



**HAL**  
open science

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

Valentine Mosbach, Lucie Poggi, Guy-Franck Richard

### ► To cite this version:

Valentine Mosbach, Lucie Poggi, Guy-Franck Richard. Trinucleotide repeat instability during double-strand break repair: from mechanisms to gene therapy. *Current Genetics*, 2019, 65, pp.17 - 28. 10.1007/s00294-018-0865-1 . pasteur-02864590

**HAL Id: pasteur-02864590**

**<https://pasteur.hal.science/pasteur-02864590>**

Submitted on 11 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1  
2  
3 **Trinucleotide repeat instability during double-strand break repair:**  
4  
5 **from mechanisms to gene therapy**  
6  
7  
8  
9

10  
11  
12  
13 Valentine Mosbach<sup>1,3</sup>, Lucie Poggi<sup>1,2,3,4</sup> and Guy-Franck Richard<sup>1,3</sup>  
14  
15

16  
17  
18  
19  
20  
21 <sup>1</sup> Institut Pasteur, Department Genomes & Genetics, 25 rue du Dr Roux, F-75015 Paris,  
22  
23 France  
24

25 <sup>2</sup> Sorbonne Université, Collège Doctoral, 4 Place Jussieu, F-75005 Paris, France  
26

27 <sup>3</sup> CNRS, UMR3525, F-75015 Paris, France  
28

29 <sup>4</sup> Biologics Research, Sanofi R&D, 13 Quai Jules Guesde, 94403 Vitry sur Seine, France  
30  
31

32  
33  
34 Corresponding author: Guy-Franck Richard  
35

36 Tel: +33-1-45-68-84-36  
37

38 e-Mail: gfrichar@pasteur.fr  
39  
40

41 ORCID: 0000-0002-8233-3532  
42  
43  
44  
45  
46  
47

48 **Acknowledgements**  
49

50 The authors wish to thank the continuous support of the Institut Pasteur and of the Centre  
51 National de la Recherche Scientifique (CNRS). L. P. is the recipient of a Cifre PhD  
52 fellowship from Sanofi. V. M. was the recipient of two post-doctoral fellowships from  
53 Fondation Guy Nicolas and from Fondation Hardy.  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## **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

1 SSA: single-strand annealing

2 ZFN: zinc-finger nucleases

3  
4 TALEN: transcription activator-like effector nuclease

5  
6  
7 DSB: double-strand break

8  
9  
10 SDSA: synthesis-dependent strand annealing

11  
12 UAS: upstream activating sequence

13  
14  
15 MRX complex: Mre11-Rad50-Xrs2 complex

16  
17 PAM: protospacer adjacent motif

18  
19  
20 iPSC: induced pluripotent stem cells

21  
22 sgRNA (or gRNA): single-guide RNA

23  
24  
25 *SpCas9: Streptococcus pyogenes Cas9*

26  
27  
28 *SaCas9: Staphylococcus aureus Cas9*

29  
30 HNH: homing endonuclease domain

31  
32 HEK293: human embryonic kidney cell line 293

33  
34  
35 K562: human immortalized myelogenous leukemia cell line

36  
37 AAV: adenovirus-associated vector

## Introduction

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

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

1 that they occurred through a Synthesis-Dependent Strand Annealing (SDSA) mechanism,  
2 a particular type of gene conversion that is never associated to crossover (Figure 1;  
3 Richard and Pâques 2000).  
4

5  
6  
7 Trinucleotide repeat instability may also occur by homologous recombination in the  
8 absence of an induced DSB. Such length changes arise from replication fork blocking  
9 and/or spontaneous breakage during S phase replication. It was shown that CAG/CTG  
10 repeat expansions occurred in a *srs2* yeast mutant, most probably by homologous  
11 recombination between sister chromatids (Kerrest et al. 2009). In the absence of the Srs2  
12 helicase activity, recombination intermediates were increased, as visualized by 2D gel  
13 electrophoresis. They partly disappeared when *RAD51*, the main recombinase gene in  
14 yeast, was deleted, proving that they were *bona fide* recombining molecules (Nguyen et  
15 al. 2017).  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

29 Expansions were also studied in mice deficient for the RAD52 recombination gene, but no  
30 difference in the rate of instability of a (CTG)<sub>300</sub> repeat tract was found, as compared to  
31 control mice (Savouret et al. 2003). However, RAD52 does not play the same role in  
32 mammals as it is playing in *S. cerevisiae*. In yeast cells, it is the mediator of all homologous  
33 recombination events (SSA, BIR, gene conversion) whereas it is only an accessory  
34 recombination gene whose exact function is not totally understood in mammalian cells.  
35 Therefore, it would be interesting to address the effect of BRCA1 and/or BRCA2 mutants  
36 on CAG/CTG repeat expansions, since these two genes belong to the real recombination  
37 mediator complex in human cells (Moynahan et al. 1999, 2001).  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50

51 Large CAG/CTG repeat expansions were also investigated in yeast using an experimental  
52 assay based on the insertion of a (CTG)<sub>140</sub> repeat tract between the *GAL1* UAS and its TATA  
53 box. Transcriptional activation of the downstream reporter no longer occurred if the  
54 repeat tract was too long. The average size of detected expansions ranged from 60 to more  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

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  $\delta$  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)<sub>80</sub> 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 preferred 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).



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

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)<sub>80</sub> 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)<sub>70</sub> repeat expansions of moderate lengths were increased during S phase in a *mre11Δ* 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)<sub>140</sub> repeat expansions were decreased in a *mre11Δ* mutant, from 10<sup>-5</sup> to 10<sup>-6</sup> 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)<sub>20</sub> 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

1 eukaryotes, although its precise role is not totally clear, the mismatch repair machinery  
2 appears to be an important player of repeat instability by its propensity to recognize  
3 mismatches in hairpins formed by trinucleotide repeats while being unable to repair them  
4 (Pearson et al. 1997; Owen et al. 2005; Tomé et al. 2009, 2013; Williams and Surtees 2015;  
5 Slean et al. 2016; Viterbo et al. 2016). It is reasonable to assume that DNA nickases now  
6 available will help to study the possible involvement of single stranded DNA nicks on  
7 CAG/CTG trinucleotide repeat instability.  
8  
9

### 10 **GAA/TTC repeat instability occurs by template switching**

11 A genetic assay was designed in yeast to study large-scale expansions of a (GAA)<sub>78-150</sub>  
12 repeat tract inserted into an artificial intron of the *URA3* gene, larger repeat lengths  
13 inhibiting intron splicing, therefore inactivating the gene (Shishkin et al., 2009).  
14 Expansions reaching more than 300 triplets were observed, as well as small  
15 insertions/deletions or substitutions outside the repeat tract. Large chromosomal  
16 deletions including the *URA3* gene and its flanking sequences were also detected. *RAD50*  
17 or *RAD52* deletion had no effect on the expansion rate, ruling out the implication of  
18 homologous recombination in this process. On the contrary, the absence of replication  
19 fork-stabilizing proteins increased the expansion rate while it was decreased in the  
20 absence of postreplication DNA repair proteins or the Sgs1 DNA helicase. This strongly  
21 suggests that template switching during replication fork progression through GAA  
22 repeats was responsible for the observed GAA expansions (Shishkin et al., 2009). More  
23 recently, advances in long-read DNA sequencing technologies allowed to identify complex  
24 genomic rearrangements originating from improper repair of naturally occurring DSBs at  
25 GAA repeats. Various chromosomal rearrangements involving gene conversion between  
26 Ty retrotransposons and formation of neochromosomes by BIR were described. These  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

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

3  
4 It is worth noting that recombination-independent recognition of DNA homology  
5 associated to mutation in *Neurospora crassa* (and probably in *Ascobolus immersus* too) is  
6 enhanced by GAC/GTC trinucleotides (Gladyshev and Kleckner 2017). It would be  
7 interesting to know if other triplets also interfere with homology recognition and whether  
8 such a mechanism could be involved in trinucleotide repeat instability.  
9

10  
11  
12 In conclusion, although both CAG/CTG and GGA/TTC repeats are apparently able to  
13 trigger DSB formation in yeast, expansions involve different sets of genes, therefore  
14 different molecular pathways. These differences may be due to: i) distinct secondary  
15 structures formed by both types of triplet repeats, GAA tracts folding into triplex DNA  
16 whereas CTG repeats form imperfect hairpins; ii) the nature of DNA damage triggered by  
17 these structures, double- vs single-strand breaks or gaps; iii) the amount of single-  
18 stranded DNA exposed following such damage; iv) differences in chromatin conformation  
19 depending on the repeat tract sequence and structure. All these assumptions being not  
20 mutually exclusive, understanding the genetic complexity of trinucleotide repeat  
21 instability will probably require alternative methods to those applied so far.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45

#### 46 **Gene editing of trinucleotide repeat expansions**

47  
48 No cure is available for any triplet repeat disorder, although several preclinical and  
49 clinical trials have been attempted. Given that microsatellite disorders are always  
50 associated to an expansion of the repeat array, deleting or shortening the expanded array  
51 to non-pathological lengths should suppress symptoms of the pathology. Indeed, when a  
52 trinucleotide repeat contraction occurred during transmission from father to daughter of  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1 an expanded myotonic dystrophy allele, clinical examination of the 17-year old daughter  
2 showed no sign of the symptoms (O'Hoy et al. 1993). In another study, a reversible model  
3 of DM1 transgenic mice, was relying on a mutant GFP gene under the control of the TetOn  
4 promoter, fused to the DMPK 3' UTR. After doxycycline treatment arrest, the GFP-DMPK  
5 transgene expression was stopped and sick mice reverted to normal (Mahadevan et al.  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26 these disorders.

27 Four families of highly specific nucleases may be used to edit trinucleotide repeats:  
28 meganucleases, Zinc-Finger Nucleases (ZFN), Transcription Activator Like Effector  
29 Nucleases (TALEN) and CRISPR-Cas9. Meganucleases are highly specific DNA  
30 endonucleases whose recognition site covers more than 12 bp, originally discovered in  
31 group I self-splicing introns in *S. cerevisiae* mitochondria (Dujon 1989). ZFNs were  
32 engineered from the fusion of a zinc-finger DNA binding domain to the FokI nuclease  
33 domain (Kim et al. 1996). ZFNs are active as heterodimers in which two arms need to  
34 dimerize in order to induce a DSB. TALENs are fusion proteins between a TAL effector  
35 derived from *Xanthomonas* bacteria and FokI, and also function as heterodimers (Cermak  
36 et al. 2011). The Cas9 protein is an RNA-guided nuclease belonging to the CRISPR system  
37 of bacterial acquired immune system. It needs the presence of a Protospacer Adjacent  
38 Motif (PAM) next to its guide sequence to induce one single-strand break on each DNA  
39 strand, resulting in a DSB (Doudna and Charpentier 2014). *Streptococcus pyogenes* Cas9  
40 (*SpCas9*) was engineered by an aspartate-to-alanine substitution (D10A) in the RuvC  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1 catalytic domain to convert the double-strand endonuclease into a single-strand nickase  
2 (Cong et al, 2013). The same approach was used at the HNH catalytic site to generate the  
3 symmetrical nickase cutting the opposite DNA strand (N863A). Depending on their  
4 bacterial origin, Cas9 proteins recognize different PAM and exhibit different activities.  
5 ZFN, TALEN and Cas9 were used to delete or shorten trinucleotide repeats, using two  
6 different approaches: i) induce two DSBs upstream and downstream the repeat tract to  
7 completely delete it, or ii) induce a DSB inside the repeat tract in order to shorten it  
8 (Figure 2).  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21

## 22 **Huntington's disease**

23 Huntington's disease is a dominant disorder caused by the expansion of a CAG repeat tract  
24 in the first exon of the *HTT* gene. In a first study, iPSCs (induced pluripotent stem cells)  
25 derived from Huntington patients harboring 72 CAG triplets were electroporated with a  
26 modified bacterial artificial chromosome containing 11.5 kb of the genomic region  
27 surrounding *HTT* first exon harboring 21 CAG triplets as well as an eGFP reporter cassette  
28 and a neomycin resistance gene. Out of 203 analyzed clones, only two showed the  
29 incorporation of the wild-type locus by homologous recombination. In these two clones,  
30 there was no detectable toxic huntingtin and modified cells retained the modifications  
31 when differentiated into neurons (An et al. 2012) (Table1).  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

47 In another study, patient derived fibroblasts of variable CAG length were transfected with  
48 the D10A nickase and two guide RNAs, each targeting upstream and downstream the CAG  
49 repeat tract. Excision of the CAG repeat in the transfected non-clonal population showed  
50 decreased levels of the *HTT* mRNA and protein, from 68% to 82% depending on the cell  
51 line, suggesting that at least one allele was efficiently deleted, on the average. Four out of  
52 13 predicted exonic off-target sites were tested and no mutation was detected  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

(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/*SpCas9* 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

1 established. The introduction of Cas9 and a pair of guide RNAs each targeting a specific  
2 locus upstream and downstream the DM1 repeats in patient cells resulted in the deletion  
3 of the CTG repeats, the suppression of RNA foci and splicing defects (Van Agtmaal et al.  
4 2017; Provenzano et al. 2017). Those two studies used different cell types, respectively  
5 myogenic DM1 myoblast and DM1 fibroblasts and different target loci and achieved  
6 respectively 46% and 14% of successfully edited cells. Indels were found in both cases at  
7 cut sites and few loci were tested for off-target effects.

8 One last strategy consisted in inserting a polyA signal upstream the CTG tract to prevent  
9 its transcription. This was carried out by making a DSB between exon 9 and 10 of the  
10 DMPK gene, induced by a TALEN, while co-transfecting the polyA cassette (Xia et al.  
11 2015). Successfully edited cells showed phenotype reversion including foci  
12 disappearance and normal splicing of MBNL1 and MBNL2.

### 31 **Fragile X syndrome**

32 The fragile X syndrome is caused by the expansion of a CGG repeat tract in the 5' UTR of  
33 the *FMR1* gene which leads through an undetermined mechanism to the methylation of  
34 the *FMR1* promoter. FXS iPSCs (more than 450 CGG) were transfected with *SpCas9* and a  
35 guideRNA targeting the region upstream the repeat tract (Park et al. 2015). Four potential  
36 off target sites were tested and no mutation was detected. Two successfully edited clones  
37 over 100 tested were obtained. In these two clones, promoter hypermethylation was  
38 abolished and *FMR1* expression was reactivated. A similar study was conducted by cutting  
39 upstream and downstream the CGG repeats using *SpCas9*. The authors observed a  
40 decrease in the methylation profile of the *FMR1* promoter in one of their analysed clones  
41 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). *SpCas9* 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*. *SaCas9* was also transfected in mice fibroblasts but its expression level was much lower than *SpCas9* 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

1 integrate in the genome at all DSB sites generated by the nuclease. They can subsequently  
2 be amplified and serve as primers for genome-wide sequencing of their insertion sites.  
3  
4 This analysis revealed that off-targets are difficult to predict, ranging from little cleavage  
5 outside the target to as many off-target as on-target DSBs, depending on the gRNA chosen.  
6  
7 Cleavage can occur on sites bearing up to seven mismatches and no canonical PAM (Tsai  
8 et al. 2015). CIRCLE-seq is a simpler and more sensitive method to detect off-target sites  
9 *in vitro*, but requires the purified nuclease (Tsai et al. 2017). Using this approach, genomic  
10 DNA that was cleaved by the nuclease in a test tube was amplified and sequenced. This  
11 method is very sensitive but may not be relevant for *in vivo* assays and may depend on  
12 each cell type and chromatin state. Recently, the VIVO method was set up for *in vivo*  
13 validation of off-target sites found by CIRCLE-seq, demonstrating that careful choice of  
14 the gRNA may strongly reduce off-target effects, while keeping a good on-target efficacy  
15 (Akcakaya et al. 2018). The same team engineered a more specific version of *SpCas9*,  
16 called HF1, by mutating residues involved in the binding to the target DNA strand. Cas9-  
17 HF1 retains on-target activity comparable to wild-type on 85% of gRNAs tested and  
18 rendered all or nearly all off-target events not detectable by GUIDE-seq (Kleinstiver et al.  
19 2016). No such extensive off-target study was carried out in any of the aforementioned  
20 articles. Such approaches must be encouraged in future assessments of gene therapy  
21 strategies for trinucleotide repeat disorders.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49

### 50 **Limitations of nuclease approaches: vectorization**

51 Nuclease vectorization is clearly a problem that also needs to be addressed. Adenovirus-  
52 associated vectors (AAV) are popular in gene therapy because they exhibit low  
53 integration frequency, but they have a limited cargo capacity making it impossible to  
54 deliver a full length *SpCas9* with its cognate guide, or a TALEN. In this case, each of the  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1 two TALEN arms must be delivered by two different vectors, lowering the efficacy of the  
2 transduction. Alternative non-viral delivery systems such as cationic lipid transfection  
3 particles was efficient to deliver a Cas9-gRNA complex as well as a TALEN both *in vitro*  
4 and *in vivo*, achieving 20% efficacy in genome modification in mice (Zuris et al. 2015).  
5 AAV-based delivery could also potentially increase the rate of off-target site cleavage due  
6 to prolonged expression of the nuclease. To circumvent this problem, a self-limiting  
7 CRISPR-Cas9 system was implemented *in vivo* by inserting the sequence recognized by  
8 the nuclease on the plasmid encoding it such that the expression plasmid would be cut  
9 and eliminated following *SpCas9* expression (Ruan et al. 2017).

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

## 41 **Conclusion**

42 Little is known yet about the immune response toward these nucleases. A very recent  
43 work identified pre-existing immunity against Cas9 from *Streptococcus pyogenes* and  
44 *Staphylococcus aureus* (Charlesworth et al, 2018). The authors showed that 70% of  
45 healthy adults have antibodies directed to the nuclease and that *SaCas9* induced a T-cell  
46 response in adult blood. A strong immune response may be a potential drawback to the  
47 use of Cas9 in future gene therapy.

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

1 cellular response to double-strand breaks. Two studies have recently shown that during  
2 gene editing, cells with a functional p53 pathway were counterselected, due to cell arrest  
3 triggered by p53 upon DSB formation. Therefore, checkpoint activity should be tightly  
4 controlled when developing cell-based therapies utilizing CRISPR–Cas9 (Haapaniemi et  
5 al. 2018; Ihry et al. 2018).  
6  
7

8  
9  
10  
11 These first reports of gene therapy attempts of trinucleotide repeat disorders are  
12 certainly promising and already give us insights into crucial factors to be considered when  
13 evaluating the success of a gene therapy approach: off-target sites number and frequency,  
14 nuclease efficacy, cell type to be targeted and vectorization method. Successful gene  
15 editing was achieved in a mouse model for Duchenne muscular dystrophy, by three  
16 independent teams. Using AAV delivery of Cas9, they obtained partial restoration of  
17 dystrophin levels that were sufficient to allow partial muscle strength recovery (Long et  
18 al. 2016; Nelson et al. 2016; Tabebordbar et al. 2016). Forthcoming experiments in a  
19 mouse model for trinucleotide repeat disorders will establish if a similar success may be  
20 achieved.  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39

#### 40 **Reference list**

- 41  
42 Agtmaal EL van, André LM, Willemse M, et al (2017) CRISPR/Cas9-Induced (CTG·CAG)<sub>n</sub>  
43 Repeat Instability in the Myotonic Dystrophy Type 1 Locus: Implications for Therapeutic  
44 Genome Editing. *Molecular Therapy* 25:24–43. doi: 10.1016/j.ymthe.2016.10.014  
45 Akcakaya P, Bobbin ML, Guo JA, et al (2018) In vivo CRISPR-Cas gene editing with no  
46 detectable genome-wide off-target mutations. *bioRxiv* 272724. doi: 10.1101/272724  
47 An MC, Zhang N, Scott G, et al (2012) Genetic Correction of Huntington’s Disease  
48 Phenotypes in Induced Pluripotent Stem Cells. *Cell Stem Cell* 11:253–263. doi:  
49 10.1016/j.stem.2012.04.026  
50 Blackwood JK, Okely EA, Zahra R, et al (2010) DNA tandem repeat instability in the  
51 *Escherichia coli* chromosome is stimulated by mismatch repair at an adjacent CAG·CTG  
52 trinucleotide repeat. *Proc Natl Acad Sci USA* 107:22582–22586. doi:  
53 10.1073/pnas.1012906108  
54 Callahan JL, Andrews KJ, Zakian VA, Freudenreich CH (2003) Mutations in yeast replication  
55 proteins that increase CAG/CTG expansions also increase repeat fragility. *Mol Cell Biol*  
56 23:7849–7860  
57 Campuzano V, Montermini L, Molto MD, et al (1996) Friedreich’s Ataxia: autosomal  
58  
59  
60  
61  
62  
63  
64  
65

recessive disease caused by an intronic GAA triplet repeat expansion. *Science* 271:1423–1427

1  
2 Cermak T, Doyle EL, Christian M, et al (2011) Efficient design and assembly of custom  
3 TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res*  
4 39:e82. doi: 10.1093/nar/gkr218  
5  
6 Cinesi C, Aeschbach L, Yang B, Dion V (2016) Contracting CAG/CTG repeats using the  
7 CRISPR-Cas9 nickase. *Nature Communications* 7:13272. doi: 10.1038/ncomms13272  
8  
9 Colleaux L, d’Auriol L, Betermier M, et al (1986) Universal code equivalent of a yeast  
10 mitochondrial intron reading frame is expressed into *E. Coli* as a specific double strand  
11 break endonuclease. *Cell* 44:521–533  
12  
13 Dabrowska M, Juzwa W, Krzyzosiak WJ, Olejniczak M (2018) Precise Excision of the CAG  
14 Tract from the Huntingtin Gene by Cas9 Nickases. *Front Neurosci* 12:. doi:  
15 10.3389/fnins.2018.00075  
16  
17 Doudna JA, Charpentier E (2014) Genome editing. The new frontier of genome engineering  
18 with CRISPR-Cas9. *Science* 346:1258096. doi: 10.1126/science.1258096  
19  
20 Dujon B (1989) Group I introns as mobile genetic elements: facts and mechanistic  
21 speculations--a review. *Gene* 82:91–114  
22  
23 Freudenreich CH, Kantrow SM, Zakian VA (1998) Expansion and length-dependent fragility  
24 of CTG repeats in yeast. *Science* 279:853–856  
25  
26 Fu Y-H, Kuhl DPA, Pizzuti A, et al (1991) Variation of the CGG repeat at the fragile X site  
27 results in genetic instability: resolution of the Sherman paradox. *Cell* 67:1047–1058  
28  
29 Gladyshev E, Kleckner N (2017) Recombination-independent recognition of DNA homology  
30 for repeat-induced point mutation. *Curr Genet* 63:389–400. doi: 10.1007/s00294-016-  
31 0649-4  
32  
33 Haapaniemi E, Botla S, Persson J, et al (2018) CRISPR–Cas9 genome editing induces a p53-  
34 mediated DNA damage response. *Nature Medicine* 1. doi: 10.1038/s41591-018-0049-z  
35  
36 Ihry RJ, Worringer KA, Salick MR, et al (2018) p53 inhibits CRISPR–Cas9 engineering in  
37 human pluripotent stem cells. *Nature Medicine* 1. doi: 10.1038/s41591-018-0050-6  
38  
39 Jankowski C, Nasar F, Nag DK (2000) Meiotic instability of CAG repeat tracts occurs by  
40 double-strand break repair in yeast. *Proc Natl Acad Sci USA* 97:2134–2139. doi:  
41 10.1073/pnas.040460297  
42  
43 Kerrest A, Anand R, Sundararajan R, et al (2009) SRS2 and SGS1 prevent chromosomal  
44 breaks and stabilize triplet repeats by restraining recombination. *Nature Structural and*  
45 *Molecular Biology* 16:159–167  
46  
47 Kim H-M, Narayanan V, Mieczkowski PA, et al (2008) Chromosome fragility at GAA tracts  
48 in yeast depends on repeat orientation and requires mismatch repair. *EMBO J* 27:2896–  
49 2906. doi: 10.1038/emboj.2008.205  
50  
51 Kim JC, Harris ST, Dinter T, et al (2017) The role of break-induced replication in large-scale  
52 expansions of (CAG)*n*/(CTG)*n*repeats. *Nat Struct Mol Biol* 24:55–60. doi:  
53 10.1038/nsmb.3334  
54  
55 Kim JC, Mirkin SM (2013) The balancing act of DNA repeat expansions. *Current Opinion in*  
56 *Genetics & Development* 23:280–288. doi: 10.1016/j.gde.2013.04.009  
57  
58 Kim YG, Cha J, Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusions to  
59 Fok I cleavage domain. *Proc Natl Acad Sci U S A* 93:1156–60  
60  
61 Kleinstiver BP, Pattanayak V, Prew MS, et al (2016) High-fidelity CRISPR–Cas9 nucleases  
62 with no detectable genome-wide off-target effects. *Nature* 529:490–495. doi:  
63 10.1038/nature16526  
64  
65 Kostriken R, Strathern JN, Klar AJ, et al (1983) A site-specific endonuclease essential for  
66 mating-type switching in *Saccharomyces cerevisiae*. *Cell* 35:167–74. doi: 0092-  
67 8674(83)90219-2 [pii]

- 1 Lahiri M, Gustafson TL, Majors ER, Freudenreich CH (2004) Expanded CAG repeats  
2 activate the DNA damage checkpoint pathway. *Mol Cell* 15:287–293. doi:  
3 10.1016/j.molcel.2004.06.034
- 4 Lengsfeld BM, Rattray AJ, Bhaskara V, et al (2007) Sae2 Is an Endonuclease that Processes  
5 Hairpin DNA Cooperatively with the Mre11/Rad50/Xrs2 Complex. *Molecular Cell*  
6 28:638–651. doi: 10.1016/j.molcel.2007.11.001
- 7 Li Y, Polak U, Bhalla AD, et al (2015) Excision of Expanded GAA Repeats Alleviates the  
8 Molecular Phenotype of Friedreich’s Ataxia. *Mol Ther* 23:1055–1065. doi:  
9 10.1038/mt.2015.41
- 10 Liu G, Chen X, Bissler JJ, et al (2010) Replication-dependent instability at (CTG)<sub>x</sub> (CAG)  
11 repeat hairpins in human cells. *Nat Chem Biol* 6:652–9. doi: 10.1038/nchembio.416
- 12 Llorente B, Smith CE, Symington LS (2008) Break-induced replication: what is it and what is  
13 it for? *Cell Cycle* 7:859–864. doi: 10.4161/cc.7.7.5613
- 14 Long C, Amoasii L, Mireault AA, et al (2016) Postnatal genome editing partially restores  
15 dystrophin expression in a mouse model of muscular dystrophy. *Science* 351:400–403. doi:  
16 10.1126/science.aad5725
- 17 Lydeard JR, Jain S, Yamaguchi M, Haber JE (2007) Break-induced replication and  
18 telomerase-independent telomere maintenance require Pol32. *Nature* 448:820–823. doi:  
19 10.1038/nature06047
- 20 Mahadevan MS, Yadava RS, Yu Q, et al (2006) Reversible model of RNA toxicity and  
21 cardiac conduction defects in myotonic dystrophy. *Nat Genet* 38:1066–1070. doi:  
22 10.1038/ng1857
- 23 McGinty RJ, Mirkin SM (2018) Cis- and Trans-Modifiers of Repeat Expansions: Blending  
24 Model Systems with Human Genetics. *Trends Genet* 34:448–465. doi:  
25 10.1016/j.tig.2018.02.005
- 26 McGinty RJ, Rubinstein RG, Neil AJ, et al (2017) Nanopore sequencing of complex genomic  
27 rearrangements in yeast reveals mechanisms of repeat-mediated double-strand break  
28 repair. *Genome Res* 27:2072–2082. doi: 10.1101/gr.228148.117
- 29 McMurray CT (2010) Mechanisms of trinucleotide repeat instability during human  
30 development. *Nat Rev Genet* 11:786–99. doi: 10.1038/nrg2828
- 31 Meservy JL, Sargent RG, Iyer RR, et al (2003) Long CTG Tracts from the Myotonic  
32 Dystrophy Gene Induce Deletions and Rearrangements during Recombination at the APRT  
33 Locus in CHO Cells. *Mol Cell Biol* 23:3152–3162. doi: 10.1128/MCB.23.9.3152-  
34 3162.2003
- 35 Miller JW, Urbinati CR, Teng-umnuay P, et al (2000) Recruitment of human musclebind  
36 proteins to (CUG)<sub>n</sub> expansions associated with myotonic dystrophy. *EMBO J* 19:4439–  
37 4448
- 38 Monteys AM, Ebanks SA, Keiser MS, Davidson BL (2017) CRISPR/Cas9 Editing of the  
39 Mutant Huntingtin Allele In Vitro and In Vivo. *Mol Ther* 25:12–23. doi:  
40 10.1016/j.ymthe.2016.11.010
- 41 Mosbach V, Poggi L, Viterbo D, et al (2018) TALEN-Induced Double-Strand Break Repair  
42 of CTG Trinucleotide Repeats. *Cell Rep* 22:2146–2159. doi: 10.1016/j.celrep.2018.01.083
- 43 Moynahan ME, Chiu JW, Koller BH, Jasin M (1999) Brca1 controls homology-directed DNA  
44 repair. *Mol Cell* 4:511–8
- 45 Moynahan ME, Pierce AJ, Jasin M (2001) BRCA2 is required for homology-directed repair  
46 of chromosomal breaks. *Mol Cell* 7:263–72
- 47 Neil Alexander J., Kim Jane C., Mirkin Sergei M. (2017) Precarious maintenance of simple  
48 DNA repeats in eukaryotes. *BioEssays* 39:1700077. doi: 10.1002/bies.201700077
- 49 Nelson CE, Hakim CH, Ousterout DG, et al (2016) In vivo genome editing improves muscle  
50 function in a mouse model of Duchenne muscular dystrophy. *Science* 351:403–407. doi:  
51 10.1126/science.1257631

10.1126/science.aad5143

- 1 Nguyen JHG, Viterbo D, Anand RP, et al (2017) Differential requirement of Srs2 helicase  
2 and Rad51 displacement activities in replication of hairpin-forming CAG/CTG repeats.  
3 *Nucleic Acids Res* 45:4519–4531. doi: 10.1093/nar/gkx088
- 4 O’Hoy KL, Tsilfidis C, Mahadevan MS, et al (1993) Reduction in size of the myotonic  
5 dystrophy trinucleotide repeat mutation during transmission. *Science* 259:809–12
- 6 Ouellet DL, Cherif K, Rousseau J, Tremblay JP (2017) Deletion of the GAA repeats from the  
7 human frataxin gene using the CRISPR-Cas9 system in YG8R-derived cells and mouse  
8 models of Friedreich ataxia. *Gene Ther* 24:265–274. doi: 10.1038/gt.2016.89
- 9 Owen BAL, Yang Z, Lai M, et al (2005) (CAG)(n)-hairpin DNA binds to Msh2-Msh3 and  
10 changes properties of mismatch recognition. *Nat Struct Mol Biol* 12:663–670. doi:  
11 10.1038/nsmb965
- 12 Park C-Y, Halevy T, Lee DR, et al (2015) Reversion of FMR1 Methylation and Silencing by  
13 Editing the Triplet Repeats in Fragile X iPSC-Derived Neurons. *Cell Rep* 13:234–241. doi:  
14 10.1016/j.celrep.2015.08.084
- 15 Pearson CE, Ewel A, Acharya S, et al (1997) Human MSH2 binds to trinucleotide repeat  
16 DNA structures associated with neurodegenerative diseases. *Hum Mol Genet* 6:1117–1123
- 17 Provenzano C, Cappella M, Valaperta R, et al (2017) CRISPR/Cas9-Mediated Deletion of  
18 CTG Expansions Recovers Normal Phenotype in Myogenic Cells Derived from Myotonic  
19 Dystrophy 1 Patients. *Mol Ther Nucleic Acids* 9:337–348. doi:  
20 10.1016/j.omtn.2017.10.006
- 21 Richard GF, Cyncynatus C, Dujon B (2003) Contractions and expansions of CAG/CTG  
22 trinucleotide repeats occur during ectopic gene conversion in yeast, by a MUS81-  
23 independent mechanism. *J Mol Biol* 326:769–782
- 24 Richard GF, Dujon B, Haber JE (1999) Double-strand break repair can lead to high  
25 frequencies of deletions within short CAG/CTG trinucleotide repeats. *Mol Gen Genet*  
26 261:871–882
- 27 Richard GF, Goellner GM, McMurray CT, Haber JE (2000) Recombination-induced CAG  
28 trinucleotide repeat expansions in yeast involve the MRE11-RAD50-XRS2 complex.  
29 *EMBO J* 19:2381–2390. doi: 10.1093/emboj/19.10.2381
- 30 Richard GF, Kerrest A, Dujon B (2008) Comparative genomics and molecular dynamics of  
31 DNA repeats in eukaryotes. *Microbiol Mol Biol Rev* 72:686–727
- 32 Richard G-F, Pâques F (2000) Mini- and microsatellite expansions: the recombination  
33 connection. *EMBO Reports* 1:122–126
- 34 Richard G-F, Viterbo D, Khanna V, et al (2014) Highly specific contractions of a single  
35 CAG/CTG trinucleotide repeat by TALEN in yeast. *PLoS ONE* 9:e95611. doi:  
36 10.1371/journal.pone.0095611
- 37 Ruan G-X, Barry E, Yu D, et al (2017) CRISPR/Cas9-Mediated Genome Editing as a  
38 Therapeutic Approach for Leber Congenital Amaurosis 10. *Mol Ther* 25:331–341. doi:  
39 10.1016/j.ymthe.2016.12.006
- 40 Savouret C, Brisson E, Essers J, et al (2003) CTG repeat instability and size variation timing  
41 in DNA repair-deficient mice. *Embo J* 22:2264–73
- 42 Shin JW, Kim K-H, Chao MJ, et al (2016) Permanent inactivation of Huntington’s disease  
43 mutation by personalized allele-specific CRISPR/Cas9. *Hum Mol Genet* 25:4566–4576.  
44 doi: 10.1093/hmg/ddw286
- 45 Shishkin AA, Voineagu I, Matera R, et al (2009) Large-scale expansions of Friedreich’s  
46 ataxia GAA repeats in yeast. *Mol Cell* 35:82–92. doi: 10.1016/j.molcel.2009.06.017
- 47 Slean MM, Panigrahi GB, Castel AL, et al (2016) Absence of MutS $\beta$  leads to the formation of  
48 slipped-DNA for CTG/CAG contractions at primate replication forks. *DNA Repair (Amst)*  
49 42:107–118. doi: 10.1016/j.dnarep.2016.04.002

- 1 Sundararajan R, Gellon L, Zunder RM, Freudenreich CH (2010) Double-strand break repair  
2 pathways protect against CAG/CTG repeat expansions, contractions and repeat-mediated  
3 chromosomal fragility in *Saccharomyces cerevisiae*. *Genetics* 184:65–77. doi:  
4 10.1534/genetics.109.111039
- 5 Tabebordbar M, Zhu K, Cheng JKW, et al (2016) In vivo gene editing in dystrophic mouse  
6 muscle and muscle stem cells. *Science* 351:407–411. doi: 10.1126/science.aad5177
- 7 Takahashi K, Tanabe K, Ohnuki M, et al (2007) Induction of pluripotent stem cells from adult  
8 human fibroblasts by defined factors. *Cell* 131:861–872. doi: 10.1016/j.cell.2007.11.019
- 9 Tomé S, Holt I, Edelmann W, et al (2009) MSH2 ATPase domain mutation affects  
10 CTG\**CAG* repeat instability in transgenic mice. *PLoS Genet* 5:e1000482. doi:  
11 10.1371/journal.pgen.1000482
- 12 Tomé S, Manley K, Simard JP, et al (2013) MSH3 polymorphisms and protein levels affect  
13 CAG repeat instability in Huntington’s disease mice. *PLoS Genet* 9:e1003280. doi:  
14 10.1371/journal.pgen.1003280
- 15 Tsai SQ, Nguyen NT, Malagon-Lopez J, et al (2017) CIRCLE-seq: a highly sensitive in vitro  
16 screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat Meth* 14:607–614. doi:  
17 10.1038/nmeth.4278
- 18 Tsai SQ, Zheng Z, Nguyen NT, et al (2015) GUIDE-seq enables genome-wide profiling of  
19 off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* 33:187–97. doi:  
20 10.1038/nbt.3117
- 21 Usdin K, House NC, Freudenreich CH (2015) Repeat instability during DNA repair: Insights  
22 from model systems. *Critical Reviews in Biochemistry and Molecular Biology*  
23 doi:10.3109/10409238.2014.999192. doi: 10.3109/10409238.2014.999192
- 24 Verkerk AJMH, Pieretti M, Sutcliffe JS, et al (1991) Identification of a gene (FMR-1)  
25 containing a CGG repeat coincident with a breakpoint cluster region exhibiting length  
26 variation in fragile X syndrome. *Cell* 65:905–914
- 27 Viterbo D, Michoud G, Mosbach V, et al (2016) Replication stalling and heteroduplex  
28 formation within CAG/CTG trinucleotide repeats by mismatch repair. *DNA Repair (Amst)*  
29 42:94–106. doi: 10.1016/j.dnarep.2016.03.002
- 30 Williams GM, Surtees JA (2015) MSH3 Promotes Dynamic Behavior of Trinucleotide Repeat  
31 Tracts In Vivo. *Genetics* 200:737–754. doi: 10.1534/genetics.115.177303
- 32 Xia G, Gao Y, Jin S, et al (2015) Genome modification leads to phenotype reversal in human  
33 myotonic dystrophy type 1 induced pluripotent stem cell-derived neural stem cells. *Stem*  
34 *Cells* 33:1829–38. doi: 10.1002/stem.1970
- 35 Xie N, Gong H, Suhl JA, et al (2016) Reactivation of FMR1 by CRISPR/Cas9-Mediated  
36 Deletion of the Expanded CGG-Repeat of the Fragile X Chromosome. *PLOS ONE*  
37 11:e0165499. doi: 10.1371/journal.pone.0165499
- 38 Yamamoto A, Lucas JJ, Hen R (2000) Reversal of neuropathology and motor dysfunction in a  
39 conditional model of Huntington’s disease. *Cell* 101:57–66. doi: 10.1016/S0092-  
40 8674(00)80623-6
- 41 Ye Y, Kirkham-McCarthy L, Lahue RS (2016) The *Saccharomyces cerevisiae* Mre11-Rad50-  
42 Xrs2 complex promotes trinucleotide repeat expansions independently of homologous  
43 recombination. *DNA Repair (Amst)* 43:1–8. doi: 10.1016/j.dnarep.2016.04.012
- 44 Yu J, Vodyanik MA, Smuga-Otto K, et al (2007) Induced Pluripotent Stem Cell Lines  
45 Derived from Human Somatic Cells. *Science* 318:1917–1920. doi:  
46 10.1126/science.1151526
- 47 Yu S, Pritchard M, Kremer E, et al (1991) Fragile X genotype characterized by an unstable  
48 region of DNA. *Science* 252:1179–81
- 49 Yudkin D, Hayward BE, Aladjem MI, et al (2014) Chromosome fragility and the abnormal  
50 replication of the FMR1 locus in fragile X syndrome. *Hum Mol Genet* 23:2940–2952. doi:  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65



10.1093/hmg/ddu006

1 Zu T, Duvick LA, Kaytor MD, et al (2004) Recovery from polyglutamine-induced  
2 neurodegeneration in conditional SCA1 transgenic mice. *J Neurosci* 24:8853–8861. doi:  
3 10.1523/JNEUROSCI.2978-04.2004

4  
5 Zuris JA, Thompson DB, Shu Y, et al (2015) Cationic lipid-mediated delivery of proteins  
6 enables efficient protein-based genome editing in vitro and in vivo. *Nat Biotechnol* 33:73–  
7 80. doi: 10.1038/nbt.3081  
8  
9

10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

**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

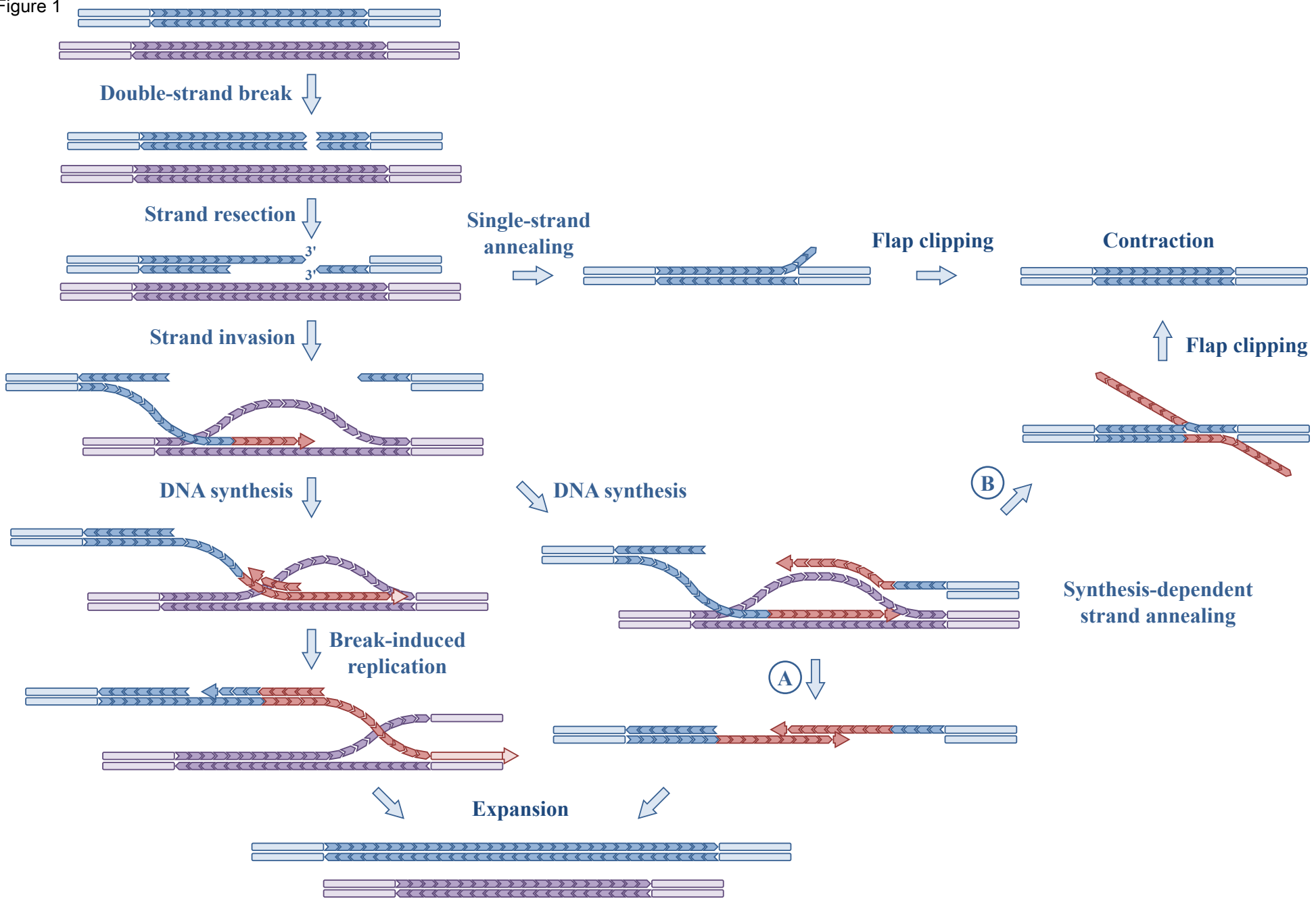


Figure 2

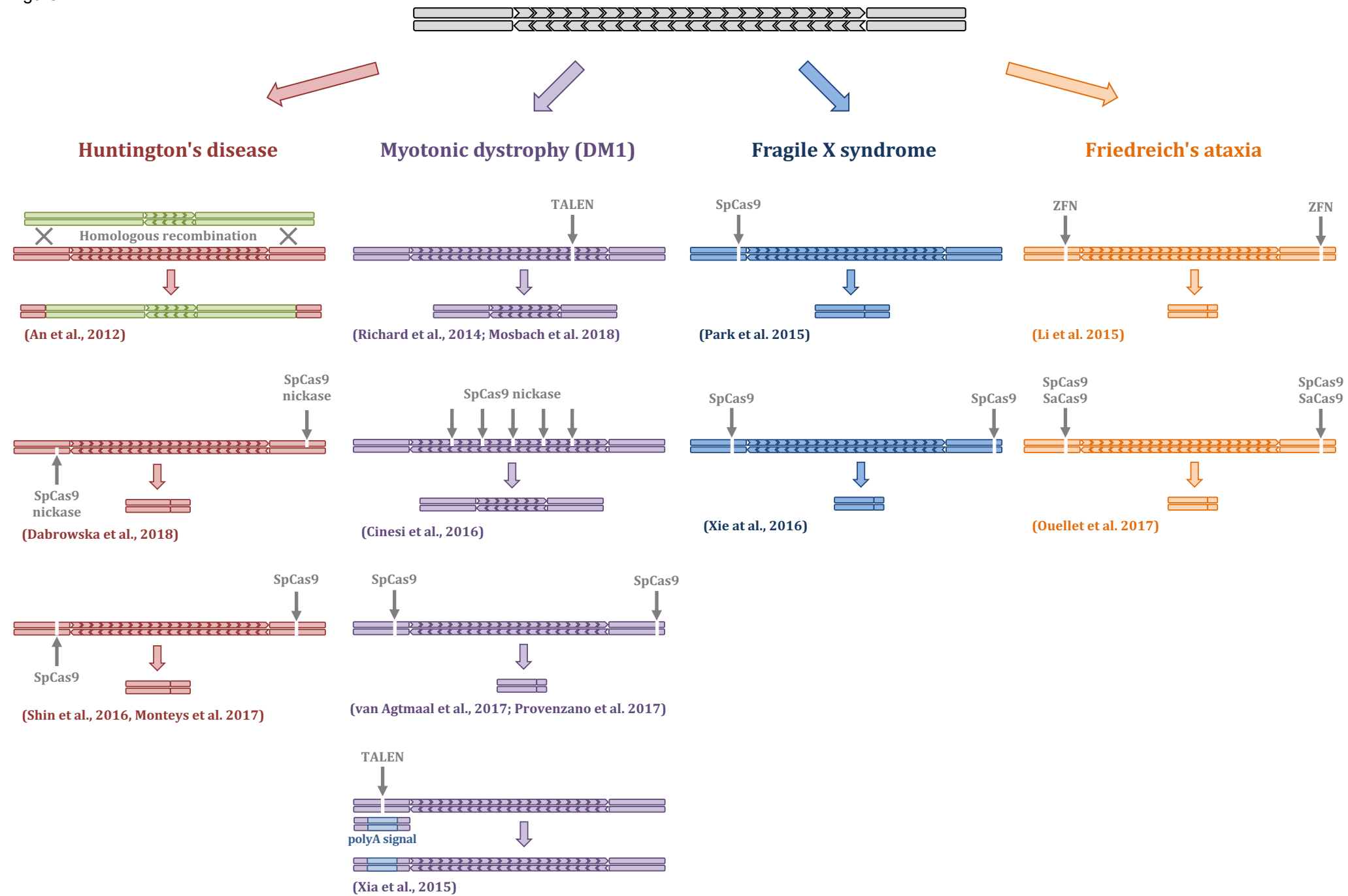


Table 1

<b>Disease</b>	Huntington's disease			
<b>Reference</b>	An <i>et al</i> , 2012	Dabrowska <i>et al</i> , 2018	Shin <i>et al</i> , 2016	Monteys <i>et al</i> , 2017
<b>Cell type</b>	HD iPS cells	HD fibroblasts	HD fibroblasts	HD fibroblasts BachHD mice
<b>Nuclease used</b>	None (spontaneous homologous recombination)	Paired D10A nickases	<i>SpCas9</i>	<i>SpCas9</i>
<b>Successful edition</b>	1% (203 clones analysed)	NA (bulk analysis)	NA (bulk analysis)	NA (bulk analysis)
<b>Off-target analysis</b>	NA	Indels at cut site 4 off target sites analysed: unchanged	None	11 top off target sites: unchanged
<b>Phenotype of edited cells</b>	No detectable toxic huntingtin	No detectable toxic huntingtin	HTT mRNA and protein levels decreased	HTT mRNA and protein levels decreased

<b>Disease</b>	Myotonic Dystrophy type I				
<b>Reference</b>	Richard <i>et al.</i> , 2014	Cinesi <i>et al.</i> , 2016	Provenzano <i>et al.</i> , 2017	Van Atgmaal <i>et al.</i> , 2017	Xia <i>et al.</i> , 2015
<b>Cell type</b>	<i>Saccharomyces cerevisiae</i>	HEK293 GFP(CAG) <sub>89</sub>	Immortalized myogenic DM1 fibroblast	Immortalized DM1 myoblast (DM11)	DM1 neural stem cells
<b>Nuclease</b>	TALEN	ZFN <i>SpCas9</i> D10A Cas9 nickase	<i>eSpCas9</i>	<i>SpCas9</i>	TALEN
<b>Successful edition</b>	99%	3%	14% (85 clones analysed)	46% (103 clones analysed)	After selection: 4 out of 10 colonies
<b>Off-target analysis</b>	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)
<b>Phenotype of edited cells</b>	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

<b>Disease</b>	Fragile X syndrome		Friedreich's ataxia	
<b>Reference</b>	Park <i>et al</i> , 2015	Xie <i>et al</i> , 2016	Li <i>et al</i> , 2015	Ouellet <i>et al</i> , 2017
<b>Cell type</b>	FXS iPS cells	FXS iPS cells	FRDA fibroblasts and lymphoblasts	Transgenic mouse fibroblasts and whole animal muscles
<b>Nuclease</b>	<i>SpCas9</i>	<i>SpCas9</i>	ZFN	<i>SpCas9</i> and <i>SaCas9</i>
<b>Successful edition</b>	2% (100 clones analysed)	5 clones analyzed	6.7% (344 fibroblasts analysed) 2.3% (305 lymphoblasts analysed)	15% for the best gRNA combination in fibroblasts (33 clones analysed) No quantification <i>in vivo</i>
<b>Off-target analysis</b>	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
<b>Phenotype of edited cells</b>	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**