M (1999) Brca1 controls homology-directed DNA repair. Mol Cell 4:511–8

Moynahan ME, Pierce AJ, Jasin M (2001) BRCA2 is required for homology-directed repair of chromosomal breaks. Mol Cell 7:263–72

Neil Alexander J., Kim Jane C., Mirkin Sergei M. (2017) Precarious maintenance of simple DNA repeats in eukaryotes. BioEssays 39:1700077. doi: 10.1002/bies.201700077

Nelson CE, Hakim CH, Ousterout DG, et al (2016) In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. Science 351:403–407. doi:

10.1126/science.aad5143
Nguyen JHG, Viterbo D, Anand RP, et al (2017) Differential requirement of Srs2 helicase and Rad51 displacement activities in replication of hairpin-forming CAG/CTG repeats. Nucleic Acids Res 45:4519–4531. doi: 10.1093/nar/gkx088
O'Hoy KL, Tsilfidis C, Mahadevan MS, et al (1993) Reduction in size of the myotonic dystrophy trinucleotide repeat mutation during transmission. Science 259:809–12
Ouellet DL, Cherif K, Rousseau J, Tremblay JP (2017) Deletion of the GAA repeats from the human frataxin gene using the CRISPR-Cas9 system in YG8R-derived cells and mouse models of Friedreich ataxia. Gene Ther 24:265–274. doi: 10.1038/gt.2016.89
Owen BAL, Yang Z, Lai M, et al (2005) (CAG)(n)-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. Nat Struct Mol Biol 12:663–670. doi: 10.1038/nsmb965
Park C-Y, Halevy T, Lee DR, et al (2015) Reversion of FMR1 Methylation and Silencing by Editing the Triplet Repeats in Fragile X iPSC-Derived Neurons. Cell Rep 13:234–241. doi: 10.1016/j.celrep.2015.08.084
Pearson CE, Ewel A, Acharya S, et al (1997) Human MSH2 binds to trinucleotide repeat DNA structures associated with neurodegenerative diseases. Hum Mol Genet 6:1117–1123
Provenzano C, Cappella M, Valaperta R, et al (2017) CRISPR/Cas9-Mediated Deletion of CTG Expansions Recovers Normal Phenotype in Myogenic Cells Derived from Myotonic Dystrophy 1 Patients. Mol Ther Nucleic Acids 9:337–348. doi: 10.1016/j.omtn.2017.10.006
Richard GF, Cyncynatus C, Dujon B (2003) Contractions and expansions of CAG/CTG trinucleotide repeats occur during ectopic gene conversion in yeast, by a MUS81-independent mechanism. J Mol Biol 326:769–782
Richard GF, Dujon B, Haber JE (1999) Double-strand break repair can lead to high frequencies of deletions within short CAG/CTG trinucleotide repeats. Mol Gen Genet 261:871–882
Richard GF, Goellner GM, McMurray CT, Haber JE (2000) Recombination-induced CAG trinucleotide repeat expansions in yeast involve the MRE11-RAD50-XRS2 complex. EMBO J 19:2381–2390. doi: 10.1093/emboj/19.10.2381
Richard GF, Kerrest A, Dujon B (2008) Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. Microbiol Mol Biol Rev 72:686–727
Richard G-F, Pâques F (2000) Mini- and microsatellite expansions: the recombination connection. EMBO Reports 1:122–126
Richard G-F, Viterbo D, Khanna V, et al (2014) Highly specific contractions of a single CAG/CTG trinucleotide repeat by TALEN in yeast. PLoS ONE 9:e95611. doi: 10.1371/journal.pone.0095611
Ruan G-X, Barry E, Yu D, et al (2017) CRISPR/Cas9-Mediated Genome Editing as a Therapeutic Approach for Leber Congenital Amaurosis 10. Mol Ther 25:331–341. doi: 10.1016/j.ymthe.2016.12.006
Savouret C, Brisson E, Essers J, et al (2003) CTG repeat instability and size variation timing in DNA repair-deficient mice. Embo J 22:2264–73
Shin JW, Kim K-H, Chao MJ, et al (2016) Permanent inactivation of Huntington's disease mutation by personalized allele-specific CRISPR/Cas9. Hum Mol Genet 25:4566–4576. doi: 10.1093/hmg/ddw286
Shishkin AA, Voineagu I, Matera R, et al (2009) Large-scale expansions of Friedreich's ataxia GAA repeats in yeast. Mol Cell 35:82–92. doi: 10.1016/j.molcel.2009.06.017
Slean MM, Panigrahi GB, Castel AL, et al (2016) Absence of MutSβ leads to the formation of slipped-DNA for CTG/CAG contractions at primate replication forks. DNA Repair (Amst) 42:107–118. doi: 10.1016/j.dnarep.2016.04.002

Sundararajan R, Gellon L, Zunder RM, Freudenreich CH (2010) Double-strand break repair pathways protect against CAG/CTG repeat expansions, contractions and repeat-mediated chromosomal fragility in Saccharomyces cerevisiae. Genetics 184:65–77. doi: 10.1534/genetics.109.111039

Tabebordbar M, Zhu K, Cheng JKW, et al (2016) In vivo gene editing in dystrophic mouse muscle and muscle stem cells. Science 351:407–411. doi: 10.1126/science.aad5177

Takahashi K, Tanabe K, Ohnuki M, et al (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131:861–872. doi: 10.1016/j.cell.2007.11.019

Tomé S, Holt I, Edelmann W, et al (2009) MSH2 ATPase domain mutation affects CTG*CAG repeat instability in transgenic mice. PLoS Genet 5:e1000482. doi: 10.1371/journal.pgen.1000482

Tomé S, Manley K, Simard JP, et al (2013) MSH3 polymorphisms and protein levels affect CAG repeat instability in Huntington's disease mice. PLoS Genet 9:e1003280. doi: 10.1371/journal.pgen.1003280

Tsai SQ, Nguyen NT, Malagon-Lopez J, et al (2017) CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. Nat Meth 14:607–614. doi: 10.1038/nmeth.4278

Tsai SQ, Zheng Z, Nguyen NT, et al (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat Biotechnol 33:187–97. doi: 10.1038/nbt.3117

Usdin K, House NC, Freudenreich CH (2015) Repeat instability during DNA repair: Insights from model systems. Critical Reviews in Biochemistry and Molecular Biology doi:10.3109/10409238.2014.999192. doi: 10.3109/10409238.2014.999192

Verkerk AJMH, Pieretti M, Sutcliffe JS, et al (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. Cell 65:905–914

Viterbo D, Michoud G, Mosbach V, et al (2016) Replication stalling and heteroduplex formation within CAG/CTG trinucleotide repeats by mismatch repair. DNA Repair (Amst) 42:94–106. doi: 10.1016/j.dnarep.2016.03.002

Williams GM, Surtees JA (2015) MSH3 Promotes Dynamic Behavior of Trinucleotide Repeat Tracts In Vivo. Genetics 200:737–754. doi: 10.1534/genetics.115.177303

Xia G, Gao Y, Jin S, et al (2015) Genome modification leads to phenotype reversal in human myotonic dystrophy type 1 induced pluripotent stem cell-derived neural stem cells. Stem Cells 33:1829–38. doi: 10.1002/stem.1970

Xie N, Gong H, Suhl JA, et al (2016) Reactivation of FMR1 by CRISPR/Cas9-Mediated Deletion of the Expanded CGG-Repeat of the Fragile X Chromosome. PLOS ONE 11:e0165499. doi: 10.1371/journal.pone.0165499

Yamamoto A, Lucas JJ, Hen R (2000) Reversal of neuropathology and motor dysfunction in a conditional model of Huntington's disease. Cell 101:57–66. doi: 10.1016/S0092-8674(00)80623-6

Ye Y, Kirkham-McCarthy L, Lahue RS (2016) The Saccharomyces cerevisiae Mre11-Rad50-Xrs2 complex promotes trinucleotide repeat expansions independently of homologous recombination. DNA Repair (Amst) 43:1–8. doi: 10.1016/j.dnarep.2016.04.012

Yu J, Vodyanik MA, Smuga-Otto K, et al (2007) Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. Science 318:1917–1920. doi: 10.1126/science.1151526

Yu S, Pritchard M, Kremer E, et al (1991) Fragile X genotype characterized by an unstable region of DNA. Science 252:1179–81

Yudkin D, Hayward BE, Aladjem MI, et al (2014) Chromosome fragility and the abnormal replication of the FMR1 locus in fragile X syndrome. Hum Mol Genet 23:2940–2952. doi:

10.1093/hmg/ddu006
Zu T, Duvick LA, Kaytor MD, et al (2004) Recovery from polyglutamine-induced neurodegeneration in conditional SCA1 transgenic mice. J Neurosci 24:8853–8861. doi: 10.1523/JNEUROSCI.2978-04.2004
Zuris JA, Thompson DB, Shu Y, et al (2015) Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. Nat Biotechnol 33:73–80. doi: 10.1038/nbt.3081

Figure 1: Double-strand break repair mechanisms leading to repeat contraction or expansion.

After a DSB was made into (or close to) a trinucleotide repeat tract, the broken molecule is resected by several nucleases and helicases leading to 3'-hydroxyl single-stranded ends. These ends may engage into different types of homologous recombination. Direct annealing of the two ends by SSA leads to repeat tract contraction after flap clipping (right). DNA synthesis during BIR generates repeat expansions (bottom). Synthesis-dependent strand annealing is resolved by unwinding and out-of-frame annealing of the recombination intermediate, possibly leading to repeat expansion (A) or repeat contraction (B). Note that none of these mechanisms requires crossover formation or resolution.

Figure 2: Methods used for deleting or contracting trinucleotide repeats in human cells. Expanded trinucleotide repeat tracts were targeted by different nucleases in four human disorders. In each case, one or more approach was used to contract or delete the repeat tract. The nuclease expressed is shown in gray, along with arrows indicating whether the DSB (or SSB) was made within or outside the repeat tract. Repair outcomes following homologous recombination or non-homologous end joining are drawn. Corresponding references are shown under each approach.

Figure 1

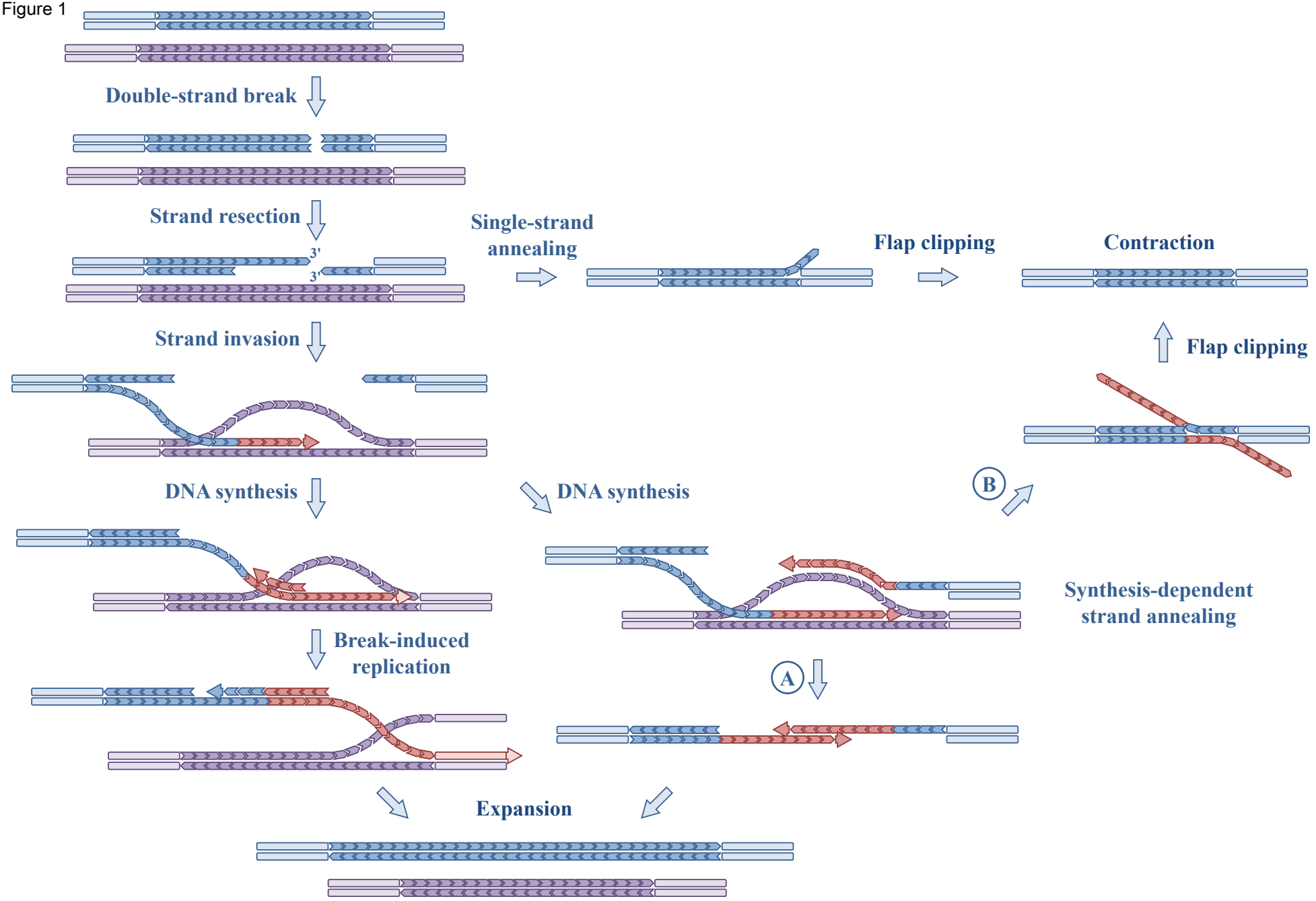
Double-strand break
Strand resection
3'
3'
Strand invasion
DNA synthesis
Break-induced replication
Expansion
Single-strand annealing
Flap clipping
Contraction
Flap clipping
B
DNA synthesis
Synthesis-dependent strand annealing
A

Figure 2

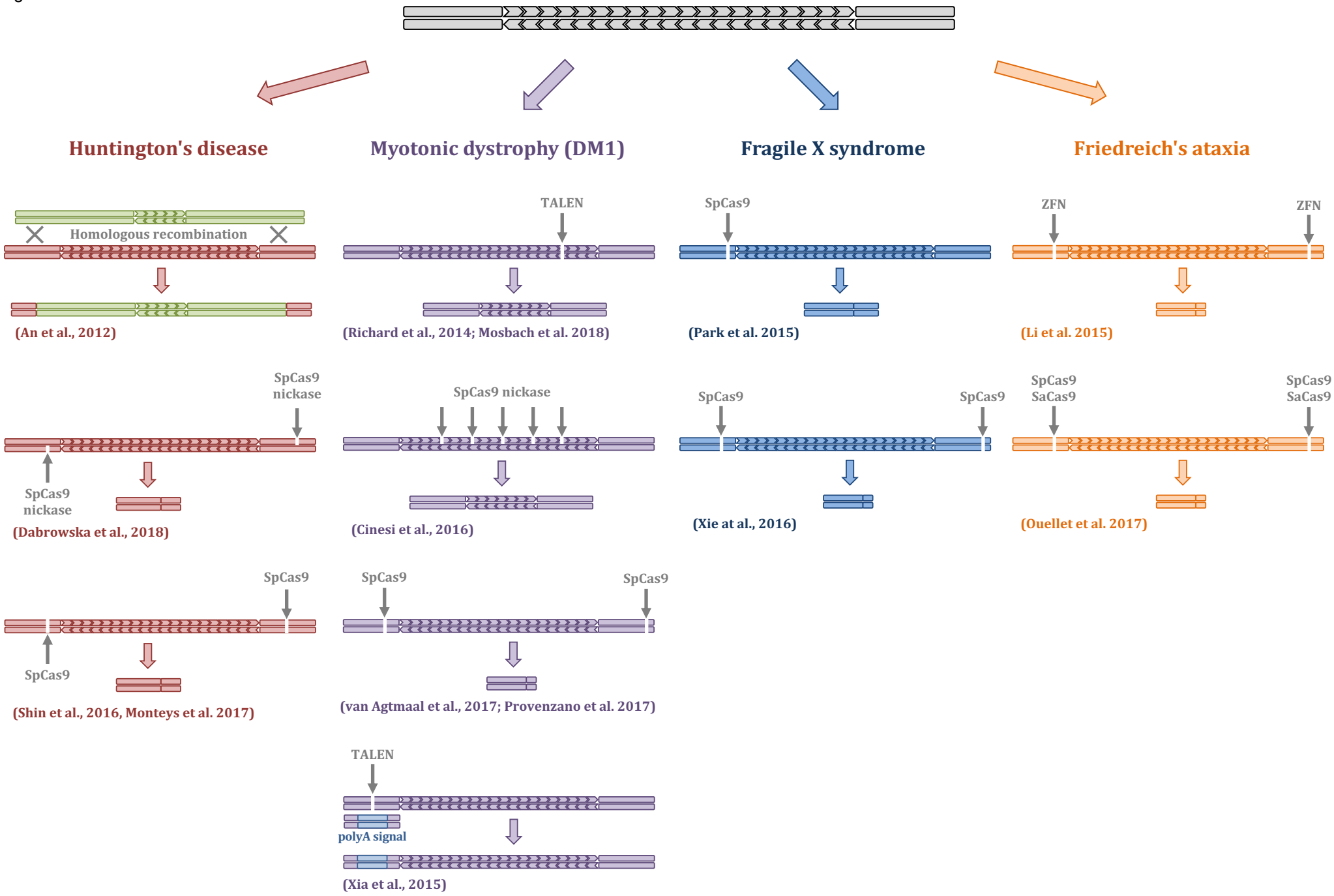
Huntington's disease
Myotonic dystrophy (DM1)
Fragile X syndrome
Friedreich's ataxia
Homologous recombination
(An et al., 2012)
TALEN
(Richard et al., 2014; Mosbach et al. 2018)
SpCas9
(Park et al. 2015)
ZFN
ZFN
(Li et al. 2015)
SpCas9 nickase
SpCas9 nickase
(Dabrowska et al., 2018)
SpCas9 nickase
(Cinesi et al., 2016)
SpCas9
SpCas9
(Xie at al., 2016)
SpCas9
SaCas9
SpCas9
SaCas9
(Ouellet et al. 2017)
SpCas9
SpCas9
(Shin et al., 2016, Monteys et al. 2017)
SpCas9
SpCas9
(van Agtmaal et al., 2017; Provenzano et al. 2017)
TALEN
polyA signal
(Xia et al., 2015)

Table 1

Disease	Huntington's disease			
Reference	An *et al,* 2012	Dabrowska *et al,* 2018	Shin *et al,* 2016	Monteys *et al,* 2017
Cell type	HD iPS cells	HD fibroblasts	HD fibroblasts	HD fibroblasts BacHD mice
Nuclease used	None (spontaneous homologous recombination)	Paired D10A nickases	*Sp*Cas9	*Sp*Cas9
Successful edition	1% (203 clones analysed)	NA (bulk analysis)	NA (bulk analysis)	NA (bulk analysis)
Off-target analysis	NA	Indels at cut site 4 off target sites analysed: unchanged	None	11 top off target sites: unchanged
Phenotype of edited cells	No detectable toxic huntingtin	No detectable toxic huntingtin	HTT mRNA and protein levels decreased	HTT mRNA and protein levels decreased

Disease	Myotonic Dystrophy type I				
Reference	Richard *et al.*, 2014	Cinesi *et al,* 2016	Provenzano *et al,* 2017	Van Atgmaal *et al,* 2017	Xia *et al,* 2015
Cell type	*Saccharomyces cerevisiae*	HEK293 GFP(CAG)$_{89}$	Immortalized myogenic DM1 fibroblast	Immortalized DM1 myoblast (DM11)	DM1 neural stem cells
Nuclease	TALEN	ZFN *Sp*Cas9 D10A Cas9 nickase	e*Sp*Cas9	*Sp*Cas9	TALEN
Successful edition	99%	3%	14% (85 clones analysed)	46% (103 clones analysed)	After selection: 4 out of 10 colonies
Off-target analysis	Whole genome sequencing: no change	Number of CTG repeats at 7 different loci remained unchanged	Indels (1-151 bp) observed at cut sites Sequencing of the top 7 off-target sites of each sgRNA: unchanged	Indels at cut site Sequencing of the top 4 off-target loci: unchanged on model cell lines	Indels at cut site (40% cases)
Phenotype of edited cells	NA	NA	No foci Normal splicing of SERCA1 and INSR	No foci No MBNL1 aggregate Normal splicing of BIN1 and DMD	No foci Normal splicing of MBNL1&2 and MAPT

Disease	Fragile X syndrome		Friedreich's ataxia	
Reference	Park *et al,* 2015	Xie *et al,* 2016	Li *et al,* 2015	Ouellet *et al,* 2017
Cell type	FXS iPS cells	FXS iPS cells	FRDA fibroblasts and lymphoblasts	Transgenic mouse fibroblasts and whole animal muscles
Nuclease	*Sp*Cas9	*Sp*Cas9	ZFN	*Sp*Cas9 and *Sa*Cas9
Successful edition	2% (100 clones analysed)	5 clones analyzed	6.7% (344 fibroblasts analysed) 2.3% (305 lymphoblasts analyzed)	15% for the best gRNA combination in fibroblasts (33 clones analysed) No quantification *in vivo*
Off-target analysis	49 and 112 bp deletion at cut site Sequencing of the 4 top off-target loci: unchanged on model cell lines	Indels at cut site	Indels at cut site. Ten off target analysed: unchanged	Indels at cut site. No off-target study
Phenotype of edited cells	Decrease of FMR1 promoter methylation FMR1 mRNA and protein levels restored	Decrease of FMR1 promoter methylation for one clone. FMR1 mRNA and protein levels restored	FXN mRNA and protein levels restored. Neural cells showed restored levels of aconitase	Depending on the deletion event, FXN protein level was sometimes increased.

NA: Not Applicable

Table 1: Comparison between 12 gene editing studies aimed at correcting trinucleotide repeat disorders