DNA repair in the triplet repeat diseases

Lesley Jones[1], Henry Houlden[2] and Sarah J Tabrizi[3]

1. MRC Centre for Neuropsychiatric Genetics and Genomics, Institute of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff CF24 4HQ, UK
2. Department of Molecular Neuroscience and MRC Centre for Neuromuscular Diseases, Institute of Neurology, Queen Square, London WC1N 3BG, UK
3. UCL Huntington's disease centre, Dept of Neurodegenerative Disease, Institute of Neurology, University College London, London WC1N 3BG, UK

Corresponding author: Lesley Jones
MRC Centre for Neuropsychiatric Genetics and Genomics, Haydn Ellis Building, Maindy Road, Cardiff University, Cardiff CF24 4HQ, UK

email: jonesl1@cf.ac.uk
phone: 02920688469

Abstract

Background

Inherited diseases caused by unstable repeated DNA sequences are rare but together represent a substantial cause of morbidity. Triplet repeat diseases are severe, usually life-shortening, neurological disorders caused by nucleotide expansions and most have no disease-modifying treatment. Longer repeat expansions are associated with genetic anticipation, earlier disease onset in successive generations, and earlier disease onset: however, not all the difference in age at onset of these diseases is accounted for by repeat length, implying the existence of additional modifying factors. As modifying factors alter the disease in people they must lie in pathways that can potentially be modulated to treat disease.

Recent developments

A recent genome wide association study detected genetic modifiers of age at onset in Huntington's disease, which replicated in the spinocerebellar ataxias, finding a significant association between DNA damage response/repair pathways and the age at onset of disease. These data indicate a common genetic mechanism modulating age at onset in polyglutamine diseases that might extend to other repeat expansion disorders. Genetic defects in DNA repair underlie other neurodegenerative disorders such as ataxia-telangiectasia and recent work has demonstrated that double-stranded DNA breaks are critical in modulating early gene expression, which provides a mechanistic link between DNA repair and neurodegeneration. Mismatch and base-excision repair have both been demonstrated to be key in the somatic expansion of repeated sequences in mouse models of repeat disease, and somatic expansion of the expanded CAG tract in *HTT* is known to correlate with age at onset of Huntington's disease and other triplet repeat disorders.

Where next

To understand both the common genetic architecture of these diseases and any further individual disease genetic susceptibilities requires further genetic analysis using more variants and larger samples, followed by sequencing approaches to define the phenotype-modifying variants. This must then be translated, using cell biology, to elucidate the mechanisms through which the genetic variants operate. Genes that have a role in

the DNA damage response may underpin a common DNA repeat-based mechanism and provide new therapeutic targets and hence therapeutics that act in multiple repeat diseases.

Search strategy and selection criteria

We searched PubMed titles and abstracts using combinations of the terms “huntingt*”, “spinocerebellar ataxia”, “trinucleotide repeat”, “triplet repeat”, “repeat” or “repeat disease” AND “DNA integrity”, “DNA repair”, “genome integrity” or “genome repair” from January 1st 2012 to September 30th 2016, until no new references were identified. We identified further relevant papers by examination of the reference lists of these papers and through searches of our files. The final reference list was generated on the basis of relevance to the topic of this Rapid Review.

Introduction

The inherited diseases caused by unstable repeated DNA sequences were first characterised in the 1990s. They are individually rare, with Fragile X the commonest at 1/4000 males, myotonic dystrophy and Huntington's disease (HD) around 1/10,000 and most spinocerebellar ataxias (SCAs) around 1/100,000 and some seen in only a handful of cases[1], but together they represent a substantial source of morbidity. There is substantial geographical variation in prevalence for HD and the SCAs. Most are life-shortening with debilitating symptoms and no available disease modifying treatments. While they have a similar mutational mechanism, the repeated sequences occur in different genomic contexts and even in the polyglutamine diseases, where the repeated codon is translated to glutamine, the proteins are functionally unrelated. The nature and expression pattern of the repeat containing proteins is likely to underlie the clinical differences between these diseases[1], but substantial phenotypic variability occurs within each disease which remains only partially explained. This variability can be exploited to gain insights into disease mechanism though genetics[2].

The repeat diseases can be subdivided into two main categories – those where the repeated sequence is translated into a protein product, and those where the repeat lies outside the coding sequence (Table 1). The non-coding disease associated repeat sequences are usually longer than those in the coding repeats They all display genetic anticipation, the earlier onset of disease in successive generations of families, caused by the germline expansion of the repeat[3]. There is also expansion of the repeat in dividing and non-dividing cells that is tissue, cell-type and disease specific[1]. Expansion of the repeat is ameliorated if the repeated sequence is interrupted by other codons. Despite the repeat associations to specific loci having been known since the 1990s, the mechanistic cascade from repeat to clinical phenotype in most of these diseases remains unclear,hindering the development of new treatments. There are some common pathogenic mechanisms. The repeat may prevent expression of the gene, as in Fragile X and Friedreich's ataxia[4]. Pathogenic RNA foci occur in myotonic dystrophy and myotonic dystrophy-like 2 and give rise to characteristic splicing deficits[5] and have been reported in other repeat diseases[6]. Repeat-associated non-ATG (RAN) translation, first observed in myotonic dystrophy and spinocerebellar ataxia 8 (SCA8)[7], has also been

observed in Huntington's disease (HD)[8], fronto-temporal dementia/amyotrophic lateral sclerosis (FTD/ALS), caused by the *C9ORF72* hexanucleotide repeat, and other repeat disorders[9]. While in the *C9ORF72* repeat disease these dipeptides are neurotoxic[10], their role in pathogenicity in other repeat diseases is unknown.

These mechanisms may also operate in the polyglutamine diseases[9]. However, the proteins containing expanded polyglutamine tracts aggregate and form characteristic insoluble protein inclusions in neural and other cells. Such insoluble inclusions are also widely seen amongst other neurodegenerations[11], leading the field to hypothesise that the protein inclusions, or their soluble oligomers, are likely to be pathogenic. This remains controversial as the final proof, that preventing aggregation can prevent disease in people, has not been demonstrated[12], though there are some recent, tantalising, hints. In early clinical trials aducanumab, an antibody that binds and reduces amyloid-β in mouse models and subjects with early Alzheimer's disease (AD), showed cognitive benefits[13]. ATXN1 oligomers have been shown to drive toxicity in spinocerebellar ataxia 1 (SCA1), and induce local spread of pathology[14], that was partially inhibited using an immunotherapy[15]. There has been extensive study of the biological consequences of expanded polyglutamine with a wide range of potentially deleterious outcomes detected[16,17] but it is unclear which of these are important in manifestation of disease. New genetic evidence indicates that the DNA damage response and DNA repair (Box 1) affect the clinical presentation of HD and multiple spinocerebellar ataxias (SCAs)[18–20] implicating common modifiers that act on the mutated repeat itself. Together with evidence implicating these processes in repeat disease biology this sheds light on mechanism and highlights new targets for therapeutic intervention.

The DNA damage response and neurological disease

The DNA damage response (Box 1) can be both deleterious and protective for neurological diseases. Mutations in genes involved with the DNA damage response were first noted to cause neurological disease in ataxia telangiectasia (A-T), a rare recessive childhood neurodegeneration. Mutations in ATM serine/threonine kinase (*ATM)* cause A-T: ATM controls cell-cycle arrest after DNA double-strand breaks, often leading to apoptosis and thus neurodegeneration[21]. Mutations in other genes that cause incorrect resolution of double-strand DNA breaks also result in profound developmental nervous system pathology

such as AT-like disease, ATR-Seckel syndrome and Nijmegen breakage syndrome[22,23]. These diseases also have widespread extra-neural effects, in contrast to diseases that result from mutations in genes involved in single-strand DNA break repair, whose effects are usually limited to the nervous system, albeit still with severe clinical outcomes[23]. Spinocerebellar ataxia with axonal neuropathy is caused by mutations in tyrosyl-DNA phosphodiesterase 1 (*TDP1)* and the recessive ataxias with oculomotor apraxia (AOA) 1, 2and 4 are caused by mutations in aprataxin (*APTX)*, senataxin (*SETX)*, and polynucleotide kinase 3'-phosphatase (*PNKP)*[24],respectively. *TDP1* repairs stalled topoisomerase I-DNA complexes, *APTX* and *PNKP*[25] operate on nucleotides and *SETX* encodes a helicase involved in transcriptional termination[26,27].

Most of these recessive diseases result in ataxia with a prominent cerebellar degeneration, also seen in the spinocerebellar ataxias caused by CAG repeat expansions, and it remains an outstanding question why this should be so. The nervous system is vulnerable to DNA damage because of its dependence on, and high levels of, oxidative metabolism, which generates free radicals with the potential to cause single-strand breaks in DNA[28,23]. Reduced capacity to repair such single-strand breaks through subtle modulation of functional activity induced by variation in genes in the DNA repair machinery might therefore lead to neuronal susceptibility. A recent novel insight from Madhabhushi et al.[29] showed that DNA damage and repair can directly affect neuronal gene expression: activity-dependent transcription of early response genes in neurons triggered the formation of Topoisomerase IIβ (TopoIIβ) double-strand DNA breaks in their promoters. These gene products, such as *c-Fos*, regulate multiple downstream pathways and influence synapses to exert downstream effects on functions such cognition, learning and memory[23]. Subtle variation in these DNA repair proteins may alter the timing or repair of double strand DNA breaks. Notably, individuals carrying mutations in tyrosyl-DNA phosphodiesterase 2 (*TDP2*) manifest with intellectual disability, epilepsy, and ataxia and the loss of TDP2, which repairs topoisomerase induced DNA breaks, leads to hypersensitivity to Topo IIβ-mediated double strand DNA breaks[30].

Conversely, DNA damage response factors can maintain appropriate neurological function and be neuroprotective. Increased DNA double-strand breaks have been linked to ageing and pathogenesis in

neurodegenerative disorders such as AD[31], and recently BRCA1, which resolves double-strand DNA breaks during homologous recombination[22] has been shown to be neuroprotective in AD mouse models [32]. This complements earlier findings in the repeat disorders. Cell models expressing mutant huntingtin (HTT) accumulate both single-strand and double-strand DNA breaks with a concomitant activation of the DNA damage response[33]. Mutant HTT binds Ku70, a core component of non-homologous end joining[22], and overexpression can rescue the phenotype in the R6/2 model of HD[34]. BRCA1 is recruited to sites of DNA damage by γ-H2AX and in HD cell lines less BRCA1 was recruited and the nuclear distribution of γ-H2AX to neuronal DNA damage was reduced: this effect was rescued by overexpression of BRCA1[35]. Both mutant HTT and ATXN1 bind high mobility group protein B (HMGB) proteins that are components of base excision repair (BER)[22,36]. In fly and mouse models of SCA1, carrying expanded repeats in *Atxn1*, neuronal pathology was rescued by expression of HMGB1, which acted to reverse mitochondrial DNA damage repair in the Atxn1-knock in mouse brain[36,37].

Genetic modifiers in the triplet repeat diseases

One way of overcoming the difficulties of interpreting the biology is to return to the study of people carrying the repeat expansions. In these natural experiments[2] it is possible to search for genetic loci that modify disease in a beneficial or deleterious way, to reveal the underlying biology likely to be important in altering the manifestation of disease: variation that renders disease onset earlier or later, or alters the progression or severity of disease is likely to lie in a biological pathway, that if manipulated using drugs, might well have a similar effect on the phenotype (disease). There are practical issues in pursuing such studies in Mendelian disease: by their nature such diseases are rare and therefore collecting sufficiently powerful samples is challenging. This is now being overcome through networks and consortia which aim to collect together large international patient cohorts, with both DNA and, critically, systematically collected clinical information, such as the Enroll study in HD (https://www.enroll-hd.org/) and the SPATAX consortium in the spinocerebellar ataxias[38,39]. Even so, in many diseases sample sizes will always be relatively small and the approach used in the pursuit of genetic loci underpinning common diseases, of collecting larger and larger samples[40], may never be possible. The recent successful search for loci that modify age at onset in HD demonstrates that this approach is feasible[18] Genetic variation that modifies rare Mendelian disease may be

common and have substantial effect sizes, as such variants may not be under population selection pressure, as in common disease, and thus be easier to find. HD is relatively common amongst rare diseases[41,42], and the collection of cohorts with DNA and systematic clinical information such as Registry[43] allowed an appropriately powered genome-wide association study to be performed. Three independent genome-wide significant loci were associated with age at motor onset, one on chromosome 8 and two close together on chromosome 15, along with a significant enrichment of signal in the network of DNA repair-related genes[18].

DNA repair related mechanisms have been implicated as modulators of somatic expansion of the disease-associated repeated sequences in mouse models of HD, DM[23,44], Fragile X[45] and Friedreich's ataxia[46]. Both the inverse correlation of age at onset – CAG repeat length and somatic expansion are widely seen in repeat diseases (Table 1). In the CAG-repeat associated spinocerebellar ataxias, testing SNPs from the DNA repair pathway genes implicated in HD[18], a significant genetic signal was observed, in the same direction as that observed in HD[19]. Thus it appears that at least some genetic modifiers in the DNA repair pathways seen in these diseases are likely to be acting at the level of the mutation type, the repeated sequence itself, rather than affecting the functions of individual repeat disease proteins. This may occur through the somatic expansion seen in many repeat disorders. In DM1, somatic expansions can be detected in blood as well as other tissues. This enabled Monckton and colleagues to identify polymorphisms in *MSH3* in a cohort of Costa Rican DM1 cases, associated with variation in somatic instability in blood, though no association with age at onset was detected[47]. Nevertheless this provides an intriguing direct link between DNA repair gene polymorphisms and somatic instability of repeats and supports mismatch repair (MMR) as central to repeat expansion.

DNA repeat expansions and the DNA damage response

Repeat expansions in DNA are affected directly by activities of the DNA damage response[48] (Box 1). The repeats undergo expansion on transmission through the germline, in both dividing and terminally differentiated somatic cells, and the repeat size increases with age[48]. Strand breakage in the repeat is repaired and it is at this point the repeat sequences are thought to expand[44,49]. The length of the repeat expansion is positively correlated with the propensity to further somatic expansion in HD[50] and the greatest

expansions occur in the striatum which has been proposed as the underlying reason that the striatum is particularly susceptible to degeneration in HD[51]. Proteins containing polyglutamine expansions can bind nuclear proteins that operate in DNA repair, such as VCP[52], raising the possibility that accumulation of the expanded polyglutamine proteins themselves induces DNA damage, which may exacerbate disease pathogenesis in a vicious cycle.

The structure of the repeats influences the likelihood of expansion. Trinucleotide repeats can adopt multiple incorrectly paired structures including hairpins, loops, triplet helices and G-quadruplexes[49,53] (Figure 1); such bulky non-B structures in the DNA may be stable at large sizes[54], and thus act as substrates for the DNA damage response. The CTG·CAG repeat sequence adopts multiple transient slipped–DNA junctions giving unpaired bases that might well be the target of DNA repair: consistent with this the prevalence of slipped strand features correlated with instability levels in DM1 tissue[55]. RNA-DNA hybrids (R-loops) formed during transcription-coupled nucleotide excision repair (NER) prevent repeat contractionand knocking down SETX, which resolves R-loops, enhances repeat instability[56]. Damage to individual bases also requires repair and oxidative damage from the generation of reactive oxygen species through mitochondrial dysfunction and exitotoxicity, seen in HD, can lead to the formation of aberrant DNA adducts and induction of the DNA damage response[16]. Such damage might also be potentiated by the expanded repeats themselves through repeat induced mutagenesis[57]. The mechanisms of germline and somatic expansion may involve different pathways as replication is also associated with repeat expansion[58]. Classically, in non-dividing cells, such as neurons, DNA repair activities and expansion are thought to be associated with DNA damage, transcription and chromatin dynamics, though recent evidence suggests that environmental stress may induce DNA rereplication and promote repeat expansion[59].

Mismatch Repair

Mismatch repair (MMR) activity on DNA modulates somatic expansion of repeat tracts[48] (Figure 1), and elements of the classical MMR pathways may also act as downstream effectors of other DNA repair mechanisms. In mammalian cells MMR is performed by two complexes: MutSα, which contains mutS

homolog 2 (MSH2) and mutS homolog 6 (MSH6) and preferentially targets mismatched bases and MutSβ which contains MSH2 and mutS homolog 3 (MSH3) and preferentially targets small deletion/insertions[22,44]. The MutS complexes recruit the endonuclease MutLα, that contains mutL homolog 1 (MLH1) and PMS1 homolog 2 (PMS2) and cleaves the DNA of the lesioned strand[44]; MutLβ (MLH1/PMS1) and MutLγ (MLH1/MLH3) can also perform this role. MutSβ can cause both somatic and intergenerational CAG·CTG repeat instability, but the evidence for MutSα is less consistent[44]. Knocking out repair genes prevents somatic expansion and ameliorates the phenotype of HD mice[60,61]. Susceptibility to somatic expansion was mapped to *Mlh1* and *Mlh3* in mouse chromosome substitution experiments[62] and HD mouse crosses in different background strains showed increased levels of MSH3 were associated with repeat expansion[63]. In cells carrying 800 CAG·CTG repeats knockdown of MSH2 and MSH3 prevented repeat expansion [56]. Similar phenomena are also apparent in mouse models of Friedreich's ataxia and Fragile X carrying GAA·TTC and CGG·CCG expansions respectively[64,65], and DNA damage response genes are downregulated in Fragile X patient blood[66]. An intriguing further finding shows that histone deacetylase (HDAC) enzymes promote repeat expansion via the MutSβ pathway[67]. CAG repeat sequences show enhanced convergent transcription[68] (transcription taking place on both strands and moving towards each other), that involves MMR components but induces cell death via ATR focus formation at CAG repeats[69]. It is therefore possible that the final pathways of neurodegeneration and cell death in the repeat diseases parallel those of other neurological diseases, such as the A-T like diseases that can be caused directly by mutations in *ATR*.

Base excision repair

The response to oxidative damage of DNA by base-excision repair influences repeat expansion (Figure 1). 8-oxoguanine glycosylase (OGG1) removes 8-oxoguanine bases from DNA, damaged through the action of reactive oxygen species. Crossing HD knock in mice with *Ogg1*$^{-/-}$ mice reduces somatic expansion of the *Htt* CAG repeat and delays phenotype, and treating mice with a reactive oxygen species scavenger to prevent DNA oxidation also reduces somatic expansion and correlates with improvement of the motor phenotype[70]. Flap endonuclease 1 (FEN1) also has a role in BER and repeat expansion[71]. During BER of 8-oxoguanine in the DNA of CAG repeats, OGG1 and mutY DNA glycosylase (MUTYH), that removes adenine incorporated

opposite unrepaired 8-oxoguanine bases, generate incisions on opposite DNA strands that may permit repeat expansion[72] though events downstream of DNA cleavage are also involved in expansions: the p53R2 protein product of *RRM2B* is induced in R6/2 HD mouse brain regions that show somatic expansion of repeats, but not in those that do not[73]. Notably *RRM2B* is in the genome-wide significant peak on chromosome 8 in the HD-GeM study[18] and has nominal associations in other repeat diseases[19].

The Fanconi Anaemia repair pathway

The operation of the Fanconi Anaemia (FA) pathway[74] in trinucleotide repeat diseases is unexplored. However, the chromosome 15 locus associated with age at onset of HD[18] contains *FAN1*. As a DNA nuclease it is a candidate for modifying HD onset through mechanisms outlined above or by promoting or ameliorating DNA expansion at the CAG repeat. FAN1 cleaves DNA at interstrand cross-links and repair occurs in complex with other FA pathway members including FANCD2 and the MMR proteins[74]. Mutations in genes associated with interstrand cross-link repair, with which FAN1 interacts, cause Fanconi anaemia, but loss of function mutations in *FAN1* lead to a recessive renal syndrome, karyomegalic interstitial nephritis[75], while heterozygous truncating mutations in *FAN1* cause some familial colorectal cancers[76], in common with other MMR pathway mutations. In addition to its known function in repairing interstrand cross-links, FAN1 recognises branched structures mimicking DNA-repair rather than specific DNA sequences[74,77]. Repeat sequences, including CAG repeats, adopt non-helical structures in DNA such as G-quadruplexes[49] and thus FAN1 may target these structures, rather than recognising the sequence itself (Figure 1). This activity is likely dependent on other MMR proteins to effect DNA repair, consistent with the manipulation of MMR genes ameliorating the HD mouse phenotypes[44]. Given that *MLH1* is in a locus on chromosome 3 that has a signal just below genome-wide significance in the HD GWAS[7,] and MLH1 interacts with FAN1, these may be central players in a novel FAN1 activity that binds repeats and modulates their instability.

Conclusions and future directions

The demonstration that genetic modifiers exist in the Mendelian triplet repeat disorders shows that finding genetic modifiers in rare genetic diseases is possible. It highlights areas of biology that modulate disease in people: in the triplet repeat disorders specific aspects of the DNA damage response are highlighted:

mismatch repair, base-excision repair and the Fanconi anaemia pathway. These recent findings raise two very interesting questions: first, does the observation of genetic association with DNA repair processes occur across all the repeat diseases, and second, is the common mechanism through which it operates somatic expansion of repeats? In order to establish the relevance of DNA-repair and other modifiers to the repeat diseases more genetic studies – including both genome-wide association and sequencing - to fully power and identify risk SNPs and loci are required, across all the repeat diseases: this is ongoing in HD. Many of these diseases have very long repeats, which are currently difficult to measure accurately but new sequencing technologies, including long read and single-cell sequencing, should overcome this and allow a more accurate determination of the exact sequence in the repeat including any interruptions[78]. The current work in HD and the polyglutamine diseases provides a model that indicates combining multiple diseases in analysis may be a possible route to increased power[19].

Elucidating the mechanistic consequences of such variation will be challenging. Many of the diseases already have models in cells and animals that can be refined with knowledge of modifiers and common biology allows development of common downstream assays that reflect disease biology demonstrated to be important in people. However, as yet there is no information about the direct molecular effects of the genetic variation detected so whether the function of the DNA damage response is enhanced or inhibited by the observed changes is unknown. As in complex diseases, clues can be gathered from aggregating data about likely expression changes and functional effects of coding changes in genes and pathways using algorithms such as co-expression networks[79], protein interaction networks[80] and developments of them.

Downstream from genetic discovery, in order to establish whether somatic expansion or other DNA repair/integrity mechanisms are responsible for the genetic signal, cell biology and animal models will be necessary. Human tissue will be needed to investigate gene expression and demonstrate the effects of the variants *in vivo*. One immediate priority, just beginning, is establishing the effects of manipulating the Fanconi anaemia pathway. Potential emerging pathogenic mechanisms that require further investigation include the operation of repeat-induced mutagenesis, where DNA repair activities lead to further

mutagenesis around repeat loci[57], and the relevance of chromatin structure to repeat exposure and dynamics[81]. Detailed mechanisms examining the exact nature of the DNA damage response essential in neurodegeneration will allow the development of new drugs and possibly the repurposing of already existing treatments targeting the DNA damage response in cancers[22].

Finally, a common mechanism in multiple diseases offers hope that treatments can be developed that will be applicable across diseases. This is analogous to the situation in cancers where common biological pathways are dysregulated in multiple forms of the disease and the same chemotherapeutic agents can be used as part of polytherapies in a number of different cancers[82]. These recent findings open a new window on repeat expansion disease, with obvious avenues for therapeutic exploitation.

Contributions

All authors contributed to the writing of the manuscript.

Declaration of interests

The authors have nothing to disclose.

Acknowledgements

There was no role for any funding source.

Box 1: ***DNA damage and the DNA damage response***

Lesions in DNA are the inevitable result of both exogenous and endogenous processes. Repairing DNA lesions, whatever their cause, is fundamental to genome integrity. Mutations induced by damaged bases, structural modifications of DNA through supercoiling and looping out of strands and interstrand cross-links all occur, and require resolution to maintain the genome: unrepaired lesions lead to cell death or uncontrolled division. Clinically, inherited lesions in the genes of the DNA damage response confer susceptibility to cancers which has recently been reviewed in detail by Pearl and colleagues[22]. The major pathways of the DNA damage response (DDR) are shown below. It should be noted that while the pathways are distinct, many of the proteins within them operate in multiple DDR pathways and it is important to bear this in mind when considering their potential effects in mediating and modulating neuropathology.

Abbreviation	Repair pathway	Type	Repair
FA	Fanconi anaemia	Double strand	Interstrand crosslinks, possibly others[74]
HR	Homologous recombination	Double strand	Template-directed end-joining from other chromatid[83]
NHEJ	Non-homologous end joining	Double strand	Ligates double-strand breaks without a template[84]
BER	Base excision repair	Single strand	Removes damaged bases, fills and ligates the strand[85]
DR	Direct repair	Single strand	Direct repair of damaged bases without removal[86]
MMR	mismatch repair	Single strand	Corrects mismatches in replication and short insertions and deletions[87]
NER	Nucleotide excision repair	Single strand	Removes DNA modifications that cause structural distortions[88]
TLS	Translesion synthesis		Synthesises DNA at sites of damage during replication[89]

Location	Disease	Gene	Repeated sequence	Non-expanded	Expanded	Somatic expansion	Comments
5' UTR	Fragile X syndrome[90,91]	*FMR1*	CGG	<50	>200	Y (M, H)	FRAXE and other rare fragile sites all CCG
	Fragile X associated tremor/ataxia[91] syndrome	*FMR1*	CGG	<50	50-200	Y (H, M)	
	Spinocerebellar ataxia 12[92]	*ATXN12*	CAG	7-28	46-78	rare	May also be exonic
	Frontotemporal dementia/Amyotrphic lateral sclerosis[93]	*C9ORF72*	GGGGCC	2-17	10-50kb	Y (H)	
Exon	Dentatorubralpallidolusian atrophy[94]	*ATN1*	CAG	<48	48+	Y (H, M)	
	Huntington's disease[48]	*HTT*	CAG	6-35	36-250	Y (H, M)	
	Huntington's disease like 2[95]	*JPH3*	CAG	6-28	40-59	Y (H)	Intronic or 3'UTR also
	Spinal and bulbar muscular atrophy[96]	*AR*	CAG	9-36	38-62	Y (H, M)	
	Spinocerebellar ataxia 1[38]	*ATXN1*	CAG	26-37	39-82	Y (H, M)	
	Spinocerebellar ataxia 2[38]	*ATXN2*	CAG	15-29	33-63	Y (H, M)	also ALS13
	Spinocerebellar ataxia 3[38]	*ATXN3*	CAG	12-35	47-84	Y (H, M)	
	Spinocerebellar ataxia 6[38]	*CACNA1A*	CAG	7-16	20-29	rare meiotic	
	Spinocerebellar ataxia 7[38]	*ATXN7*	CAG	8-14	36-62	Y (H, M)	
	Spinocerebellar ataxia 17[38]	*TBP*	CAG	29-42	43-63	meiotic	
Intron	Friedreich's ataxia[97]	*FXN*	GAA	6-30	66-1700	Y (H, M)	
	Myotonic dystrophy 2[98]	*CNBP*	CCTG	<30	55-11,000	Y (H, M)	Complex repeat structure
3' UTR	Myotonic dystrophy 1[48,98]	*DMPK*	CTG	5-37	100-2000+	Y (H, M)	
	Spinocerebellar ataxia 8[99]	*ATXN8*	(CTA)(CTG)	<33	80+	Y (H)	May also be exonic, instability in non-expanded and expanded repeats

Table 1: Characteristics of selected disease causing repeat loci

Disease causing repeat loci mentioned in the text are included thus multiple rare diseases caused by expansion of repeat codons giving rise to shorter polyalanine tracts[100] are not included. H = seen in human tissues, M = seen in mouse model tissues.

Figure legend

Figure 1 DNA repair and repeat expansion mechanisms

Mismatch repair (a), transcription-coupled repair (b), the Fanconi anaemia pathway (c) and base excision repair (d) may all have a role in repeat instability. Slipped strands of different sizes in CAG repeats , where C-G bases are Watson-Crick paired and stabilise the looped out structures but the intervening bases are not paired. Unpaired bases also occur at the ends of loop structures and at bulges in the DNA. These unpaired bases are susceptible to damage which can lead to base excision repair, illustrated in (d). TC-NER (b) can occur as the DNA strands separate for transcription and the DNA on the non-transcribed strand is unwound and exposed. Elements of the transcriptional machinery can cleave the DNA and stalled transcription promotes the formation of R-loops which predispose to repeat instability. More speculatively, FAN1, a structure specific nuclease, and possibly other elements of the Fanconi anaemia pathway (c), might recognise and bind to bulky structures formed by the repeat sequences such as G-quadruplexes leading to DNA cleavage rendering repair necessary and predisposing to repeat instability. There is currently no mechanistic work to support this hypothesis. All the DNA structures with unpaired bases are likely to have an increased propensity for DNA damage, and extrinsic and intrinsic factors such as oxidative stress which can damage DNA are likely to cause more damage. Such damage and subsequent BER (d) is known to lead to DNA repair by gap-filling synthesis and predispose to instability of repeats.

References

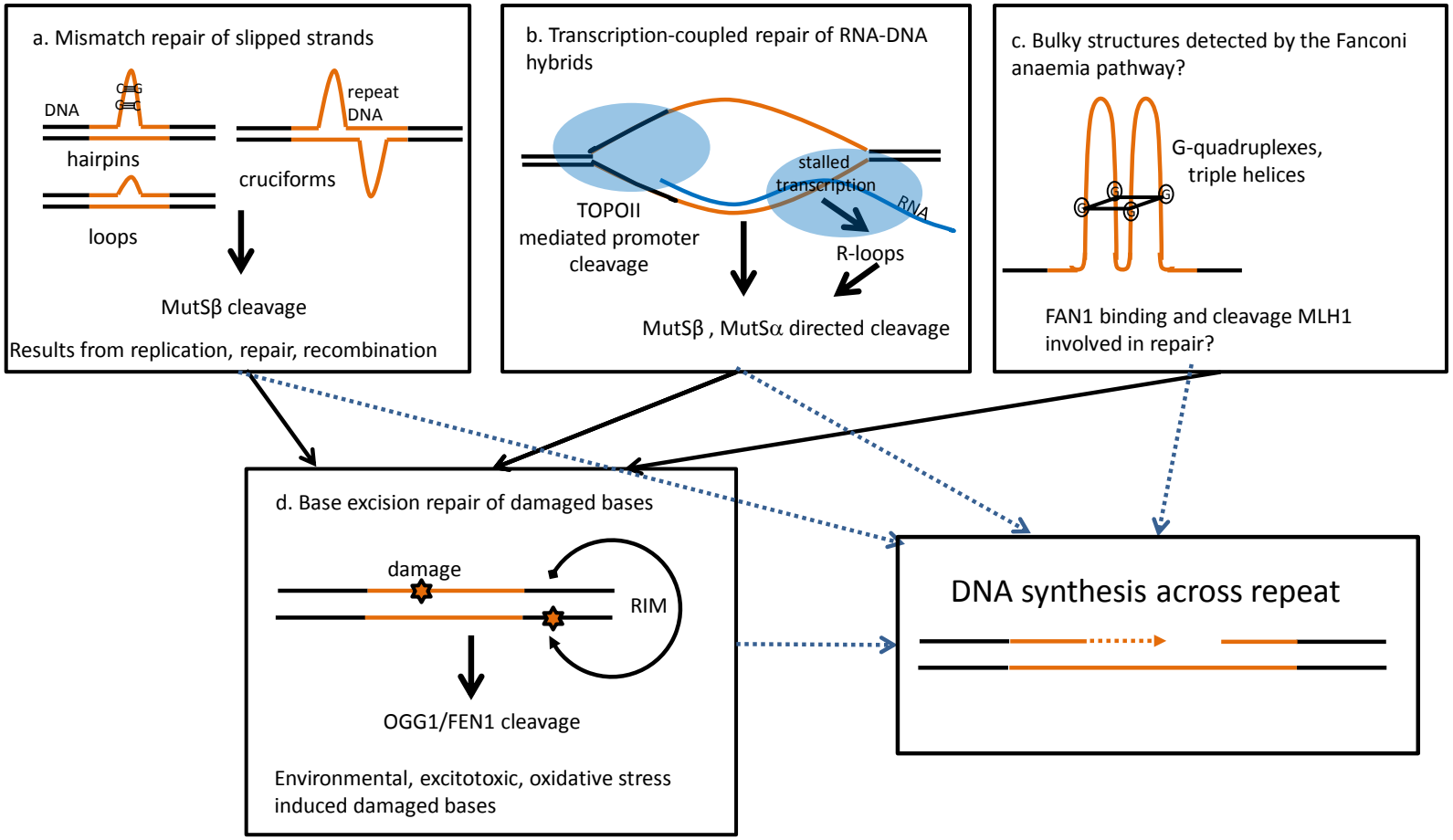
1 Orr HT, Zoghbi HY. Trinucleotide repeat disorders. *Annu Rev Neurosci* 2007; **30**: 575–621.
2 Gusella JF, MacDonald ME, Lee J-M. Genetic modifiers of Huntington's disease. *Mov Disord* 2014; **29**: 1359–65.
3 Trang H, Stanley SY, Thorner P, *et al.* Massive CAG repeat expansion and somatic instability in maternally transmitted infantile spinocerebellar ataxia type 7. *JAMA Neurol* 2015; **72**: 219–23.
4 Nageshwaran S, Festenstein R. Epigenetics and triplet repeat neurological diseases. *Front Neurol* 2015; **6**. DOI:10.3389/fneur.2015.00262.
5 Cooper TA, Wan L, Dreyfuss G. RNA and Disease. *Cell* 2009; **136**: 777–93.
6 Echeverria G V, Cooper TA. RNA-binding proteins in microsatellite expansion disorders: Mediators of RNA toxicity. *Brain Res* 2012; **1462**: 100–11.
7 Zu T, Gibbens B, Doty NS, *et al.* Non-ATG-initiated translation directed by microsatellite expansions. *Proc Natl Acad Sci U S A* 2011; **108**: 260–5.
8 Bañez-Coronel M, Ayhan F, Tarabochia AD, *et al.* RAN Translation in Huntington Disease. *Neuron* 2015; **88**: 667–77.
9 Cleary JD, Ranum LPW. Repeat associated non-ATG (RAN) translation: new starts in microsatellite expansion disorders. *Curr Opin Genet Dev* 2014; **26**: 6–15.
10 Tran H, Almeida S, Moore J, *et al.* Differential Toxicity of Nuclear RNA Foci versus Dipeptide Repeat Proteins in a Drosophila Model of C9ORF72 FTD/ALS. *Neuron* 2015; **87**: 1207–14.
11 Knowles TPJ, Vendruscolo M, Dobson CM. The amyloid state and its association with protein misfolding diseases. *Nat Rev Mol Cell Biol* 2014; **15**: 384–96.
12 Harrison JR, Owen MJ. Alzheimer's disease: the amyloid hypothesis on trial. *Br J Psychiatry* 2016; **208**: 1–3.
13 Sevigny J, Chiao P, Bussiere T, *et al.* The antibody aducanumab reduces Abeta plaques in Alzheimer's disease. *Nature* 2016; **537**: 50–6.
14 Lasagna-Reeves CA, Rousseaux MWC, Guerrero-Muñoz MJ, *et al.* A native interactor scaffolds and stabilizes toxic ATAXIN-1 oligomers in SCA1. *Elife* 2015; **4**: e07558.
15 Lasagna-Reeves CA, Rousseaux MWC, Guerrero-Munoz MJ, *et al.* Ataxin-1 oligomers induce local spread of pathology and decreasing them by passive immunization slows Spinocerebellar ataxia type 1 phenotypes. *Elife* 2015; **4**: e10891.
16 Bates GP, Dorsey R, Gusella JF, *et al.* Huntington disease. *Nat Rev Dis Prim* 2015; **1**: 15005.
17 Guo JL, Lee VMY. Cell-to-cell transmission of pathogenic proteins in neurodegenerative diseases. *Nat Med* 2014; **20**: 130–8.
18 Huntington's GM of, Consortium D (GeM-H. Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease. *Cell* 2015; **162**: 516–26.
19 Bettencourt C, Moss DH, Flower M, *et al.* DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. *Ann Neurol* 2016; published online April 4. DOI:10.1002/ana.24656.
20 Martins S, Pearson CE, Coutinho P, *et al.* Modifiers of (CAG)n instability in Machado--Joseph disease (MJD/SCA3) transmissions: an association study with DNA replication, repair and recombination genes. *Hum Genet* 2014; **133**: 1311–8.
21 Paull TT. Mechanisms of ATM Activation. *Annu Rev Biochem* 2015; **84**: 711–38.
22 Pearl LH, Schierz AC, Ward SE, Al-Lazikani B, Pearl FMG. Therapeutic opportunities within the DNA damage response. *Nat Rev Cancer* 2015; **15**: 166–80.
23 McKinnon PJ. DNA repair deficiency and neurological disease. *Nat Rev Neurosci* 2009; **10**: 100–12.
24 Bras J, Alonso I, Barbot C, *et al.* Mutations in PNKP Cause Recessive Ataxia with Oculomotor Apraxia Type 4. *Am J Hum Genet* 2015; **96**: 474–9.
25 Weinfeld M, Mani RS, Abdou I, Aceytuno RD, Glover JNM. Tidying up loose ends: the role of polynucleotide kinase/phosphatase in DNA strand break repair. *Trends Biochem Sci* 2011; **36**: 262–71.
26 Hatchi E, Skourti-Stathaki K, Ventz S, *et al.* BRCA1 Recruitment to Transcriptional Pause Sites Is Required for R-Loop-Driven DNA Damage Repair. *Mol Cell* 2015; **57**: 636–47.
27 Yüce Ö, West SC. Senataxin, Defective in the Neurodegenerative Disorder Ataxia with Oculomotor

Apraxia 2, Lies at the Interface of Transcription and the DNA Damage Response. *Mol Cell Biol* 2013; **33**: 406–17.
28 Barzilai A. The Contribution of the DNA Damage Response to Neuronal Viability. *Antioxid Redox Signal* 2006; **9**: 211–8.
29 Madabhushi R, Gao F, Pfenning AR, *et al.* Activity-Induced DNA Breaks Govern the Expression of Neuronal Early-Response Genes. *Cell* 2015; **161**: 1592–605.
30 Gomez-Herreros F, Schuurs-Hoeijmakers JHM, McCormack M, *et al.* TDP2 protects transcription from abortive topoisomerase activity and is required for normal neural function. *Nat Genet* 2014; **46**: 516–21.
31 Suberbielle E, Sanchez PE, Kravitz A V, *et al.* Physiologic brain activity causes DNA double-strand breaks in neurons, with exacerbation by amyloid-beta. *Nat Neurosci* 2013; **16**: 613–21.
32 Suberbielle E, Djukic B, Evans M, *et al.* DNA repair factor BRCA1 depletion occurs in Alzheimer brains and impairs cognitive function in mice. *Nat Commun* 2015; **6**: 8897.
33 Illuzzi J, Yerkes S, Parekh-Olmedo H, Kmiec EB. DNA breakage and induction of DNA damage response proteins precede the appearance of visible mutant huntingtin aggregates. *J Neurosci Res* 2009; **87**: 733–47.
34 Enokido Y, Tamura T, Ito H, *et al.* Mutant huntingtin impairs Ku70-mediated DNA repair. *J Cell Biol* 2010; **189**: 425–43.
35 Jeon GS, Kim KY, Hwang YJ, *et al.* Deregulation of BRCA1 leads to impaired spatiotemporal dynamics of gamma-H2AX and DNA damage responses in Huntington's disease. *Mol Neurobiol* 2012; **45**: 550–63.
36 Qi M-L, Tagawa K, Enokido Y, *et al.* Proteome analysis of soluble nuclear proteins reveals that HMGB1/2 suppress genotoxic stress in polyglutamine diseases. *Nat Cell Biol* 2007; **9**: 402–14.
37 Ito H, Fujita K, Tagawa K, *et al.* HMGB1 facilitates repair of mitochondrial DNA damage and extends the lifespan of mutant ataxin-1 knock-in mice. *EMBO Mol Med* 2015; **7**: 78–101.
38 Tezenas du Montcel S, Durr A, Bauer P, *et al.* Modulation of the age at onset in spinocerebellar ataxia by CAG tracts in various genes. *Brain* 2014; **137**: 2444–55.
39 Jacobi H, du Montcel ST, Bauer P, *et al.* Long-term disease progression in spinocerebellar ataxia types 1, 2, 3, and 6: a longitudinal cohort study. *Lancet Neurol* 2015; **14**: 1101–8.
40 Spencer CCA, Su Z, Donnelly P, Marchini J. Designing Genome-Wide Association Studies: Sample Size, Power, Imputation, and the Choice of Genotyping Chip. *PLoS Genet* 2009; **5**: e1000477.
41 Wexler NS, Collett L, Wexler AR, *et al.* Incidence of adult Huntington's disease in the UK: a UK-based primary care study and a systematic review. *BMJ Open* 2016; **6**. DOI:10.1136/bmjopen-2015-009070.
42 Evans SJW, Douglas I, Rawlins MD, Wexler NS, Tabrizi SJ, Smeeth L. Prevalence of adult Huntington's disease in the UK based on diagnoses recorded in general practice records. *J Neurol Neurosurg Psychiatry* 2013; **84**: 1156–60.
43 Orth M, Network TEH 's D. Observing Huntington's disease: the European Huntington's Disease Network's REGISTRY. *J Neurol Neurosurg Psychiatry* 2011; **82**: 1409–12.
44 Iyer RR, Pluciennik A, Napierala M, Wells RD. DNA Triplet Repeat Expansion and Mismatch Repair. *Annu Rev Biochem* 2015; **84**: 199–226.
45 Lokanga RA, Zhao X-N, Usdin K. The mismatch repair protein MSH2 is rate limiting for repeat expansion in a fragile X premutation mouse model. *Hum Mutat* 2014; **35**: 129–36.
46 Ezzatizadeh V, Pinto RM, Sandi C, *et al.* The mismatch repair system protects against intergenerational GAA repeat instability in a Friedreich ataxia mouse model. *Neurobiol Dis* 2012; **46**: 165–71.
47 Morales F, Vásquez M, Santamaría C, Cuenca P, Corrales E, Monckton DG. A polymorphism in the MSH3 mismatch repair gene is associated with the levels of somatic instability of the expanded CTG repeat in the blood DNA of myotonic dystrophy type 1 patients. *DNA Repair (Amst)* 2016; **40**: 57–66.
48 Schmidt MHM, Pearson CE. Disease-associated repeat instability and mismatch repair. *DNA Repair (Amst)* 2016; **38**: 117–26.
49 McMurray CT. Mechanisms of trinucleotide repeat instability during human development. *Nat Rev Genet* 2010; **11**: 786–99.
50 Veitch NJ, Ennis M, McAbney JP, Shelbourne PF, Monckton DG. Inherited CAG·CTG allele length is a

major modifier of somatic mutation length variability in Huntington disease. *DNA Repair (Amst)* 2007; **6**: 789–96.
51 Shelbourne PF, Keller-McGandy C, Bi WL, *et al.* Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain. *Hum Mol Genet* 2007; **16**: 1133–42.
52 Fujita K, Nakamura Y, Oka T, *et al.* A functional deficiency of TERA/VCP/p97 contributes to impaired DNA repair in multiple polyglutamine diseases. *Nat Commun* 2013; **4**: 1816.
53 Mirkin SM. Expandable DNA repeats and human disease. *Nature* 2007; **447**: 932–40.
54 Volker J, Plum GE, Gindikin V, Klump HH, Breslauer KJ. Impact of bulge loop size on DNA triplet repeat domains: Implications for DNA repair and expansion. *Biopolymers* 2014; **101**: 1–12.
55 Axford MM, Wang Y-H, Nakamori M, Zannis-Hadjopoulos M, Thornton CA, Pearson CE. Detection of slipped-DNAs at the trinucleotide repeats of the myotonic dystrophy type I disease locus in patient tissues. *PLoS Genet* 2013; **9**: e1003866.
56 Nakatani R, Nakamori M, Fujimura H, Mochizuki H, Takahashi MP. Large expansion of CTG*CAG repeats is exacerbated by MutSbeta in human cells. *Sci Rep* 2015; **5**: 11020.
57 Shah KA, Mirkin SM. The hidden side of unstable DNA repeats: Mutagenesis at a distance. *DNA Repair (Amst)* 2015; **32**: 106–12.
58 Iyama T, Wilson DM 3rd. DNA repair mechanisms in dividing and non-dividing cells. *DNA Repair (Amst)* 2013; **12**: 620–36.
59 Chatterjee N, Lin Y, Santillan BA, Yotnda P, Wilson JH. Environmental stress induces trinucleotide repeat mutagenesis in human cells. *Proc Natl Acad Sci U S A* 2015; **112**: 3764–9.
60 Wheeler VC, Lebel L-A, Vrbanac V, Teed A, te Riele H, MacDonald ME. Mismatch repair gene Msh2 modifies the timing of early disease in HdhQ111 striatum. *Hum Mol Genet* 2003; **12**: 273–81.
61 Dragileva E, Hendricks A, Teed A, *et al.* Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. *Neurobiol Dis* 2009; **33**: 37–47.
62 Pinto RM, Dragileva E, Kirby A, *et al.* Mismatch Repair Genes <italic>Mlh1</italic> and <italic>Mlh3</italic> Modify CAG Instability in Huntington's Disease Mice: Genome-Wide and Candidate Approaches. *PLoS Genet* 2013; **9**: e1003930.
63 Tomé S, Manley K, Simard JP, *et al.* MSH3 Polymorphisms and Protein Levels Affect CAG Repeat Instability in Huntington's Disease Mice. *PLoS Genet* 2013; **9**: 1–16.
64 Bourn RL, De Biase I, Pinto RM, *et al.* Pms2 suppresses large expansions of the (GAA.TTC)n sequence in neuronal tissues. *PLoS One* 2012; **7**: e47085.
65 Zhao X-N, Kumari D, Gupta S, *et al.* Mutsbeta generates both expansions and contractions in a mouse model of the Fragile X-associated disorders. *Hum Mol Genet* 2015; **24**: 7087–96.
66 Xu H, Rosales-Reynoso MA, Barros-Nunez P, Peprah E. DNA repair/replication transcripts are down regulated in patients with Fragile X Syndrome. *BMC Res Notes* 2013; **6**: 90.
67 Gannon A-MM, Frizzell A, Healy E, Lahue RS. MutSbeta and histone deacetylase complexes promote expansions of trinucleotide repeats in human cells. *Nucleic Acids Res* 2012; **40**: 10324–33.
68 Lin Y, Wilson JH. Nucleotide excision repair, mismatch repair, and R-loops modulate convergent transcription-induced cell death and repeat instability. *PLoS One* 2012; **7**: e46807.
69 Chatterjee N, Lin Y, Wilson JH. Mismatch repair enhances convergent transcription-induced cell death at trinucleotide repeats by activating ATR. *DNA Repair (Amst)* 2016; **42**: 26–32.
70 Budworth H, Harris FR, Williams P, *et al.* Suppression of Somatic Expansion Delays the Onset of Pathophysiology in a Mouse Model of Huntington?s Disease. *PLoS Genet* 2015; **11**: e1005267.
71 Liu Y, Wilson SH. DNA base excision repair: a mechanism of trinucleotide repeat expansion. *Trends Biochem Sci* 2012; **37**: 162–72.
72 Cilli P, Ventura I, Minoprio A, *et al.* Oxidized dNTPs and the OGG1 and MUTYH DNA glycosylases combine to induce CAG/CTG repeat instability. *Nucleic Acids Res* 2016; **44**: 5190–203.
73 Lokanga RA, Senejani AG, Sweasy JB, Usdin K. Heterozygosity for a hypomorphic Polbeta mutation reduces the expansion frequency in a mouse model of the Fragile X-related disorders. *PLoS Genet* 2015; **11**: e1005181.
74 Ceccaldi R, Sarangi P, D'Andrea AD. The Fanconi anaemia pathway: new players and new functions. *Nat Rev Mol Cell Biol* 2016; **17**: 337–49.

75 Zhou W, Otto EA, Cluckey A, *et al.* FAN1 mutations cause karyomegalic interstitial nephritis, linking chronic kidney failure to defective DNA damage repair. *Nat Genet* 2012; **44**: 910–5.
76 Segui N, Mina LB, Lazaro C, *et al.* Germline Mutations in FAN1 Cause Hereditary Colorectal Cancer by Impairing DNA Repair. *Gastroenterology* 2015; **149**: 563–6.
77 Pennell S, Declais A-C, Li J, *et al.* FAN1 activity on asymmetric repair intermediates is mediated by an atypical monomeric virus-type replication-repair nuclease domain. *Cell Rep* 2014; **8**: 84–93.
78 McFarland KN, Liu J, Landrian I, *et al.* SMRT Sequencing of Long Tandem Nucleotide Repeats in SCA10 Reveals Unique Insight of Repeat Expansion Structure. *PLoS One* 2015; **10**: e0135906.
79 Parikshak NN, Gandal MJ, Geschwind DH. Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders. *Nat Rev Genet* 2015; **16**: 441–58.
80 Szklarczyk D, Franceschini A, Wyder S, *et al.* STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 2015; **43**: D447–52.
81 Viterbo D, Michoud G, Mosbach V, Dujon B, Richard G-F. Replication stalling and heteroduplex formation within CAG/CTG trinucleotide repeats by mismatch repair. *DNA Repair (Amst)* 2016; **42**: 94–106.
82 Longley DB, Harkin DP, Johnston PG. 5-Fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer* 2003; **3**: 330–8.
83 San Filippo J, Sung P, Klein H. Mechanism of eukaryotic homologous recombination. *Annu Rev Biochem* 2008; **77**: 229–57.
84 Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* 2010; **79**: 181–211.
85 Robertson AB, Klungland A, Rognes T, Leiros I. DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cell Mol Life Sci* 2009; **66**: 981–93.
86 Eker APM, Quayle C, Chaves I, van der Horst GTJ. DNA repair in mammalian cells: Direct DNA damage reversal: elegant solutions for nasty problems. *Cell Mol Life Sci* 2009; **66**: 968–80.
87 Kunkel TA, Erie DA. DNA mismatch repair. *Annu Rev Biochem* 2005; **74**: 681–710.
88 de Laat WL, Jaspers NGJ, Hoeijmakers JHJ. Molecular mechanism of nucleotide excision repair. *Genes Dev* 1999; **13**: 768–85.
89 Waters LS, Minesinger BK, Wiltrout ME, D'Souza S, Woodruff R V, Walker GC. Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. *Microbiol Mol Biol Rev* 2009; **73**: 134–54.
90 Metsu S, Rooms L, Rainger J, *et al.* FRA2A Is a CGG Repeat Expansion Associated with Silencing of <italic>AFF3</italic>. *PLoS Genet* 2014; **10**: e1004242.
91 Lokanga RA, Entezam A, Kumari D, *et al.* Somatic Expansion in Mouse and Human Carriers of Fragile X Premutation Alleles. *Hum Mutat* 2013; **34**: 157–66.
92 Dong Y, Wu J-J, Wu Z-Y. Identification of 46 CAG repeats within PPP2R2B as probably the shortest pathogenic allele for SCA12. *Parkinsonism Relat Disord* 2015; **21**: 398–401.
93 van Blitterswijk M, DeJesus-Hernandez M, Niemantsverdriet E, *et al.* Association between repeat sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): a cross-sectional cohort study. *Lancet Neurol* 2013; **12**: 978–88.
94 Tsuji S. Dentatorubral-pallidoluysian atrophy. *Handb Clin Neurol* 2012; **103**: 587–94.
95 Todd PK, Paulson HL. RNA-mediated neurodegeneration in repeat expansion disorders. *Ann Neurol* 2010; **67**: 291–300.
96 Tanaka F, Reeves MF, Ito Y, *et al.* Tissue-Specific Somatic Mosaicism in Spinal and Bulbar Muscular Atrophy Is Dependent on CAG-Repeat Length and Androgen Receptor–Gene Expression Level. *Am J Hum Genet* 1999; **65**: 966–73.
97 Evans-Galea M V, Lockhart PJ, Galea CA, Hannan AJ, Delatycki MB. Beyond loss of frataxin: the complex molecular pathology of Friedreich ataxia. *Discov Med* 2014; **17**: 25–35.
98 Udd B, Krahe R. The myotonic dystrophies: molecular, clinical, and therapeutic challenges. *Lancet Neurol* 2012; **11**: 891–905.
99 Martins S, Seixas AI, Magalhães P, Coutinho P, Sequeiros J, Silveira I. Haplotype diversity and somatic instability in normal and expanded SCA8 alleles. *Am J Med Genet Part B Neuropsychiatr Genet* 2005; **139B**: 109–14.

100 Shoubridge C, Gecz J. Polyalanine tract disorders and neurocognitive phenotypes. In: Hannan AJ, ed. Tandem Repeat Polymorphisms: Genetic Plasticity, Neural Diversity and Disease. Landes Bioscience and Springer Science+Business Media., 2013.


a. Mismatch repair of slipped strands
DNA
CAG
GTC
repeat DNA
hairpins
cruciforms
loops
MutSβ cleavage
Results from replication, repair, recombination
b. Transcription-coupled repair of RNA-DNA hybrids
stalled transcription
RNA
TOPOII mediated promoter cleavage
R-loops
MutSβ , MutSα directed cleavage
c. Bulky structures detected by the Fanconi anaemia pathway?
G-quadruplexes, triple helices
FAN1 binding and cleavage MLH1 involved in repair?
d. Base excision repair of damaged bases
damage
RIM
OGG1/FEN1 cleavage
Environmental, excitotoxic, oxidative stress induced damaged bases
DNA synthesis across repeat