UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE

IGBMC – CNRS UMR 7104 – Inserm U 964

THÈSE présentée par :

Agathi-Vasiliki GOULA

soutenue le : **26 janvier 2012**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Aspects moléculaires et cellulaires de la biologie

Implication des lésions oxydantes et du mécanisme de réparation par excision de base dans la sélectivité tissulaire de l'instabilité somatique des répétitions CAG dans la maladie de Huntington

THÈSE dirigée par :

M Yvon Trottier	Directeur de recherche, Université de Strasbourg
Mme Karine Merienne	Chargée de recherche, Université de Strasbourg

RAPPORTEURS :

Mme Geneviève Gourdon	Directeur de recherche, Université Paris Descartes
M J. Pablo Radicella	Directeur de recherche, CEA

AUTRES MEMBRES DU JURY :

M San Martin Pablo Reina	Directeur de recherche, Université de Strasbourg
Mme Sandrine Humbert	Directeur de recherche, Institut Curie

THESIS

Presented by

Agathi-Vasiliki Goula

To obtain the grade of

Doctor of Philosophy of the Université de Strasbourg

Discipline: **Life sciences**

Specialization: **Molecular and cellular aspects of biology**

Implication of oxidative lesions and base excision repair in the tissue selectivity of the somatic instability of CAG repeats in Huntington's disease

Thesis defended publicly on 26 January 2012

Thesis Director: Dr. Yvon Trottier
Thesis Director: Dr. Karine Merienne
Reporters: Dr. Geneviève Gourdon
Dr. Pablo Radicella
Internal examiner: Dr. Bernardo Reina-San-Martin
External examiner: Dr. Sandrine Humbert

"Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less."

"One never notices what has been done; one can only see what remains to be done."
Marie Curie

To my parents,

Acknowledgments

I would like to thank the members of my thesis jury: Geneviève Gourdon, Pablo Radicella, Sandrine Humbert and Bernardo Reina-San-Martin for accepting to read and grade my thesis and for accepting to attend my thesis defense in Strasbourg.

I would like to express my gratitude to my supervisors Yvon Trottier and Karine Merienne, for giving me the opportunity to work with them and for all the stimulating discussions. Thank you for all your help and guidance during the duration of my PhD. In particular, I would like to thank Karine for working closely with me and for always being available to discuss both experimental and scientific details, for her support during my PhD and motivation during thesis writing. A big thank you to Yvon for welcoming me to his team and for his help during the time I have spent in his laboratory, for his support and for the advices he has given me.

I would also like to thank Prof. Jean-Louis Mandel for supporting this project and my second trip to NIH, and especially for his support during the write-up of this thesis. I deeply appreciate your support.

I would like to thank all members of the team past and present: Fabrice for the very interesting scientific and not discussions, Chantal for the help with mice breading and cell culture and her contribution to the Parp project, Alice and Hubert for their positive outlook. I thank all of you for the nice moments inside and outside lab, special thanks to Yvon and the lab members for the nice moments we spent together at "snow lab" edition 2009 and 2010.

Thanks to Alain Martelli and Fred Coin for having initiated me and helped me with the radioactivity world. I would also like to thank the radioprotection officers Gabrielle Mengus, Dominique Biellman and James Stevenin for their useful radioadvice and help.

I would like to thank all the people of the Human Genetics lab, past and present members, for always being available to help and share knowledge and cakes ☺. Thanks to Karim, Eric, Laurence, Christine, Patricia, Solange, Mirna, Lama, Olga for help, very interesting discussions and nice moments during PhD period inside or outside lab. Thanks Michel and Moez for the after-congress times and unending discussions☺.

Thanks to Vilhelm Bohr for accepting me working twice in his unit at NIH, for his availability and encouragements. I would especially like to thank David Wilson for accepting me in his lab twice, for his enthusiasm, availability, constant motivation and energy. Many thanks to Brian, Dan and Daemyung for their help with the lab projects, their scientific tips and the nicely disturbing discussions during incubation times ☺. I also thank Al, the radioactivity officer, for his help and advices and I apologize for "shocking" him when he

was discovering at the radioactivity room cameras my samples amount☺. Many thanks as well to all the LMG members who helped me get my bearings in the lab very fast, gave me rides to the lab during the weekends and for the moments we shared around one or more beers and dinners. Thanks to James the NIH driver that was taking me back home from the lab often at nights when there was no bus anymore. Thanks you all for making my stays at the US unforgettable experiences.

I thank Valerie Schreiber for her availability for the Parp1 project. I thank Christopher Pearson for his enthusiasm, encouragements, and help with trinucleotide repeats projects.

I would also like to thank Olga, Fabrice, Farid, Laurent and Thanasis for tracking errors and helping me with the manuscript editing and for their substantial support during this home stretch.

I also thank Angela G, Evi S, Alkimini, Celine D, Peggy, Bahar, Farid, Berra, Valeria, Christine F, Celine Gr., Anne-Solène, Shankar, Mannu etc for the very interesting scientific and not discussions and all the nice times we shared.

To all my friends that were by my side during all the PhD and previous years, with each one of whom we shared very special moments and helped me to keep going on: Elena, Irini H, Irini M, Elodie, Corinne, Ann, Regis, Othonas, Ahilleas, Alexis, Filipa, Jorge, Sven, thank you for so many years of friendship despite distance and thank you for your support. Thanks to the "jeuns" group: David, Nico, Laurent, and also Christelle, ManuD, Pierre and Claude for all the deliriums and their constant motivation and ManuP for having first forward me this PhD offer. Special thanks to Virginia with whom our almost everyday communications during those years was really important to me and to her, and helped me to put the things into perspective and keep going on.

I thank Thanasis for his constant motivation, understanding and help especially during the long lab-nights, that was really important to me. I am sure you are able now to undergo a biology PhD without any undergraduate studies‼!

Many thanks to my aunts' family: Sossana, Bernard, Paul, Jean and Ilia, for all their help and support all those years, you were like a second family to me. I thank my grandma for her enthusiasm and help that was really important to me.

I thank my family, my parents and my brother for having allowed me to undertake those studies, for their support and trust to my skills. I thank my parents who participated at doing me the person I am and who taught me some basic but very important principles: integrity, respect of others and the importance of improving the everyday life of patients' as well as healthy people. I especially thank and think at my father who passed away and who keeps influencing all my life.

SUMMARY

List of Tables and Figures

List of Tables

List of Figures

Abbreviations

AAG: 3-methyladenine DNA glycosylase
AP sites: apuric/apyrimidic
APE1: AP-endonuclease 1
ATP: Adenosine Triphosphate
BRCA1/BARD1: Breast cancer 1/BRCA1-associated RING domain protein 1
CBP: CREB-binding protein
cdc9: cell division cycle 9
cDNA: complementary DNA
CELF: CUG-BP- and ETR-3-like factors
CHO cells: Chinese hamster ovary cells
CREB: CRE-binding protein
CSB: CS complementation group B
CTCF: CCCTC-binding factor
CUGBP1: CUG binding protein 1
D. melanogaster: Drosophila melanogaster
DM1: myotonic dystrophy type 1
DM2: myotonic dystrophy type 2
DMPK: Dystrophy Myotonic Protein Kinase
Dnmt1: DNA (cytosine-5)-methyltransferase 1
DRPLA: Dentatorubral-pallidoluysian atrophy
DSBR: Double strand breaks repair
E.coli: Escherichia coli
FAPY: formamidopyrimidine
FapyA: formamide pyrimidine adenine
FapyG: formamide pyrimidine guanine
FEN1: Flap endonuclease 1
FMR1: fragile X mental retardation 1
Fpg: formamidopyrimidine-DNA glycosylase
FRAXA: Fragile X A form
FRDA: Friedreich ataxia
FXN: frataxin
FXS: Fragile X syndrome
FXTAS*:* fragile X tremor ataxia syndrome
GG-NER: global genome NER
H3, H4: Histones 3, 4
H3K36(or 4)me3: Trimethylation of Lysine36(or 4) in histone H3

H3K9: Histone H3 lysine 9
HAP1: Huntingtin-associated protein 1
HAT: histone acetyltransferase
HD: Huntington's disease
HDAC: Histone deacetylases
HDL2: Huntington's disease-like 2
HeLa cells: Henrietta Lacks cells
HMGB1: high-mobility group box 1 protein
hOGG1: human OGG1
HOXA13: Homeobox protein Hox-A13
HOXD13: Homeobox protein Hox-D13
HPRT: hypoxanthine-guanine phos- phoribosyltransferase
HR: homologous recombination
Hsp70: Heat shock protein 70
Huntingtin: HTT
ID: insertion/deletion loops
Lig1/: DNA Ligase I
Lig3: DNA Ligase III
LP-BER: Long-Patch BER
MBNL: Muscleblind-like
MMR: Mismatch repair
mRNA: messenger Ribonucleic acid
Msh2: mutS homolog 2
Msh3: mutS homolog 3
Msh6: mutS homolog 6
MutSα: Msh2/Msh6 complex
MutSβ: Msh2/Msh3 complex
MYH: MutY glycosylase homologue
NEIL1 and NEIL2: endonucleases VIII-like1 and 2
NHEJ: non homologous end-joining
NTH1: endonuclease III
OGG1: 8-oxoguanine DNA glycosylase
OPMD: oculopharyngeal muscular dystrophy
PAR: poly-ADP ribose polymer
PARP-1: Poly [ADP-ribose] polymerase 1
PARP-2: Poly [ADP-ribose] polymerase 2
PCNA : Proliferating *cell nuclear antigen*
Pol α, β, δ, ε, η, θ, ι, λ: DNA polymerase alpha, beta, delta, eta, theta, iota, lamda

polyA: polyalanine
polyQ: polyglutamine
Rad2, 23B, 27, 51, 52: radiation mutant 2, 23B, 27, 51, 52
RAN : Non-ATG initiation of translation
rev7: reversion mutant 7
R-loops: DNA:RNA hybrids
RNA Pol II (or RNAP II): RNA polymerase II
ROS: reactive oxygen species
S and G2 phases: synthesis and gap2 phases
SBMA: Spinal Bulbar Muscular Atrophy
SCA 1, 2, 3, 6, 7, 8, 17: Spinocerebellar ataxias 1, 2, 3, 6, 7, 8, 17
siRNA: small interfering RNA
SIRT1: sirtuin (silent mating type information regulation 2 homolog) 1
SN-BER: single nucleotide BER
SSB: Single-strand break
SV40ori/EPR: a nucleosome excluding sequence
TC-NER: Transcription coupled NER
Tet-ON: tetracycline inducible expression system
Tg: thymidine glycol
THF: tetrahydrofuran
TNR: trinucleotide repeat
UDG: Uracil-DNA glycosylase
UV-DDB: DDB1-DDB2-containing E3-ubiquitin ligase complex
UvrA: UV irradiation A
UvrB : UV irradiation B
XPA, XPC, XPG: Xeroderma Pigmentosum A, C, G
XRCC1: X-ray repair complementing defective repair in Chinese hamster cells 1
2-meC: 2-methylcytosine
3meA: 3-methyladenine
5'dRP: 5'-deoxyribose phosphate
5OHC: 5-hydroxycytosine
5OHU: 5-hydroxyuracile 5'-OH-U
7-meG: 7-methylguanine
8-oxodA: 8-oxodesoxyadenine
8-oxodG: 8-oxodesoxyguanine
9-1-1: Rad9-Rad1-Hus1

I) Introduction

1 Repeat associated diseases

Diseases caused by unstable repeated DNA sequences represent a group of more than 40 inherited neurological, neurodegenerative or neuromuscular disorders. The diseases are caused by the abnormal expansion of repeats of varying sequence and size, depending on the disorder considered. Indeed, penta-, tetra- or tri-nucleotides, or even mini- or megasatellites have been implicated in diseases. In addition, those abnormal repeats can be located in different positions within genes, including 5' or 3' untranslated regions (5' or 3' UTR), introns and exons. These various mutations result in different pathomechanisms, ranging from loss-of-function mechanisms, due to silencing of the mutated gene (e.g. Fragile X syndrome (FXS) or Friedreich ataxia (FRDA) to gain-of-function mechanisms, due to the production of toxic mRNA (e.g. myotonic dystrophy type 1; DM1) or toxic mutant protein (e.g. Huntington's disease; HD) (Lopez Castel et al., 2010; McMurray, 2010; Pearson et al., 2005). Table 1 recapitulates the main repeat associated disorders and their mutation characteristics (Guyenet SJ, 2005; Lopez Castel et al., 2010; Pearson et al., 2005).

1.1 Trinucleotide repeat diseases

The most common subgroup of repeat-associated diseases is the trinucleotide repeat (TNR) group, which accounts for more than 20 different diseases. The repetitive sequence consists of CAG, CTG, GAA, CGG, or GAC trinucleotidic units, some examples are listed hereafter.

The larger family of TNR disorders is the so-called group of polyglutamine (polyQ) diseases, which are caused by expansion of CAG repeats in the coding region of the affected genes, thereby encoding for a polyQ expansion in the corresponding proteins. To date, nine polyQ diseases have been identified, including several dominant spinocerebellar ataxias (SCA 1,

2, 3, 6, 7 and 17), Spinal Bulbar Muscular Atrophy (SBMA), DentatoRubral-PallidoLuysian atrophy (DRPLA), and HD, the most frequent disease within the polyQ disorder group, with a prevalence of 1/10000 (Roos, 2010).

CTG expansion at the 3' UTR of the DMPK gene causes DM1, a disease characterized by a 1/8000 overall worldwide prevalence (Suominen et al., 2011). CTG repeats located in 3' UTR regions also lead to Huntington's disease-like 2 (HDL2) and SCA 8. In both diseases, bidirectional expression of the CAG/CTG repeat expansion has been observed in patients, resulting in expression of a non-translated CUG-containing mRNA and a polyQ-containing protein or peptide. Both toxic entities are susceptible to contribute to pathogeneses (Ikeda et al., 2008; Moseley et al., 2006; Wilburn et al., 2011).

Expansion of GAA repeats in intron-1 of the frataxin gene causes FRDA, the most common recessive ataxia, with a prevalence estimated to 1/50.000 in Caucasian populations (Cossee et al., 1997). Expansion of CGG repeats in the 5' UTR of the FRAXA locus is involved in FXS with a prevalence of 1/2500 in males (Hagerman, 2008), but also in fragile X tremor ataxia syndrome (FXTAS), depending on the repeat length (Cohen et al., 2006). FXS is one of the most common X-linked disease and one of the first trinucleotide mutations identified at the *FMR1* locus with pre/postnatal genetic testing available (Arveiler et al., 1988; Heilig et al., 1988; Oberle et al., 1991; Rousseau et al., 1991).

1.2 Toxicity of TNR

Repeat expansions result in gain- or loss-of-function mechanisms. In HD and other polyQ disorders, the polyQ expansion confers to mutant proteins a toxic gain of function. PolyQ proteins form toxic aggregates, sequestering specific proteins and impairing specific cellular pathways (Hands and Wyttenbach, 2011). Moreover, the aggregation of polyQ proteins induces a toxic cellular stress (Bertoni et al., 2011; Cowan et al., 2003; Merienne et al., 2003). Finally, the polyQ expansion may modify the affinity for some protein partners, affecting specific

pathways (Gil and Rego, 2008; Han et al., 2010). For instance, in HD the mutant protein -called Huntingtin (HTT)- is known to form large insoluble nuclear aggregates where some transcription factors, including CREB-binding protein (CBP), are recruited, resulting in transcriptional alteration of specific genes (Nucifora et al., 2001). Additionaly, mutant polyQ may influence binding of these proteins to partners. For instance, HTT was described to interact with HAP1 to promote axonal transport (Gunawardena et al., 2003; Li et al., 1995), disruption of the efficiency of which in the case of HD, is probably caused by modification of this interaction by the mutation, leading probably to the altered transport and neuropil aggregates observed in HD models (Gauthier et al., 2004; Gunawardena and Goldstein, 2005; Zala et al., 2008). Recently, it has proposed that the observed variations of interaction of HTT with its partners might be a consequence of the aggregation properties of the mutant protein (Davranche et al., 2011).

Moreover, expanded RNA can also be a toxic entity. RNA associated gain-of-function has been largely described in DM1. The mutant *DMPK* mRNA with CUG expansions forms RNA foci, which aberrantly interact with RNA-binding proteins such as Muscleblind-like (MBNL). This in turn results in upregulation of CELF proteins, including CUG binding protein 1 (CUGBP1). These alterations lead to splicing and transcription dysregulation in DMPK expressing cells (O'Rourke and Swanson, 2009; Ranum and Cooper, 2006; Wheeler and Thornton, 2007). Upregulation of MBNL was also reported for a SCA3 Drosophila model, where RNA with toxic CAG repeats conferred neuronal degeneration (Li et al., 2008). Flies expressing toxic ataxin-3 RNAs showed an increase in Hsp70 expression, a polyQ suppressor, which limited the RNAs toxicity, suggesting a possible dual action at the toxic RNA and protein level (Bonini, 2002; Shieh and Bonini, 2011; Warrick et al., 1999).

Interestingly, in SCA 8, both protein and RNA gain of function mechanisms may contribute to pathogenesis, due to bidirectional expression at the level of the *SCA 8* locus (Ikeda et al., 2008). Transcription in the CTG direction leads to the production of RNA foci that sequester MBNL1, while the transcript expressed in the CAG direction is translated into a small polyQ-containing protein forming ubiquitinated aggregates (Daughters et al., 2009). Antisense transcripts spanning CAG/CTG repeats have been described in SCA 7, HD, HDL2 and DM1 (Cho

et al., 2005; Chung et al., 2011; Sopher et al., 2011; Wilburn et al., 2011), suggesting that the SCA 8 situation may not be unique.

Furthermore, the recent study from Zu *et al*. raises the possibility for additional levels of toxicity (Zu et al., 2010). Recently, non-ATG initiation of translation (RAN), leading to production of protein tracts with different repetitive aminoacids, has been observed for DM1 and SCA 8. These additional toxic entities may act synergistically on pathogenesis (Zu et al., 2010). These examples highlight that the effects of CAG/CTG expansions are complex and warrant further investigation.

Finally, loss-of-function mechanism is also involved in TNR diseases. In FRDA, the intronic GAA expansion results in the transcriptional silencing of the frataxin gene (Campuzano et al., 1996). Similarly loss-of-function by transcriptional silencing has been associated to FXS (Sutcliffe et al., 1992).

1.3 Pathogenic threshold length of TNR

The size of the alleles implicated in TNR disorders is polymorphic in the normal population; however, the numbers of trinucleotide repeats does not exceed a specific threshold length. For most TNR diseases, a threshold length of 30 to 50 repeats needs to be reached to lead to pathogenesis (Gatchel and Zoghbi, 2005; Lopez Castel et al., 2010). Polyalanine (polyA) diseases, including OPMD, HOXA13 and HOXD13, as well as SCA 6, a polyQ disorder, represent an exception to this rule, with a toxic threshold of 7 to 20 GCG repeats (Brais et al., 1998; Goodman et al., 2000; Innis et al., 2004; Kato et al., 2000; Matsuyama et al., 1997; Mortlock et al., 1996; Muragaki et al., 1996; Takiyama et al., 1998).

The numbers of TNR found in patients is variable between diseases. Thousands of repeats can be transmitted to patients with DM1, FRDA or FRX. In contrast, in polyQ and polyA diseases, the maximum repeat size is relatively short, likely due to the high toxicity of polyQ and polyA proteins. For instance, in HD, while the normal repeat size ranges between 6 and 35 CAG units, the pathological size rarely exceeds a hundred of repeats. Most HD patients carry 40 to

50 CAG repeats, resulting in an onset during mid-life. Patients with larger repeats can develop a juvenile form of the disease, characterized by an onset during childhood and very rapid progression of symptoms (Duyao et al., 1993; Ranen et al., 1995; Zuhlke et al., 1993).

An interesting feature of some TNR loci is the existence of premutation alleles, as opposed to alleles carrying a full length mutation, and characterized by an intermediate toxic size. With this respect, the FRAXA locus is particularly interesting, as CGG expansions at the FRAXA locus results in FRX or FXTAS, depending on the size of the repeat (Pirozzi et al., 2011). A size of repeats between 56 and 200 lead to FXTAS results from a toxic gain of function of the mutant CGG-containing RNA (Fu et al., 1991; Jin et al., 2003; Sellier et al., 2010; Sofola et al., 2007), while FXS is caused by a CGG expansion greater than 200 units. In FXS, the expansion is methylated, leading to suppression of Fmr1 expression and a loss-of-function pathogenic mechanism.

2 Instability of TNR

A hallmark of all TNR diseases is the instability of the expanded repeat, resulting in contractions or expansions in both the germline and the somatic cells. TNR mutations are therefore dynamic mutations, inducing a continuous change of mutant allele size in successive generations through instability of germ cells and throughout patients' life through instability of somatic cells. (Table 1)

The level of TNR instability varies between diseases. For instance, somatic instability occurs during development in the case of FRX, whereas in HD, it is observed in adulthood (HD) and in DM1, it is observed during both development and adulthood. The tissues (and cells) that undergo instability are selective and specific to each disease, though an overlap can exist. For instance, the pattern of tissue selectivity of somatic instability is rather similar in all polyQ diseases. Thus, the dual axis of "when and where" TNR instability occurs is variable among diseases.

Table 1. Main repeat associated disorders and their mutation characteristics.

[According to (Lopez Castel et al., 2010)]

Disorder	Gene	Locus	Repeat	Repeat Size			Parent of Origin	Somatic Instability	Pathogenesis
				Normal	Premutation	Disease			
Coding repeats - Polyglutamine									
SBMA: Spinal and Bulbar Muscular Atrophy	AR	Xq11-q12	(CAG)n•(CTG)n	9 - 36	ND	40-53	ND	limited	GOF
HD: Huntington's disease	HTT	4p16.3	(CAG)n•(CTG)n	10 - 34	29-35	> 35	Paternal	moderate	GOF
DRPLA: Dentatorubral-pallidoluysian atrophy	ATN1	12p13.31	(CAG)n•(CTG)n	7 - 25	ND	49-88	Paternal	moderate	GOF
SCA1: Spinocerebellar ataxia 1	ATXN1	6p23	(CAG)n•(CTG)n	6 - 39	ND	39-81	Paternal	moderate	GOF
SCA2: Spinocerebellar ataxia 2	ATXN2	12q24	(CAG)n•(CTG)n	13 - 33	ND	> 34	Paternal	moderate	GOF
SCA3: Spinocerebellar ataxia 3	ATXN3	14q24.3-q31	(CAG)n•(CTG)n	13 - 44	ND	> 52	Paternal	moderate	GOF
SCA6: Spinocerebellar ataxia 6	CACNA1A	19p13	(CAG)n•(CTG)n	4 - 18	ND	20 - 30	none	ND	LOF
SCA7: Spinocerebellar ataxia 7	ATXN7	3p21.1-p12	(CAG)n•(CTG)n	4 - 35	ND	37-306	Paternal	moderate	GOF
SCA17: Spinocerebellar ataxia 17	TBP	6q27	(CAG)n•(CTG)n	25 - 42	ND	47 - 63	ND	ND	GOF
KCNN3: no disease association	KCNN3	1q21.3	(CAG)n•(CTG)n	7 - 28	ND	ND	ND	ND	ND
AIB-1: Increased prostate cancer risk	NCOA3	20q12	(CAG/CAA)n•(CTG/TTG)n	29/29 or 28/29	<29/29	>29 - 2 8	ND	limited	GOF
Coding repeats - Polyalanine									
HOXD13: Synpolydactyly	HOXD13	2q31-q32	(GCG)n•(CGC)n	15	ND	22 - 29	na	none	GOF
OPMD: Oculopharyngeal Muscular Dystrophy	PABPN1	14q11.2-q13	(GCG)n•(CGC)n	10	ND	12 - 17	na	none	GOF
CBFA1: cleidocranial dysplasia	RUNX2	6p21	(GCG)n•(CGC)n	17	ND	27	na	none	LOF
ZIC2: holoprosencephaly	ZIC2	13q32	(GCG)n•(CGC)n	15	ND	25	na	none	LOF
HOXA13: Hand-Foot-Genital Syndrome	HOXA13	7p15-p14.2	(GCG)n•(CGC)n	18	ND	24 - 26	na	none	GOF ?
FOXL2 Blepharophimosis Ptosis/ Epicanthus inversus syndrome type II	FOXL2	3q23	(GCG)n•(CGC)n	14	ND	22 - 24	na	none	LOF
ARX: infantile spasm syndrome	ARX	Xp22.13	(GCG)n•(CGC)n	10:16	ND	17-23	na	none	LOF
Coding repeats - Polyaspartic									
COMP: Multiple Skeletal dysplasias	COMP	19p13.1	(GAC)n•(GTC)n	5	ND	4-6,7	ND		LOF
Non-coding repeats									
CTG18.1: no confirmed disease association	TCF4	18q21.1	(CTG)n•(CAG)n	10 - 37	53 - 250	800 - 2100	ND	ND	ND
DM1: myotonic dystrophy type 1	DMPK/ DMPKas	19q13.2-q13.3	(CTG)n•(CAG)n	5 - 37	34-80	> 80 - 6500	Maternal	extensive	RNA
FRDA: Friedreich's ataxia	FXN	9q13	(GAA)n•(TTC)n	6 - 32	40-200	> 200	Maternal	moderate	LOF
SCA8: Spinocerebellar ataxia 8	ATXN8/ ATXN8OS	13q21	(CTG)n•(CAG)n	2 - 130	85 - 109	> 110	Exp: Mat Cnt: Pat	ND	unknown
SCA12: Spinocerebellar ataxia 12	PPP2R2B	5q31-q33	(CAG)n•(CTG)n	7 - 45	ND	55 - 78	ND	ND	ND
SCA31: Spinocerebellar ataxia 31	TK2/ BEAN	16q22	(TGGAA)n•(TTCCA)	0	ND	110	*	*	RNA
HDL2: Huntington's disease-like 2	JPH3	16q24.3	(CAG)n•(CTG)n	6 - 27	ND	51 - 57	ND	ND	RNA

Table 1. Main repeat associated disorders and their mutation characteristics. (continued)

[According to (Lopez Castel et al., 2010)]

Disorder	Gene	Locus	Repeat	Repeat Size: Normal	Repeat Size: Premutation	Repeat Size: Disease	Parent of Origin	Somat Instab	Pathogenesis
MAB21L1: not confirmed disease association	MAB21L1	13q13	(CAG)n*(CTG)n	6 - 31	ND	> 50	ND	limite	ND
DM2: myotonic dystrophy type 2	ZNF9	3q13.3-q24	(CCTG)n*(CAGG)n	104 - 176	ND	75 - 11.000	ND	extens	RNA
SCA10: Spinocerebellar ataxia 10	ATXN10	22q13	(AATCT)n*(AGATT)n	10 - 22	ND	800 - 4500	Paternal	moder	LOF
EPM1:epilepsy progressive myoclonic	CSTB	21q22.3	(CCCCGCCCCGCG)n	2 - 3	12 - 17	30 - 75	ND	ND	LOF
IDDM: Increased risk of diabetes	INS	11p15.5	(ACAGGGGT(GC)(T C)GGGG)n	110 - 130	45 - 109	30 - 44	Maternal	ND	unknown
h-Ras: increased risk of ovarian cancer	HRAS	11p15.5	28-bp VNTR	5 alleles	ND	rare alleles	ND	ND	unknown
FSHD: Facioscapulohumeral muscular dystrophy	FSHMD1A	4q35	3.3 kb D4Z4repeat	11-150	ND	<11	ND	moder	unknown
SMYD3: Cancer risk factor	SMYD3	1q44	(CCGCC)n(GGCGG)n	3	ND	3	?	?	GOF
RELN: Risk of Autism	RELN	7q22	(CGG)n(GCC)n	8-10	ND	12-13	No	?	LOF
Fragile site-associated repeats - *Folate-Sensitive Sites*									
FRAXA: fragile X syndrome	FMR1 FMR1	Xq27.3	(CGG)n*(CCG)n	6 - 52	59 - 230	230 - 2000	Exp: Mat Cur: Pat	limite	LOF, RNA ?
FXTAS: fragile X tremor/ataxia syndrome	FMR1	Xq27.3	(CGG)n*(CCG)n	6 - 52		59 - 230	Exp: Mat Cur: Pat	limite	GOF, RNA
FRAXE: fragile X syndrome	FMR2	Xq28	(CCG)n*(CGG)n	4 - 39	(31-61)	200-900	Exp: Mat Cur: Pat	ND	LOF
FRAXF: No confirmed disease association	FAM11A	Xq28	(CGG)n*(CCG)n	7 - 40	ND	306-1008	ND	ND	ND
FRA10A: No confirmed disease association	FRA10AC1	10q23.33	(CCG)n*(CGG)n	8 - 14	ND	> 300	ND	ND	ND
FRA11B: Jacobsen syndrome	CBL2	11q23.3	(CCG)n*(CGG)n	11	80	100-1000	ND	ND	LOF
FRA16A: No confirmed disease association	—	16q22	(CCG)n*(CGG)n	16 - 49	ND	1000-1900	Maternal	ND	ND
Fragile site-associated repeats - *Distamycin A-inducible*									
FRA16B: No confirmed disease association	—	16q22.1	33-bp AT-rich VNTR	7-12	ND	up to 2000	ND	limited	ND
Fragile site-associated repeats - *BrdU-inducible*									
FRA10B: No confirmed disease association	—	10q25	~42-bp AT-rich VNTR	1-74	ND	> 75	ND	limited	ND

2.1 TNR instability: where and when?

2.1.1 Germline instability of TNR

TNR diseases display repeat instability at the germ cell level, most often resulting in further expansion of the TNR size, though contractions are also observed (Aziz et al., 2011; De Michele et al., 1998; Monckton et al., 1995; Moseley et al., 2000). Interestingly, depending on the disease, paternal or maternal biases are observed, suggesting the involvement of disease-specific processes related to spermatogenesis or oogenesis (Pearson et al., 2005). Germline expansion can lead to an anticipation effect, corresponding to worsening of the disease in successive generations, reported also as the "Sherman paradox" for FXS (Sherman et al., 1985; Sherman et al., 1984). For instance, in HD, transmission by fathers of the mutation to offspring tends to lead to increased severity of the disease, with an earlier onset and a more rapid progression of symptoms. This is due to the production of mutant protein with a longer polyQ, and therefore with increased toxicity.

Most polyQ diseases are actually characterized by a paternal expansion bias occurring during spermatogenesis and exhibiting a mosaicism of repeat sizes in sperm. Precisely, expansion bias was observed in patients suffering from DRPLA (Ikeuchi et al., 1995b; Takiyama et al., 1999), HD (Andrew et al., 1993; Duyao et al., 1993; Leeflang et al., 1995; Snell et al., 1993; Telenius et al., 1995), and several SCAs [SCA 1 (Chung et al., 1993), SCA 2 (Riess et al., 1997), SCA 3 (Ikeuchi et al., 1995a; Takiyama et al., 1997), SCA 7 (David et al., 1998; Gouw et al., 1998; Monckton et al., 1999). For instance, in HD, small size increases of one to a few repeats are typically observed during both paternal and maternal transmission (Duyao et al., 1993). However, in a subset of male transmissions, larger increase occur to produce extreme HD alleles, resulting in anticipation phenomenon and causing juvenile form of the disease (Duyao et al., 1993). It has been shown that CAG expansion upon paternal transmission occurs before meiosis begins, during spermatogonial stem mitotic divisions (Yoon et al., 2003). It is believed that the great number of mitotic divisions for male gametogenesis contributes to the paternal expansion bias, suggesting that replication plays a role in the instability of the male germline.

The age does not appear to play a role in determining HD CAG instability in male germline (Aziz et al., 2011; Wheeler et al., 2007).

DM1 germline instability was described for both sexes, but the parent-of-origin effect differed depending on the mutation length. A paternal expansion bias is associated to repeat length <100, paternal and maternal expansion bias are observed for middle repeat lengths (200-600) and a maternal expansion bias is described for repeat length >600 repeats, resulting in extreme DM1 alleles of 1000 to 4000 repeats (Jansen et al., 1994; Lenzi et al., 2005; Martorell et al., 2000). It is suggested that each parent-of-origin mutation bias is driven by mechanisms specific to sperm or oocyte development. Both pre-meiotic and post-zygotic events likely contribute to the large DM1 CTG expansion upon maternal transmission (De Temmerman et al., 2004; Lenzi et al., 2005). It is believed that the highly extended time for oogenic meiosis contributes to pre-meiotic instability (De Michele et al., 1998; Kaytor et al., 1997; Sato et al., 1999).

Extensive maternal expansion bias is also observed for FRAXA fetuses during oogenesis or post-zygotically, whereas male patients show a paternal contraction bias leading to unmethylated alleles of premutation size (Malter et al., 1997; Moutou et al., 1997; Nolin et al., 1999; Reyniers et al., 1993).

A paternal contraction bias has been reported for SCA 8 (Moseley et al., 2000), FRDA (De Michele et al., 1998; Delatycki et al., 1998; Monros et al., 1997), DM1 (Monckton et al., 1995), and FRAXA (Ashley-Koch et al., 1998).

2.1.2 Somatic instability of TNR is tissue-selective

Somatic instability of TNR is tissue- or cell-specific, resulting in repeat tract length variation between or within tissues. The timing, pattern and tissue selectivity of somatic instability is variable between TNR diseases. For instance, in HD and other polyQ disorders, including SBMA, somatic CAG instability is not observed during development, but is detected in the adult (Benitez et al., 1995; Jedele et al., 1998). Conversely, in FXS, somatic CGG instability

only occurs in fetal tissues, but not postnataly (Devys et al., 1992). In DM1, somatic CTG instability is detected both in fetal and adult tissues (Martorell et al., 1997; Thornton et al., 1994; Zatz et al., 1995).

Somatic instability of TNR is tissue selective, but this tissue selectivity varies between diseases. In HD and some polyQ disorders, both the striatum and the cortex show increased instability, while instability is minimal in the cerebellum (Chong et al., 1995b; Hashida et al., 2001; Kennedy and Shelbourne, 2000; Lopes-Cendes et al., 1996; Shelbourne et al., 2007; Telenius et al., 1994). However, this pattern of somatic instability is not shared by all polyQ diseases. In SBMA, the CAG tract is stable in the central nervous system, but unstable in the muscle (Tanaka et al., 1999; Tanaka et al., 1996). In DM1, the tissues presenting high levels of instability include the heart, skeletal muscle and cortex, the cerebellum showing a reduced level of instability (Anvret et al., 1993; Ishii et al., 1996; Thornton et al., 1994). In FRDA, somatic GAA instability is elevated in the cerebellum and dorsal root ganglia (Clark et al., 2007).

In HD and other polyQ disorders, somatic instability is usually greatest in the central nervous system (Pearson et al., 2005). More specifically, it has been shown that neurons present a higher propensity for instability, when compared to non neuronal cells such as glial cells, indicating that replication-independent mechanisms contribute to CAG instability in polyQ diseases (Shelbourne et al., 2007).

In HD, DM1, and FRDA somatic instability is most prevalent in the affected tissues, suggesting that instability contributes to disease progression. In the case of HD, somatic instability is most extensive in the striatum, corresponding to the tissue that preferentially degenerates. It is believed that somatic expansion in the striatum significantly accelerates HD pathology, by leading to the production of increasingly toxic mutant proteins, and therefore acts as a disease-modifier (Gonitel et al., 2008; Shelbourne et al., 2007).

2.1.3 Somatic instability of TNR is age-dependent

Repeat instability evolves with age. Heterogeneity of the repeat tract was reported to correlate with age and to progress with age in blood cells of DM1 patients (Martorell et al., 1995; Martorell et al., 1998; Wong et al., 1995). Additionally, the difference of the size of the repeat expansions in muscle, an unstable tissue, with the size of the expansion in lymphocytes, a relatively stable tissue, in DM1 young and adult patients revealed higher differences in older subjects suggesting that somatic repeat instability is age-dependent (Zatz et al., 1995). In a similar manner somatic instability was suggested to increase with age in FRDA and DRPLA patients (De Biase et al., 2007a; De Biase et al., 2007b; Takano et al., 1996). The age of onset of disease was described to correlate with repeat length in DM1 and HD (Duyao et al., 1993; Harper, 1992; Swami et al., 2009). Mouse models of TNR diseases also showed variation of somatic instability with age (Fortune et al., 2000; Gonitel et al., 2008; Ishiguro et al., 2001; Kovtun et al., 2007; Lia et al., 1998; Sato et al., 1999). An exception to this tendency is FXS where instability of the repeat tract is present at embryonic levels whereas repeats are stable postnataly (Sato et al., 1999).

2.1.4. Somatic instability of TNR is repeat length-dependent

The length of repeats inlfuences the instability process. The length of repeats correlated with the level of somatic instability in SBMA, DM1 and HD patients (Kahlem and Djian, 2000; Martorell et al., 1995; Tanaka et al., 1999). Similar correlation was found for *FMR1* gene (Eichler et al., 1994).

2.1.5. Conclusions

Most TNR diseases show repeat instability. TNR loci are located in different positions in the corresponding genes and therefore the position and/or the sequence of repeats might have a role in instability. Repeat instability is involved in intergenerational variability of disease onset and severity and manifests as intergenerational and somatic instability. Somatic instability is tissue-selective and, in some diseases, including HD, DM1 and FRDA, the tissues or cells that are most vulnerable regarding pathogensis are also those presenting highest TNR instability. Therefore it is important to examine the implication of instability in disease evolution, since instability may lead to production of increasingly toxic mutant gene products with time.

2.2 Mechanisms implicated in TNR instability

2.2.1 Structure of TNR

2.2.1.1 DNA structure of TNR

The models proposed to explain TNR instability involve the formation of stable secondary DNA structures, which would be aberrantly processed by mechanisms associated to DNA physiology, including replication, DNA repair, transcription and recombination (Pearson et al., 2005). *In vitro* experiments have indeed shown that TNR sequences can adopt several structures (Mirkin, 2007). Single stranded DNA with CNG repeats, where "N" is one of the four desoxynucleotides, can form stable slipped strand DNA structures (or hairpin structures), adopting a mismatched base pair conformation (Gacy et al., 1995; Pearson and Sinden, 1996). The stability of the secondary structures depends on the thermodynamic energy of the base pair mismatch. The propensity to form stable hairpins ranks as follows: CCG=CAG<CTG<CGG. It was further shown that slipped-out CAG and slipped-out CTG repeats are predominantly in the random coil and hairpin conformation, respectively (Pearson et al., 2002). Accordingly, several studies propose that CTG repeats have a higher propensity to form stable hairpin, compared to CAG repeats (Hou et al., 2009; Mitas et al., 1995; Panigrahi et al., 2005). In addition, increasing

the length of CNG repeats augments the stability as well as the complexity of the secondary structures, longer repeats being more stable and presenting alternative DNA structures with more complex patterns (Gacy et al., 1995; Pearson and Sinden, 1996; Pearson et al., 1998b). More specifically, alternative structures of more than 10 repeats can show a pattern, involving multiple loops or hairpins (Panigrahi et al., 2010; Pearson et al., 1998b). Accordingly, the stability of the secondary structures is dependent upon the purity of the CNG tract, as CAT or AGG interruptions within CAG or CGG tracts, respectively, have both qualitative and quantitative effects on slipped strand formation (Pearson et al., 1998a). Finally, tetrahelical structures have been observed for CGG and CCG repeats (Fry and Loeb, 1994; Usdin and Woodford, 1995), which are stabilized by G quartet motifs (Zheng et al., 1996).

Unconventional DNA structures have also been described for GAA repeats (Bidichandani et al., 1998; LeProust et al., 2000). Unlike CNG repeats, GAA repeats can form triplex or sticky DNA structures, likely contributing to both GAA instability and silencing of frataxin expression (Gacy et al., 1998; Grabczyk and Usdin, 2000a, b; Heidenfelder et al., 2003; Sakamoto et al., 1999).

Thus, the propensity of TNR for forming stable secondary structures has been clearly shown using *in vitro* approaches, supporting their role in TNR instability; yet, the existence of such structures *in vivo* needs to be demonstrated. (Figure: 1: Unusual DNA structures formed by TNR repeats, examples.)

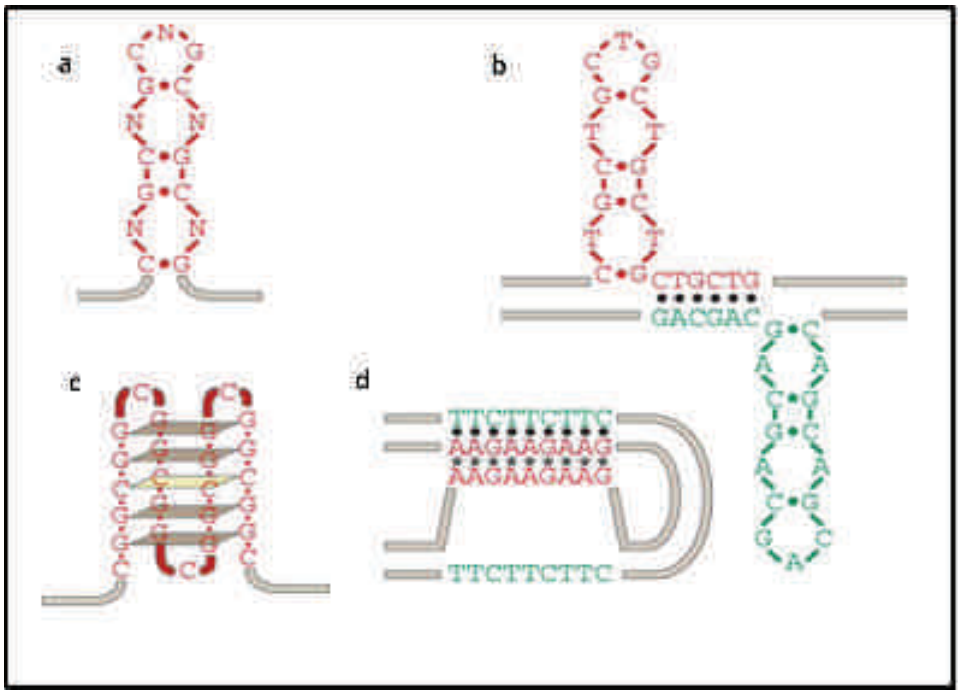


Figure 1. Unusual DNA structures formed by TNR repeats, examples.

The structure-prone strand of the repetitive run is shown in red, its complementary strand in green, and flanking DNA in beige. a, An imperfect hairpin formed by $(CNG)_n$ repeats. b, A slipped-stranded structure formed by the $(CTG)_n$ $(CAG)_n$ repeat. c, A quadruplex-like structure formed by the $(CGG)_n$ repeat. Brown rectangles indicate G quartets. d, sticky DNA formed by the $(GAA)_n$ $(TTC)_n$ repeat. [According to (Mirkin, 2007)]

2.2.1.2 RNA structure of TNR

mRNA transcribed from the genes containing TNR mutation are also prone to formation of stable secondary structures, which can be toxic and contribute to pathogenesis. CUG repeats of the *DMPK* transcript can adopt imperfect mismatch hairpin structures, the stability of which increases with repeat length (Napierala and Krzyzosiak, 1997). By using chemical and enzymatic assays, it has been shown that RNA with CNG repeats form stable hairpins, the stability of which ranks as follows: CAG<CCG<CUG<CGG (Sobczak et al., 2003). In addition, the CGG repeats

in FXTAS mRNA also form stable hairpin structures, contributing to pathogenesis (Hoem et al., 2011; Kiliszek et al., 2011; Solvsten and Nielsen, 2011; Zumwalt et al., 2007).

2.2.2 DNA repair

2.2.2.1 Base Excision Repair pathway (BER) and TNR instability

2.2.2.1.1 BER: a repair mechanism divided in two subpathways

2.2.2.1.1.1 Lesions repaired by BER

BER is involved in the elimination of DNA damages caused by spontaneous base hydrolysis, by reactive oxygen species (ROS) and by alkylating agents (Lindahl, 1993, 2000). Base oxidation, one most common cause of DNA damage, is caused by ROS, which are either endogenously produced by the cellular metabolism or have exogenous origins. One of the most common oxidant damages is the modification of guanine into 8-oxodesoxyguanine (8-oxodG). Other examples of oxidant damage are 8-oxodesoxyadenine, 5-hydroxycytosine (5OHC), 5-hydroxyuracile (5OHU), etc. (Muller et al., 1998; Wagner et al., 1992). Other types of damages include spontaneous hydrolysis, resulting in depurination and depyrimidation and to formation of AP sites (Lindahl, 1993, 2000), and base alkylation by exogenous factors, resulting in 3-methyladenine (3meA), 7-methylguanine (7-meG), 2-methylcytosine (2-meC) etc (Seeberg et al., 1995). (Figure 2: DNA bases and examples of common modifications)

2.2.2.1.1.2 DNA glycosylases involved in BER

DNA glycosylases are the initiating enzymes of BER that specifically detect and excise the modified base by incising the N-glycosidic bond of the base to the sugar-phosphate DNA strand, generating thus an AP site (apuric/apyrimidic (AP)). At that point the DNA strand is incised in 5' by an AP-lyase activity generating a single-strand break (SSB) (Fortini and Dogliotti, 2007).

The DNA glycosylases can be monofunctional enzymes, such as uracil-DNA glycosylase (UDG) and 3-methyladenine DNA glycosylase (AAG), which only recognize and excise the damaged base, i.e. uracil for UDG and 3-meA, xanthine and to a lesser extent 8-oxodG for AAG (Engelward et al., 1997; Mol et al., 1995; O'Brien and Ellenberger, 2003; Savva et al., 1995). The step of SSB formation is catalyzed by the AP-endonuclease 1 (APE1), which generates 3'-OH and 5'-deoxyribose phosphate (5'dRP) extremities at the damaged site (Wilson and Barsky, 2001). However, there are also bifunctional DNA glycosylases, including 8-oxoguanine DNA glycosylase (OGG1) and endonuclease III (NTH1), which not only recognize and excise the damaged base, but also dispose an AP-lyase activity that incises at 3' of the AP site generating a 3'-α,β, unsaturated aldehyde or a 3'dRP (Boiteux and Radicella, 1999, 2000). The endonucleases VIII-like1 and 2 (NEIL1 and NEIL2, respectively) are also bifunctional, but lead to formation of a SSB that harbors a 3'-phosphate extremity (Das et al., 2006; Hazra et al., 2002a; Hazra et al., 2002b). OGG1 is the main enzyme that recognizes 8-oxodG modification and FapyG, NTH1 recognizes mainly 5OHC, thymidine glycol (Tg), 5OHU, urea and formamide pyrimidine guanine (FapyG) modifications. NEIL1 repairs the same lesions than NTH1 and, in addition, formamide pyrimidine adenine (FapyA) modification and Neil2 recognizes AP sites and 5OHU damages (Ide and Kotera, 2004). (see Table 2: DNA glycosylase function and targeted lesion)

Therefore the DNA glycosylases have overlapping functions but different damage specificities. They either recognize specifically one modification -e.g. UDG that recognizes uracil-, or they recognize a group of modifications but with a different specificity for each lesion -e.g. OGG1 is an enzyme that mainly recognizes 8-oxodG, but also recognizes other lesions with less affinity.

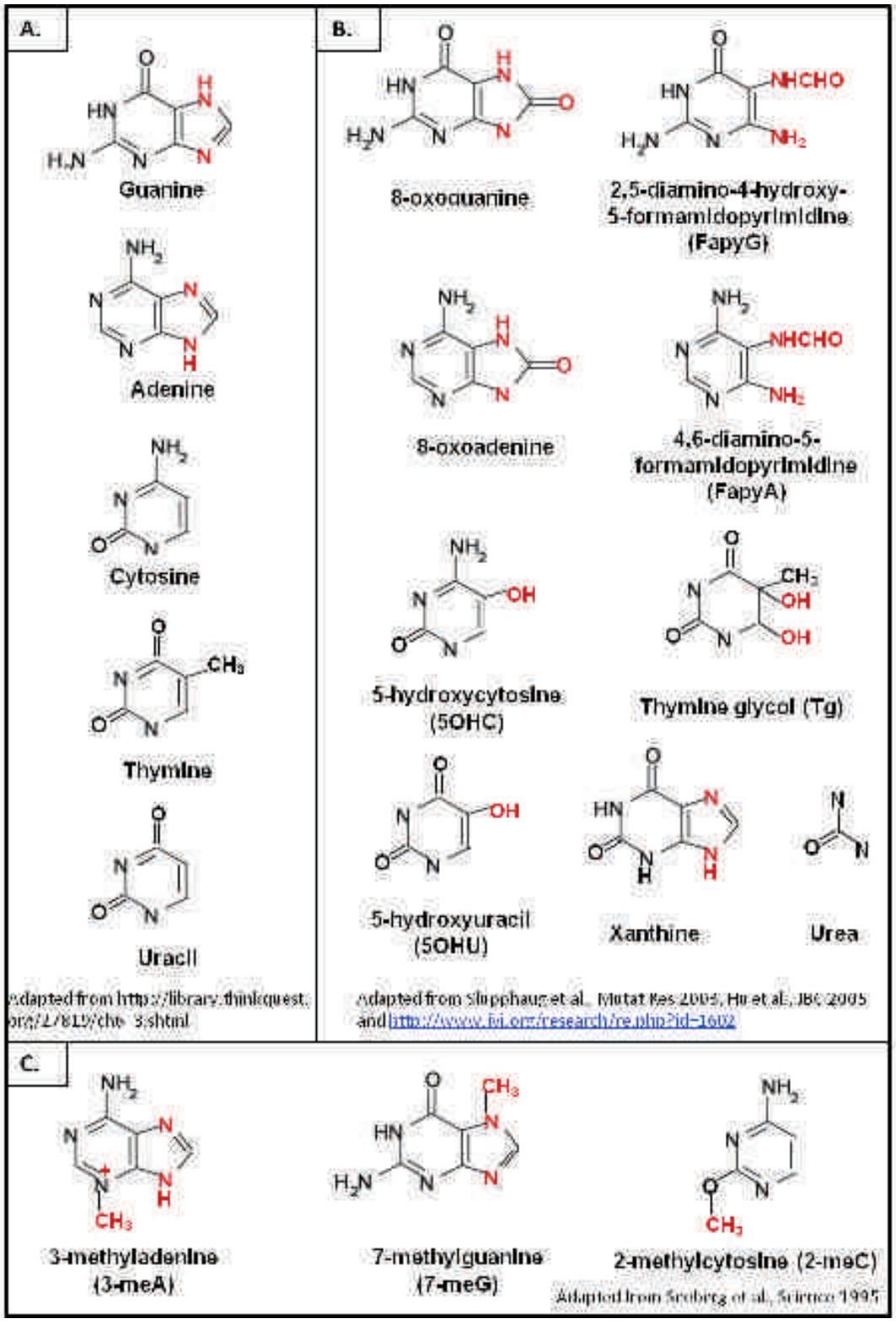


Figure 2. DNA bases and examples of common modifications. A. nucleic bases, B. common oxidized or reduced bases, C. common alkylated bases [Adapted by (Seeberg et al., 1995;

Slupphaug et al., 2003), http://library.thinkquest.org/27819/ch6_3.shtml and http://www.jyi.org/research/re.php?id=1602]

Table 2. DNA glycosylase function and targeted lesion

Glycosylase categorie	Glycosylase	Targeted lesion
monofunctional	UDG	uracil
	AAG	3-meA, xanthine, 8-oxodG
	MYH	8-oxodG
bifunctional	OGG1	8-oxodG, FapyG
	NTH1	5OHC, Tg, 5OHU, FapyG
	NEIL1	5OHC, Tg, 5OHU, FapyG, FapyA
	NEIL2	AP site, 5OHU

2.2.2.1.1.3 Processing of the AP site: SN- and LP-BER

The subsequent repair of SSB requires the presence of 3'-OH and 5'-phosphate extremities. To this purpose, Polβ due to its 5'-dRP-lyase activity excises the 5'dRP (Matsumoto and Kim, 1995). The 3'-dRP and 3'-α,β, unsaturated aldehyde extremities formed following the OGG1 and NTH1 cleavage are directly excised by APE1 (Izumi et al., 2000; Klungland et al., 1999; Parsons et al., 2004; Wiederhold et al., 2004), whereas the 3'-phosphate extremity following Neil1 and Neil2 cleavage is processed by PNK (Das et al., 2006; Wiederhold et al., 2004).

In order to fill the gap between the two opposite DNA extremities on the DNA strand, a number of nucleotides can be added. The number of nucleotides incorporated at the lesion site determines the repair pathway that is used. When a single nucleotide is incorporated, the single nucleotide BER (SN-BER) pathway is used, whereas when several nucleotides are incorporated, the lesion is processed via Long-Patch BER (LP-BER). During SN-BER, Polβ (Polλ

and maybe Polθ (Braithwaite et al., 2010; Prasad et al., 2009)) incorporates the missing nucleotide and promotes the recruitment of DNA Ligase III (Lig3) to seal the nick, and of X-ray repair complementing defective repair in Chinese hamster cells 1 (XRCC1) which interacts with both Polβ and Lig3 allowing efficient BER (Dianova et al., 2004; Parsons et al., 2005). On the other hand, LP-BER is used when the 5'-dRP is refractory to excision by Polβ lyase activity. In this pathway, several DNA Polymerases, including Polβ, Polδ and Polε (all of them participating also in replication), as well as Polλ, Polι and Polθ can be used for multinucleotide incorporation (Yoshimura et al., 2006). The processivity of DNA polymerases is promoted by proliferating cell nuclear antigen (PCNA). The incorporation of multiple nucleotides results in displacement of the adjacent 5' strand, which forms a 5' flappy strand that is excised by the flap endonuclease FEN1. The two edges are finally sealed by DNA Ligase I (Lig1). (see Figure3: BER pathway and interplay between SN- and LP-BER)

The different proteins implicated in repair interact between each other, which enhances their repair activities. The role of DNA Polβ in strand displacement synthesis and in controlling the size of the excised flap was suggested by Dianov and coworkers (Dianov et al., 1999). Using reconstituted repair assays, DNA Polβ was reported not only to cooperate with FEN1 but also to influence the length of the incorporated nucleotides in the lesion depending on the relative concentration of those two proteins (Prasad et al., 2000). Later, the coordination of those two proteins during LP-BER was shown to function through a "Hit and Run" mechanism, with the two proteins acting alternatively during repair in reconstituted assays (Liu et al., 2005). The notion of Polβ/DNA binary complexes and its association to APE1 which stimulated the Polβ gap-filling and lyase activity argued rather for a coordination and interaction of BER proteins (Liu et al., 2007). Several other BER actors were shown to interplay. For instance, PCNA was shown to interact with DNA Ligase I, Polδ and Polε in LP-BER. DNA Ligase I specific implication in LP-BER was also argued by its interaction with the Rad9-Rad1-Hus1 complex (9-1-1) (Wang et al., 2006), a PCNA analogue, and its interaction with FEN1 (Smirnova et al., 2005). More recently, DNA polymerases β and λ were reported to interact with DNA glycosylases in mouse embryonic fibroblasts, suggesting an early coordination between repair proteins already from the initiating steps of repair (Braithwaite et al., 2010). Although DNA Ligase III and DNA Ligase I

are preferentially associated to SN- and LP-BER, respectively, this distinction seems to be more shaded as DNA Ligase I can be also partially involved in SN-BER (Pascucci et al., 2002).

While SN- and LP-BER can be both reconstituted *in vitro*, the definitive demonstration of the *in vivo* existence of both pathways is more recent. Studies based on repair of uracil and oxidant damages such as 8-oxodG and Tg by *E.coli* and mammal cell extracts indicated that SN-BER mainly drives repair in cells (Dianov et al., 1998; Dianov et al., 1992; Dianov et al., 2000). SN-BER in cells was suggested to be DNA Polβ-dependent. Interestingly, mouse fibroblasts deficient for Polβ still had the ability to perform SN-BER in a slower repair reaction by using Polδ and ε (Stucki et al., 1998). However, the rate of lesions processed by SN-BER was decreased compared to cells with Polβ, whereas the rate of lesions processed by LP-BER was increased (Dianov et al., 1998). LP-BER was observed using mouse embryonic fibroblasts extracts to repair methylation-induced damage (Horton et al., 2000) or AP sites (Dogliotti et al., 2001). Repair of AP sites using human lymphoid cells also involved LP-BER (Podlutsky et al., 2001). In post-mitotic cells, the use of LP-BER was sustained by the increased level of Polβ, δ, or ε in rat neurons (Hubscher et al., 1977; Raji et al., 2002). Interestingly, a study revealed a Polβ-dependent but PCNA-independent LP-BER mechanism specific to post-mitotic neurons in rat brain (Wei and Englander, 2008). Those results suggest that LP-BER can be used in cells.

Few mechanisms have been implicated in BER subpathway selection. Repair of oxidized or reduced AP sites requires FEN1, suggesting the preferential use of LP-BER under those conditions (Klungland and Lindahl, 1997). Another determinant of subpathway selection is the relative ATP availability (Petermann et al., 2003). A low level of ATP stimulates the use of LP-BER in a Polβ and PARP1-dependent reaction, which allowed regeneration of ATP required for ligation. Finally, in this assay, XRCC1 and DNA Ligase III influenced the choice of BER subpathway (Petermann et al., 2006). In fact, XRCC1 stimulated the strand displacement activity of Pol β in case of ATP shortage, and thus the use of LP-BER. However, in the presence of ATP, XRCC1 effect was counteracted by stimulation of DNA Ligase III activity, thus promoting the use of SN-BER. Those results give some important clues for the preferential use of SN- or LP-BER *in vitro*; nevertheless what is influencing the final subpathway choice *in vivo* needs further investigation.

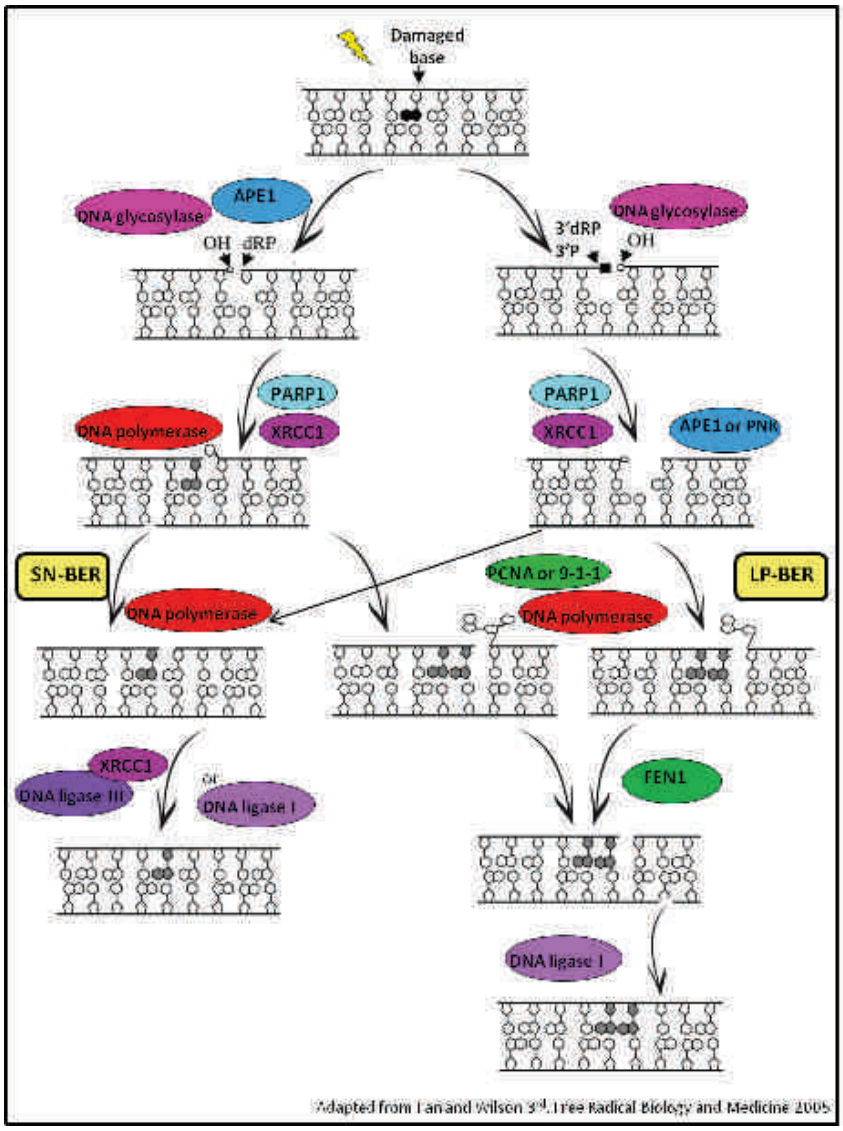


Figure 3. BER pathway and interplay between SN- and LP-BER. A DNA damage is excised by a monofunctional DNA glycosylase and APE1 leads to production of 3'OH and 5-dRP extremities. On the other site 3'dRP or 3'-phosphate (P) extremities generated by some bifunctional glycosylases, are processed by APE1 or PNK. In SN-BER a DNA polymerase incorporates 1 nucleotide and DNA ligase III/XRCC1 or DNA ligase I seals the nick. Addition of >1 nucleotide by DNA polymerase lead to formation of a 5'flappy strand excised by FEN1, in the case of LP-BER. Ligation of the nick is performed by DNA ligase I. PARP1 is a nick sensor, repair protein recruiter and coordinator; PCNA, 9-1-1 are repair proteins coordinators. [Adapted from (Fan and Wilson, 2005)]

2.2.2.1.1.4 Cofactors involved in BER

BER can be influenced by several cofactors. One of them is the high-mobility group box 1 protein (HMGB1), which was characterized as a DNA "bending" molecule to facilitate interaction of DNA with repair proteins. HMGB1 specifically interacts with the dRP intermediate substrate and stimulates APE1 and FEN1 activity during LP-BER (Prasad et al., 2007) and inhibits SN-BER (Liu et al., 2010b). Another BER cofactor is PARP-1, which recognizes SSB extremities, homodimerizes or heterodimerizes with PARP-2, and following this activation forms a poly-ADP ribose polymer (PAR) that produces ATP, recruits repair factors at the lesion, and then dissociates from the DNA. It has been proposed that in case of energy depletion, synthesis of PAR is stimulated and repair through LP-BER is promoted as PARP-1 stimulates Polβ nucleotide incorporation and interacts with DNA Ligase III and XRCC1 (Caldecott et al., 1996; Petermann et al., 2003). It is therefore suggested that PARP-1 is mainly used upon ATP shortage to regenerate ATP, which is necessary during ligation (Oei and Ziegler, 2000; Petermann et al., 2003). In accordance with those results, PARP-1 was shown to interact with Polβ and regulate its activity by interfering with APE1 and FEN1 (Sukhanova et al., 2010; Sukhanova et al., 2005). Upon PAR synthesis, Polβ synthesis is enhanced and LP-BER is stimulated (Sukhanova et al., 2010).

The complex formed of Rad9-Rad1-Hus1 proteins (9-1-1) is another sensor of DNA damage. It was shown to co-localize with its analogue PCNA in DNA damage foci (Bai et al., 2010). 9-1-1 interacts with MYH glycosylase, APE1, DNA polymerase β, FEN1, DNA ligase I, and was shown to promote LP-BER by stimulating APE1 and Polβ strand-displacement activity (Gembka et al., 2007), as well as DNA ligase I (Smirnova et al., 2005; Wang et al., 2006), in a "Hit and Run" mechanism (Balakrishnan et al., 2009). Therefore, BER is regulated by a whole budge of proteins, suggesting that their site- and time-specific availability could influence BER coordination and activity *in vivo*.

2.2.2.1.2 Contribution of BER to instability of CAG/CTG repeats

2.2.2.1.2.1 Yeast genetics show that players of BER are involved in CAG/CTG instability

Yeast studies were the first evidences that allowed gaining insight into the involvement of BER in TNR instability. Precisely, instability was prevented in yeast strains with wild-type or functional rad27, the homolog of mammalian FEN1 (Liu et al., 2004). Deficiency of this protein in yeast led to length-dependent CAG/CTG expansion and instability (Freudenreich et al., 1998). Similarly, haploinsufficiency of *rad27* promoted expansion of long repeat tracts in a repeat length-dependent manner, suggesting a rate limiting effect of rad27 depending on the repeat length (Yang and Freudenreich, 2007). On the other side, it is the overexpression of DNA ligase I homolog in yeast (*cdc9*) that yielded longer repeat tracts (Subramanian et al., 2005). Strikingly, inactive DNA ligase I possessing a functional binding site for PCNA led to similar results (Subramanian et al., 2005). Taking into account that the PCNA-binding site is mutually exclusive between cdc9 and rad27, those results suggest that it is the protein level as well as the interplay between several repair factors that influence repeat instability. DNA polymerases are another group of proteins that have been involved in TNR instability. Mutations in DNA Pol α, δ and ε led to destabilization of the repeat tract (Schweitzer and Livingston, 1999), whereas other polymerases such as pol ζ (the *rev7* homolog) and pol η (the *rad30* homolog) had minor effect on repeat instability (Dixon and Lahue, 2002). FEN1, Ligase I and the DNA polymerases are involved in both DNA repair and replication. Yeast studies did not allow discriminating if DNA repair and/or replication contributed to TNR instability.

2.2.2.1.2.2 Inactivation of Ogg1 *in HD transgenic mice reduces somatic CAG/CTG instability*

Inactivation of *Ogg1* in HD R6/1 transgenic mice, which express the first exon of *HTT* with a CAG/CTG expansion, showed a reduced age-dependent repeat instability in brain and liver, suggesting a role for oxidative DNA damage and BER in repeat instability (Kovtun et al.,

2007). In fact, the 8-oxoguanine modification accumulated in the ageing brain of HD mice. Several other oxidant lesions -such as 5'-hydroxyuracile (5'-OH-U), 5'-hydroxycytosine (5'-OHC) and formamidopyrimidine (FAPY)- were also tested and accumulated in the ageing brain of R6/1 mice. On the opposite, the level of other lesions such as uracil or 3-meA did not change. Interestingly, the CAG/CTG instability of R6/1 mice deficient for *Nth1* was similar to that of R6/1 mice. In a similar manner, deficiency of *Aag* had no effect on somatic instability. Those data suggest that although several types of DNA damage may accumulate in the brain of HD mice, not all of them play a role in repeat instability.

Assuming that BER and MMR (see §2.2.2.2) interact, Kovtun et al. proposed the following "toxic oxidation cycle" model (Kovtun et al., 2007). In this model, an oxidant DNA damage on the CAG/CTG repeats initiates BER. After excision of the damaged base and formation of a single stranded break by APE1, DNA Polβ is incorporating nucleotides in the gap. DNA Polβ would slip on the repeats, facilitating the formation of repeat hairpin structures, which would then be stabilized by MutSβ complex (Msh2-Msh3). This hairpin structure, which would be refractory to excision by FEN1, would be incorporated into the repeat tract, leading to repeat expansion. The load of oxidant DNA damage in tissues may increase with age, leading to the age-dependent increase of somatic CAG/CTG instability. (see Figure 4: The toxic oxidation cycle model)

The "toxic oxidation cycle" model implies that oxidant DNA damage at CAG/CTG repeats is increasing with age. Kovtun et al. showed that 8-oxodG and 5'-OH-U lesions accumulate in the ageing brain in both wild-type and R6/1 mice (Kovtun et al., 2007). However, those data do not specify if this accumulation is tissue-specific. Do the lesions accumulate more in the striatum that shows an important CAG/CTG instability as compared to tissues exhibiting more limited CAG/CTG instability? In addition, the results did not examine whether the lesions specifically accumulated with time at the CAG/CTG repeats in the *HTT* gene.

The "toxic oxidation cycle" model also suggests that LB-BER is involved in the somatic repeat instability as Polβ would incorporate multiple nucleotides during BER. *In vitro*, the amount of incorporated nucleotides in CAG substrates was more elevated when increasing DNA Polβ concentration (Kovtun et al., 2007). However no data on tissue-specific amount of DNA

Polβ are reported in order to support the involvement of DNA Polβ in the tissue-specific or age-dependent TNR instability.

Finally, the "toxic oxidation cycle" model presumes that BER promotes the formation of stable CAG/CTG secondary structures, rather than being involved in repair of the secondary structures. Whether BER promotes formation of those stable secondary structures remains an open question.

2.2.2.1.2.3 Reduction of FEN1 or DNA Ligase I in HD or DM1 mice does not affect somatic CAG/CTG instability

FEN1 and Ligase I are both involved in LP-BER. Although those two enzymes are involved in DNA repair, one has to keep in mind that they are also playing a role in replication in the lagging strand during the maturation processing of Okazaki fragment (Rossi et al., 2008).

Expansion of TNR is thought to result from polymerase slippage and formation of hairpin structures that are refractory to excision by FEN1 (Henricksen et al., 2000; Spiro et al., 1999). In order to assess the role of FEN1 in TNR instability, mouse and cell models for HD and DM1 were crossed with mice haploinsufficient for *Fen1*, as the complete loss of *Fen1* is embryonic lethal (Kucherlapati et al., 2002). Somatic CAG/CTG instability was comparable in HD mice heterozygous for Fen1 and in HD mice, regardless of the tissues analyzed, which included liver, spleen and postmitotic brain tissues (Spiro and McMurray, 2003). Interestingly, *Fen1* haploinsufficiency led to reduction of repeat deletions in the germ line, suggesting a role for FEN1 in germline instability. In addition, transfection of wild-type human cells with a mutant FEN1, resulting in inhibition of flap strand cleavage activity, led to contraction of TNR repeats, supporting a role for FEN1 in repeat stability (Spiro and McMurray, 2003).

Alike in HD mice, somatic CTG/CAG instability was unchanged in DM1 mice heterozygous for *Fen1* (van den Broek et al., 2006). In addition, the analysis of early DM1 mouse embryos did not reveal any effect of the deficiency or haploinsufficiency of *Fen1* in CTG/CAG instability (van den Broek et al., 2006).

The final step of LP-BER is DNA sealing of the nick by DNA ligase I (Levin et al., 2000). To address the role of DNA Ligase I *in vivo*, DM1 mice were crossed with knock-in mice carrying a *DNA Ligase I* mutation (Harrison et al., 2002; Tome et al., 2011). In fact the *DNA Ligase I* mutation was from a patient with growth retardation, sun sensitivity and immunodeficiencies (Barnes et al., 1992; Webster et al., 1992) that results in 3-5% residual ligase activity in cells. Somatic CTG/CAG instability in DM1 mice expressing the Ligase I mutant was similar to that in DM1 mice (Tome et al., 2011). However, DM1 mice mutant for Ligase I showed a maternal instability bias, leading to increased contractions and decreased expansions. This effect was observed for both the heterozygous or homozygous states, suggesting a rate-limiting effect of Ligase I level on instability.

These surprising results suggest that *in vivo* FEN1 and Ligase I do not play a major role in repeat somatic CAG/CTG instability. However, the complete inactivation of *Fen1* and *DNA LigI* could not be achieved in mice due to embryonic lethality of full knock-outs (Bentley et al., 1996; Kucherlapati et al., 2002; Larsen et al., 2003). Thus, we cannot exclude that the residual activities or protein levels of FEN1 and Ligase I in HD or DM1 mice are sufficient to induce CAG/CTG instability. In general, assessing the role of BER proteins in CAG/CTG instability using mouse genetics is a difficult task as inactivation of the main BER genes, including *Polβ* (Cabelof et al., 2003; Sobol et al., 1996), *Ape1* (Xanthoudakis et al., 1996), *Xrcc1* (Tebbs et al., 1999), *Fen1* or *DNA LigI* is embryonic lethal.

2.2.2.1.2.4 In vitro *and cell-based studies to gain into mechanistic insights*

2.2.2.1.2.4.1 Accessibility of oxidative lesions at repeats

CAG repeats form stable hairpin structures *in vitro* (Owen et al., 2005). Since the toxic oxidation model by Kovtun et al. proposes that oxidative DNA damage accumulates at CAG/CTG repeats (Kovtun et al., 2007), Jarem et al. tested the hypothesis that oxidative lesions at CAG hairpins might be refractory to excision by OGG1 (Jarem et al., 2009). To this end, Jarem et al. used peroxynitrite to induce 8-oxodG lesions on CAG hairpin substrates (Jarem et al., 2009).

Interestingly, they reported that CAG hairpins represent a hot spot for oxidation as they were more prone to this damage compared to CAG/CTG duplexes. In addition, CAG hairpins were much less efficiently (700-fold) repaired by human OGG1 than CAG/CTG duplexes in reconstituted assays. Noteworthy, further *in vitro* studies reported that hOGG1 binding and activity are comparable between random sequence and repeated sequence duplexes, suggesting that the initiation of repair is similar despite the duplexes DNA sequence but that binding and repair of hairpin structures is less efficient (Jarem et al., 2011). However, APE1 stimulated repair more efficiently in random sequences compared to repeated sequences.

2.2.2.1.2.4.2 Role of FEN1 in CAG/CTG instability

FEN1 processes the excision of the 3' flap strand generated during repair or replication. However, not all DNA structures can be processed by this enzyme as secondary structures formed by CAG/CTG repeats are refractory to excision and this in a length-dependent manner (Spiro et al., 1999). More recent studies report that CTG hairpin structures or CGG repeats are also less efficiently repaired than unstructured flaps *in vitro* (Vallur and Maizels, 2010), suggesting that those structured repeats can be incorporated in DNA and thus increase the expansion size and therefore the instability level.

2.2.2.1.2.4.3 Role of LigI in CAG/CTG instability

DNA ligase I seals the final step in repair or replication. In order to address the role of DNA ligase I in repeat instability, Lopez Castel at al. used extracts of a human cell line deficient for this ligase, which was derived from a patient with a mutation resulting in 3-5% residual activity (Lopez Castel et al., 2009; Webster et al., 1992). As a control, they transfected the wild-type cDNA of Ligase I in the Ligase deficient cells. The different cell types were then transfected with a circular plasmid containing a CAG or CTG slip-out and harboring a nick upstream or

downstream of the repeats with respect to slip-out. Interestingly, incorporation of nucleotides spread in a very long region above the nick and was dependent on the Ligase I background. Repair of the substrates was less efficient for cells transfected with the Ligase I cDNA compared to the Ligase I deficient cells. CTG slipped-out substrates were globally less efficiently repaired by the different cell lines extracts compared to CAG slipped-out substrates, suggesting a role for the higher stability of CTG slip-outs compared to CAGs. In contrast to the results obtained in mice, these results suggest a role for DNA ligase I in repair of CAG/CTG substrates and in CAG/CTG instability.

In these experiments, the modification used for repair assays is a nick and not an oxidant damage. Thus, other repair mechanisms than BER might be involved in the processing of the CAG or CTG slipped-out substrates used in the assay. Other repair mechanisms such as MMR also necessitate the presence of DNA ligase I. In addition, MMR has been involved in repair of the above repeat-containing plasmid substrates (Panigrahi et al., 2010). Finally those results point out to the importance of the relative level of the repair proteins, as the concentration of DNA Ligase I was critical for the repair outcome.

2.2.2.1.2.4.4 Role of LP-BER in CAG/CTG instability

The role of LP-BER in CAG/CTG instability was explored in partially reconstituted assays (Liu et al., 2009). CAG/CTG containing oligonucleotides harboring an 8-oxodG or THF -an AP site analog- modifications were incubated with cell extracts complemented with selected BER recombinant proteins depending on the specific assay. Random sequence oligonucleotides did not show any expanded products following repair, while repair of CAG/CTG substrates resulted in production of expanded products. The production of longer oligonucleotidic products for the CAG/CTG substrates depended on the level of DNA polymerase β, FEN1 and HMGB1, which was shown to promote CAG/CTG repeat expansion probably by stabilizing repeat hairpins formed during BER (Liu et al., 2010b). Increased levels of DNA polymerase β and HMGB1 led to longer

expansion products. Strikingly, increased levels of FEN1 led to production of similar expansion products. Variation of the relative levels of DNA polymerase β and FEN1 showed that the balance between those two proteins regulates repair outcome and formation of expansion products.

These *in vitro* results indicate that the relative levels of BER proteins may be critical in regulating the propensity of TNR for instability. These results indicate that repair of oxidant lesions at CAG/CTG repeats is performed via LP-BER, as proposed by Kovtun et al. (Kovtun et al., 2007).

2.2.2.1.3 Conclusions

Despite many important studies, the mechanism by which BER contributes to TNR instability is still unclear. Precisely, we do not know yet which exact lesions contribute to repeat instability. The study by Kovtun at al. showed that 8-oxodG as well as other oxidant lesions accumulate in the ageing brain (Kovtun et al., 2007). However, only the deficiency of *Ogg1* results in reduction of repeat instability in HD mice. Those results raise the question of the functional overlapping between the cellular DNA glycosylases, and the question of the relative level and the relevance of the different lesions on DNA. Another point is that although some oxidant lesions increased with age in the HD mice brain, no information is known about the level and identity of oxidant lesions at the repeats compared to random genomic sequences. Also, whether specific lesions accumulate with ageing in the tissues presenting high somatic CAG/CTG instability is unknown.

Several studies suggest that the relative levels of repair proteins influence repair and repeat instability. However downregulation of two essential LP-BER enzymes FEN1 or DNA ligase I in HD and DM1 mice does not have an important effect on somatic CAG/CTG instability. Those results suggest that *in vivo*, in mammalian organisms, those proteins are not relevant for repeat instability, or that there are other proteins with overlapping functions that encounter the deficiency of *Fen1* and *Ligase I*, or that the residual level of those proteins in cells is

sufficient to drive efficient repair. Thus, whether BER and more specifically LP-BER is involved *in vivo* in the processing of oxidative lesions at repeats remains unclear. A clue arguing for the involvement of BER in repeat instability is that oxidant lesions occur with a higher propensity in CAG hairpin structures compared to DNA duplexes. Those hairpin-located damages are very slowly processed by DNA glycosylases and lead to impairment of the recruitment and coordination of the following BER steps, suggesting a role for oxidant lesions occurring at repeats in instability. Those hairpin structures were proposed to form spontaneously at CAG repeats but also to be formed during BER following DNA polymerase slippage. The hairpin repeat structures were also shown to be stabilized *in vitro* by Msh2-Msh3 complexes (Owen et al., 2005). (see Figure 4. The “toxic oxidation cycle” model.)

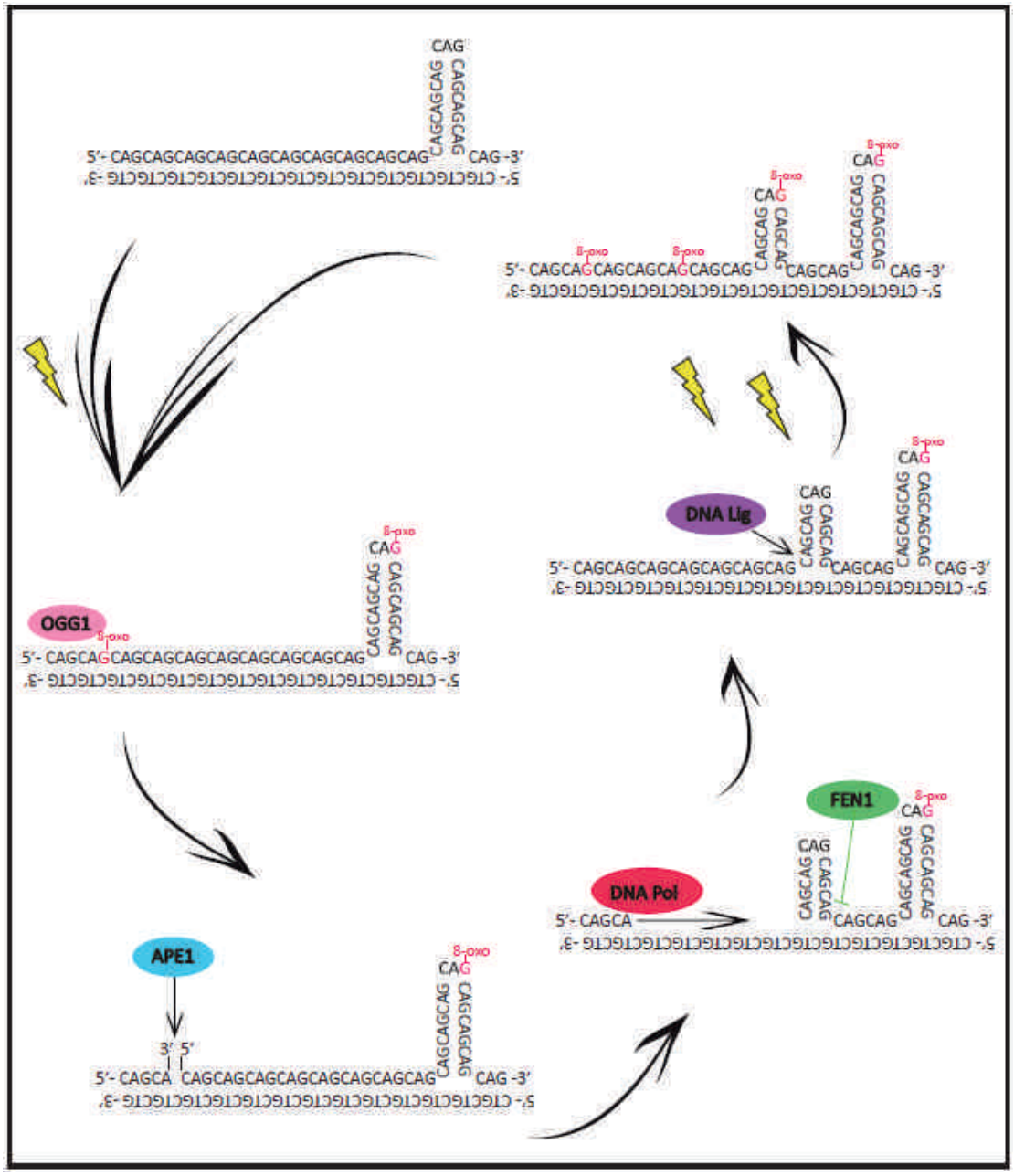


Figure 4. The "toxic oxidation cycle" model. An oxidant damage occurs at the CAG repeats and is processed by the BER machinery: a DNA glycosylase, *e.g.* OGG1, and APE1. A DNA polymerase due to strand displacement activity leads to incorporation of several nucleotides and adjacent to them to formation of a repeat hairpin –which can also be spontaneously formed. This hairpin structure is further stabilized by Msh2 and Msh3 complex and is refractory to excision from FEN1. DNA ligase I seals the nick leading to strand expansion. Oxidant damages are thought to increase with age leading to this error-prone repair and repeat instability.

2.2.2.2 MMR and TNR instability

2.2.2.2.1 MMR contributes to instability of CAG/CTG repeats

The Mismatch repair pathway (MMR) has a role in CAG/CTG repeat instability. It has been shown that inactivation of genes involved in this pathway prevents instability. Both germline and somatic instability are decreased in HD and DM1 mouse models deficient for MMR genes. Precisely, deficiency of *Msh2* in HD transgenic mice carrying the first exon of HTT with a CAG/CTG expansion, prevents repeat instability in all tissues, including the brain and germ cells, demonstrating that MSH2 contributes to CAG/CTG instability (Kovtun and McMurray, 2001; Manley et al., 1999). Furthermore, decreased somatic instability in the brain was associated with the delayed appearance of nuclear inclusions in a knock-in (KI) model of HD, suggesting that reducing somatic CAG/CTG instability is beneficial (Wheeler et al., 2003).

As MSH2 forms heterodimers with MSH3 or MSH6, resulting in MutSβ or MutSα complexes respectively (Acharya et al., 1996), the role of MSH3 and MSH6 in CAG/CTG instability in HD has been investigated (Dragileva et al., 2009). Inactivation of *Msh3* in HD KI mice inhibits somatic instability, and this correlates in the brain with the delayed appearance of nuclear inclusions. In addition, paternal intergenerational instability is decreased, but maternal intergenerational instability is not significantly changed. Somatic CAG/CTG instability is not changed in HD KI mice deficient for MSH6, but intergenerational instability is reduced.

Similar studies have been done in DM1 transgenic mice carrying a CTG/CAG repeat expansion and deficient in *Msh2*, *Msh3* or *Msh6*. As in HD mice, inactivation of *Msh2* and *Msh3* reduces somatic and germline instability (in the case of a maternal or paternal transmission of the mutant allele) (Foiry et al., 2006; Savouret et al., 2004; van den Broek et al., 2002), whereas inactivation of *Msh6* in DM1 mice slightly increases somatic instability, has no effect on intergenerational instability and slightly decreases maternal intergenerational instability. The reduced levels of MSH2 and MSH3 proteins in DM1 *Msh6* deficient mice could account for the effect on maternal intergenerational instability. Thus, MutSβ rather than MutSα contributes to CAG/CTG instability.

So far, the role of MMR in GAA/TTC repeat expansion, which is implicated in FRDA, has been investigated using bacteria- and yeast-based models. In *E. Coli*, MMR has a mitigated effect on GAA/TTC, increasing the frequency of small length changes and decreasing the frequency of large length changes (Schmidt et al., 2000). In yeast, MMR machinery increases fragility of GAA/TTC repeats, resulting in increased double-strand breaks and chromosomal rearrangements. MutSβ and MutLα, a downstream player involved in MMR, contribute to this process (Kim et al., 2008). The role of MMR in GAA/TTC instability remains to be investigated using mammalian models.

2.2.2.2.2 Mechanism of action of MMR in TNR instability

2.2.2.2.2.1. Mutsβ binds stable secondary structures formed by CAG/CTG repeats

CAG/CTG repeats form stable secondary structures as described in §2.2.1.1., resulting in bulky A*A mismatch base pairs in the stem of structured CAGs (Gacy et al., 1995) or T*T pairs within structures CTGs. The propensity of CAG/CTG repeats to form stable slip-out structures increases with the repeat length (Gacy et al., 1995; Pearson and Sinden, 1996). CAG/CTG repeats form *in vitro* short and large slip-outs (Pearson et al., 1998b; Petruska et al., 1996; Zheng et al., 1996). *In vitro,* Msh2 binds to CAG/CTG slipped-strand structures in a repeat length-dependent manner, and with a higher affinity for CAG repeats, introducing thus the concept of asymmetric recognition of the repeats (Pearson et al., 1997). In addition, MutSβ (MSH2/MSH3 dimer) binds to CAG/CTG slippled-strand (or hairpin) structures (Owen et al., 2005). These data support a role for MutSβ in CAG/CTG repeat instability.

2.2.2.2.2.2. Does TNR instability require a functional MMR?

The study by Owen et al. shows MSH2-MSH3 stabilizes CAG/CTG hairpins, and MSH2-MSH3 ATPase activity is inhibited, raising the possibility that CAG/CTG instability results from the stabilization of hairpin structures due to inhibition of MutSβ activity at CAG/CTG repeats (Owen et al., 2005). However, a more recent study shows that the binding of MutSβ on CAG hairpins does not modify the biochemical properties of the protein complex and the repeat-containing hairpins are efficiently repaired by cell extracts (Tian et al., 2009). In addition, Proliferating cell nuclear antigen (PCNA), a protein required at the initiating step of nick-directed mismatch repair (Umar et al., 1996), is necessary for efficient incision of CAG or CTG hairpins (Hou et al., 2009). If CAG hairpins inhibit the activity of MutSβ, one would expect that inactivation of the ATPase activity of MutSβ or suppression of MMR proteins implicated downstream of MutSβ would have no effect on CAG/CTG instability. Mutation of the ATPase domain of *Msh2* in a DM1 mouse model leads to contraction of CTG repeats similar to *Msh2* and *Msh3* deficiencies, suggesting that a functional MSH2 or MMR is required for instability to occur at CAG/CTG repeats (Tome et al., 2009). Additionally, deficiency of *Pms2*, a downstream component of the MMR machinery recruited by MutS complexes, results in the reduction of expansion and induction of large contractions in a DM1 mouse model (Gomes-Pereira et al., 2004), further suggesting that the instability of CAG/CTG repeats requires a functional MMR.

2.2.2.2.2.3. What lesions are targeted by MMR at TNR?

DNA structure plays an important role in MMR. MutSα and MutSβ both repair base-pair mismatches. However, repair of some mispaired DNA does not depend upon MMR (Corrette-Bennett et al., 2001; Littman et al., 1999; McCulloch et al., 2003; Umar et al., 1994). MutSα is involved in repair of non-repetitive insertion/deletion loops (ID) with 1-3 excess nucleotides, whereas MutSβ enhances repair of longer ID (<12 nucleotides) (Genschel et al., 1998; Littman

et al., 1999; Tian et al., 2009). Larger loops are repaired through MMR-independent pathways (Genschel et al., 1998; Littman et al., 1999; McCulloch et al., 2003; Umar et al., 1994).

In vitro, the propensity for hairpin structure formation and the structural complexity of hairpins increases with the length of CAG/CTG repeats (Gacy et al., 1995; Pearson and Sinden, 1996; Pearson et al., 1998b). Specifically, $(CTG)_{50}/(CAG)_{50}$ repeats form two distinct populations of slipped structures, those with less than 10 repeats per slippage, and those involving more than 10 repeats, with a complex structural pattern (Pearson et al., 1998b). Long CAG/CTG slip-outs are poorly processed by MutSβ, and can be repaired by a MMR-independent pathway (Hou et al., 2009; Panigrahi et al., 2005; Panigrahi et al., 2010). In contrast, short slip-outs (1-3 nucleotides) are efficiently processed by MutSβ; however, multiple short slip-outs escape repair, possibly due to interference in MMR (Panigrahi et al., 2010). Taken together, the data suggest that MMR contributes to CAG/CTG instability due to the inefficient processing of multiple short slip-outs (see Figure 5. MMR and TNR instability model.).

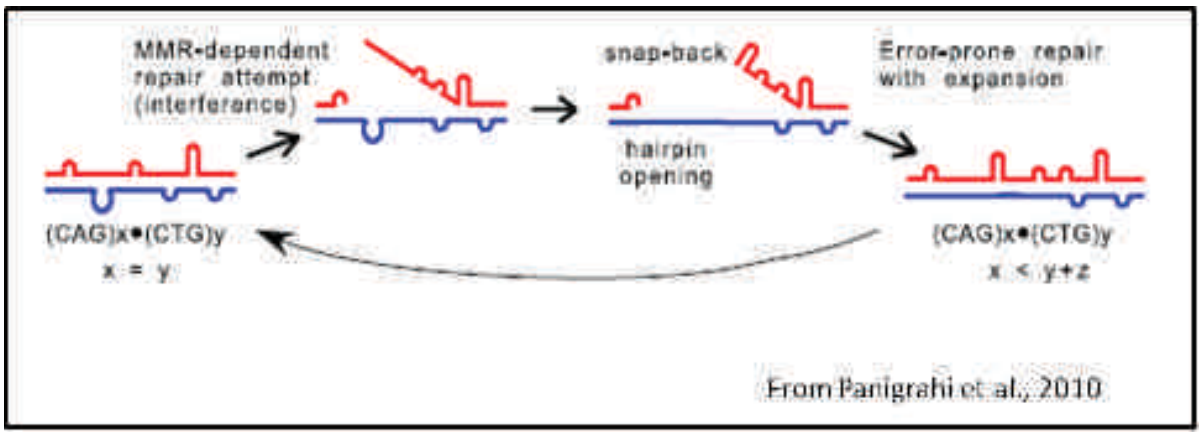


Figure 5. MMR and TNR instability model. CAG/CTG repeats form multiple slip-outs which may initiate repair process but lead to expansion product following error-pone repair. Some initiated repair could be arrested, leading to strand slippage, strand displacement and finally incorporation of structured repeats.

2.2.2.3 Role of Nucleotide Excision Repair pathway (NER) and transcription in regulating TNR instability

2.2.2.3.1 Transcription-coupled NER contributes to instability of CAG/CTG repeats

NER is a DNA repair mechanism involved in the repair of DNA helix distorting damages such as bulky DNA adducts, DNA cross-links, methylated bases as well as DNA loops (Huang et al., 1994; Kirkpatrick and Petes, 1997; Sancar, 1996; Shuck et al., 2008). Transcription coupled NER (TC-NER), as opposed to global genome NER (GG-NER), is involved in the repair of distorting DNA damage in transcribed regions of the genome, and is coupled to active transcription (Venema et al., 1990). The hallmark of TC-NER is the accelerated repair of DNA lesions that block the elongating RNA Polymerase II (RNA Pol II). While common factors are involved in lesion processing in both GG-NER and TC-NER, specific factors contribute to lesion recognition (Figure 7. NER subpathways: GG-NER and TC-NER.). It is assumed that RNA Pol II stalled at a DNA lesion triggers the recruitment of TC-NER factors, including CSA, CSB, XAB2 and HMGN1, whereas in GG-NER the damage-induced DNA distortion is recognized by the UV-DDB (DDB1-DDB2-containing E3-ubiquitin ligase complex) and XPC-RAD23B protein complexes. NER and more precisely TC-NER contribute to TNR instability.

In *E. coli,* NER is mediated by few proteins, including UvrA and UvrB, which process distorting DNA damages (Sancar, 1996). Studies in *E.coli* revealed that deficiency of UvrA, but not UvrB, increases CAG/CTG repeat instability (mostly leading to deletions) in an orientation-dependent manner (Kang et al., 1995; Parniewski et al., 1999). Interestingly, instability of CAG/CTG repeats was greater when the CAG strand was transcribed as compared to the CTG strand and was length-dependent (Kang et al., 1995; Oussatcheva et al., 2001; Parniewski et al., 1999). As a result, the authors propose that transcription transiently dissociates the triplet repeat complementary strands enabling the non-transcribed strand to fold into a hairpin conformation, which is sufficiently stable that replication by-passes the hairpin to give large deletions.

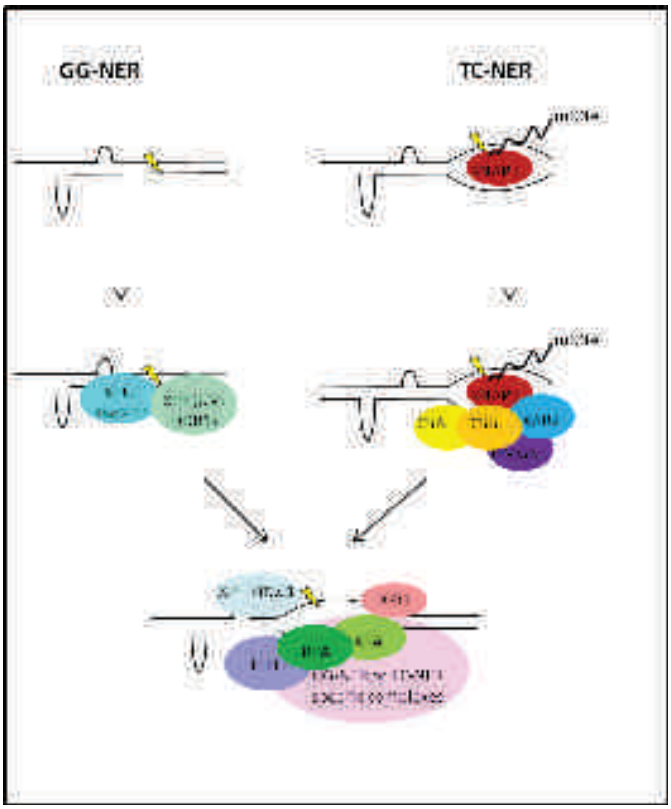


Figure 7. NER subpathways: GG-NER and TC-NER. Main factors participating in each pathway are illustrated. GG-NER: XPC-Rad23B and XPE recognize the damage. TC-NER: the damage induces RNAP II arrest which associates with CSB and CSA that recruit other repair factors. XPA, RPA and TFIIH open DNA and XPG and XPF-ERCC1 excise the damaged strand. Then the DNA pol δ/ε fil the gap. [Adapted by (Fousteri and Mullenders, 2008)]

The role of NER was also assessed in eukaryotes, including in mammalian systems. Repeat instability of CAG/CTGs was assessed in human cells by selective measure of CAG repeat contractions (Lin et al., 2006; Lin and Wilson, 2007). To this purpose, the HPRT minigene, driven by the Tet-ON inducible promoter, was modified to contain 95 CAG repeats, resulting in *HPRT*$^{-}$ HeLa cells. Contraction to <39 CAG repeats permit cells to survive *HPRT*$^{+}$ selection. Knock-down of XPA, a central component of NER, by siRNA led to reduced contraction frequency, whereas siRNA against XPC, a GGR component, did not alter the contraction rate (Lin et al., 2006). siRNA of ERCC1 and XPG decreased contraction frequencies, similar to XPA siRNA (Lin and Wilson,

2007). Finally, knock-down of CSB, which is involved in TC-NER, also decreased CAG/CTG instability (Lin and Wilson, 2007). Those data suggest that NER and specifically TC-NER are involved in CAG/CTG instability in HeLa cells. Interestingly, knock-down of MMR factors such as Msh2 and Msh3 also resulted in repeat stabilization. Noteworthy, decreasing the expression of both Msh2 and XPA did not further stabilize CAG/CTG repeats, suggesting that MMR and NER act in a same pathway.

Studies in SCA3 Drosophila melanogaster models showed that inactivation of Mutβ, the orthologue of the human Rad2/XPG, resulted in reduction of the instability level in the germline, further suggesting that NER is involved TNR instability (Jung and Bonini, 2007).

Xpc deficiency did not affect germline and somatic instability of CAG/CTG repeats in HD knock-in mice (Dragileva et al., 2009), suggesting that GG-NER does not play a predominant role in TNR instability in mice. *Xpa* deficiency did not alter germline repeat instability in SCA1 mice with a CAG/CTG expansion at the murine *SCA 1* locus (Hubert et al., 2011). Similarly, repeat instability in somatic tissues, such as the kidney and liver, was not affected by *Xpa* deficiency. However, neuronal tissues, such as the striatum, hippocampus and cortex, showed an important decrease of repeat instability. Those results suggest that NER regulates TNR instability in post-mitotic tissues. In addition, knock-out of *Csb* in R6/1 HD mice resulted in limited repeat instability in brain and germline (Kovtun et al., 2011). Apart from a role in TC-NER, it has been reported that CSB is involved in oxidant damage processing (Osterod et al., 2002; Sunesen et al., 2002; Thorslund et al., 2005; Trapp et al., 2007; Tuo et al., 2002; Tuo et al., 2001). Interestingly, R6/1 mice deficient for both *Csb* and *Ogg1* led to opposite effects than OGG1 and CSB deficiencies separately, and thus to exacerbation of instability in brain and liver, suggesting that CSB does not participate in repeat instability by oxidant lesions processing (Kovtun et al., 2011). Taken together those results suggest that TC-NER is involved in somatic TNR instability.

2.2.2.3.2 Transcription and TNR instability

All TNR loci are located in transcribed regions, suggesting that transcription might influence repeat instability. Evidence of the involvement of transcription in CAG/CTG repeat instability was first described in *E. coli* where untranscribed tracts were stable, whereas transcribed tracts were unstable (Bowater et al., 1997; Schumacher et al., 2001). In both conditions, when the CTG repeat strand was used as leading-strand synthesis template -defined as orientation I during replication-, the repeats were more stable compared to when the CAG repeat strand was used as leading-strand template -defined as orientation II-, suggesting that the role of transcription in the increase of repeat instability is orientation-dependent (Kang et al., 1995; Mochmann and Wells, 2004; Parniewski et al., 1999). In a *D. melanogaster* model of SCA 3, transcription of CAG/CTG repeats led to increased repeat instability (Jung and Bonini, 2007). Cell-based models also suggest a role for transcription in TNR instability. Study based on a HeLa cell-based model allowing for detection of contraction events at CAG/CTG repeats showed that the transcription elongation factor TFIIS contributes to repeat instability (Lin and Wilson, 2007). Furthermore, inhibition of the proteasome and downregulation of BRCA1/BARD1 proteins, which modulate RNA Pol II activity, decreased CAG/CTG instability (Lin and Wilson, 2007). Finally, in mice, a role for transcription in TNR instability is also suggested. Somatic CAG instability is only seen in HD R6 transgenic mouse lines that express the HD transgene. CAG expansion is stable in the R6/0 line, which does not express the HD transgene (Mangiarini et al., 1997).

Transcription contributes to TNR instability in cell models and in lower organisms. However, the level of expressed of expanded genes does not correlate with the level of repeat instability in somatic tissues in DM1 and HD mice (Guiraud-Dogan et al., 2007; Lia et al., 1998; Seznec et al., 2000). Studies in human fibroblasts transformed to contain 800 CTG repeats showed that one-way transcription through repeats increased the repeat instability rate and bidirectional transcription exacerbated this effect (Nakamori et al., 2011). Bidirectional transcription has been proved for DM1, SCA 8, SCA 7, HD and FRAXA (Cho et al., 2005; Chung et al., 2011; Ikeda et al., 2008; Ladd et al., 2007; Loesch et al., 2011; Nemes et al., 2000; Sopher et

al., 2011). To which extent bidirectional transcription is involved in instability *in vivo* needs to be explored. Additionally, whether bidirectional transcription contributes to toxicity of expanded repeats needs investigation, since stalling of RNAPII/RPA at repeats induced the ATR signaling pathway, leading to apoptosis in a cell-based model (Lin et al., 2010b).

GC rich DNA sequences were shown to induce transcription blockage by formation of stable DNA:RNA hybrids or R-loops (Belotserkovskii et al., 2010). Using cell-based models, it was shown that stable R-loops also form at transcribed TNR repeats (Lin et al., 2010a; Reddy et al., 2010). The propensity to form R-loops at TNR was sequence- and length-dependent (Reddy et al., 2010). Although interrupted CGG or CAG repeats did not hinder loop formation, interruptions in GAA repeats led to reduced transcription inhibition (Reddy et al., 2010; Sakamoto et al., 2001a; Sakamoto et al., 2001b). Additionally, it was shown that R-loops contribute to TNR instability. Specifically knocking-down RNAses H, which removes R-loops, reduced TNR instability in a cell-based model (Chon et al., 2009; Itaya, 1990; Lin et al., 2010a).

2.2.3.3.3 Mechanistic model for a role of transcription and TC-NER in CAG/CTG instability

The following model, which takes into account the data presented above, has been proposed (Figure 8. Transcription induced repeat instability.). Strand dissociation during transcription would promote formation of secondary structures such as hairpins or slip-outs. Transcription may also lead to formation of stable R-loops. The intrastrand secondary structures -potentially further stabilized by MutSβ- as well as R-loops may impede RNAP II progression and transcription elongation (Lin and Wilson, 2007; Salinas-Rios et al., 2011). Arrest of RNAP II at non canonical DNA structures may result in gratuitous TC-NER, as proposed (Hanawalt and Spivak, 2008), thereby leading to TNR instability. In addition, TC-NER and MMR possibly interplay to promote TNR instability. (Figure 8. Transcription induced TNR instability)

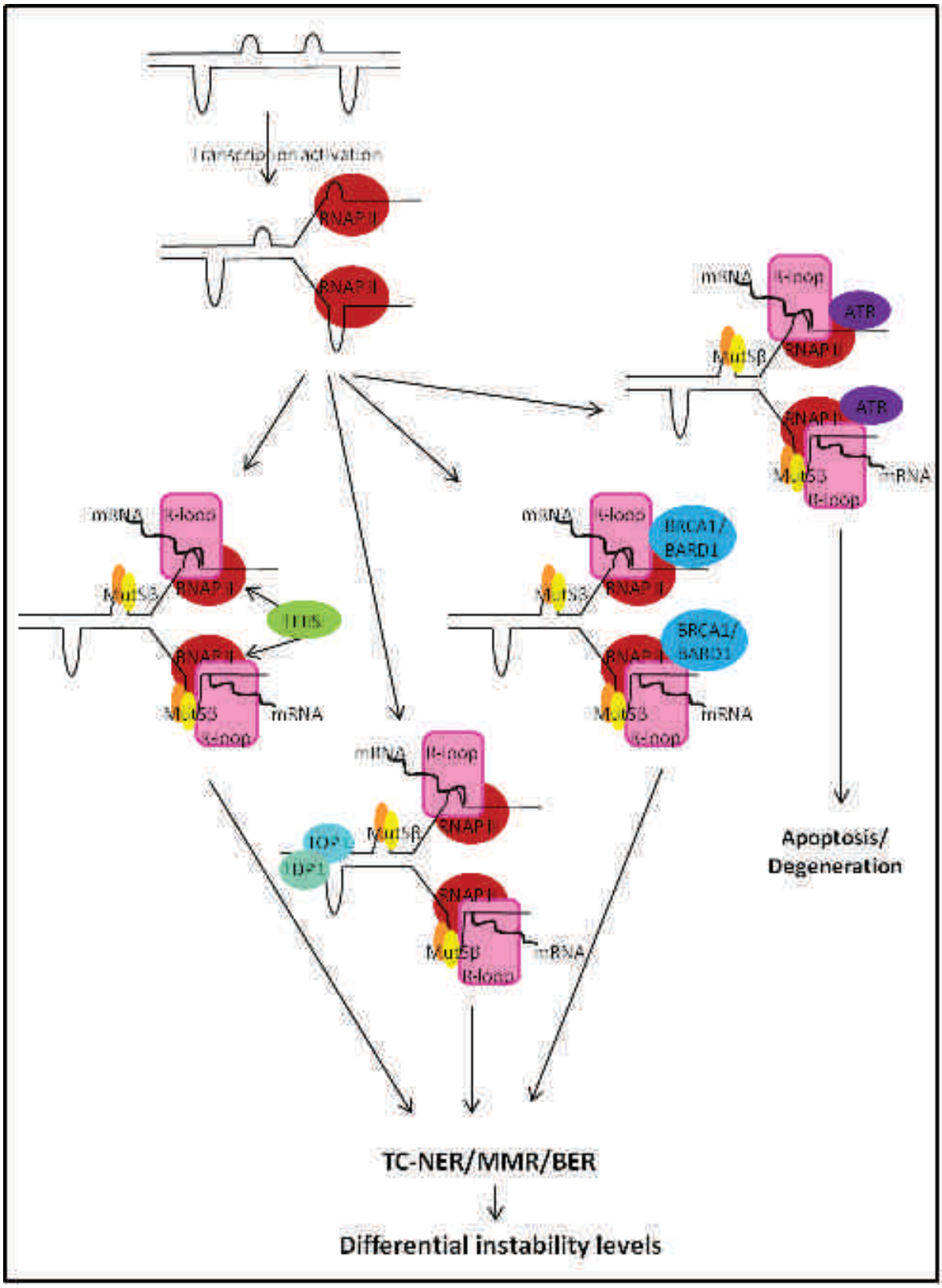

RNAPII
RNAPII
mRNA
R-loop
ATR
MutSβ
RNAPII
mRNA
R-loop
ATR
RNAPII
MutSβ
mRNA
R-loop
RNAPII
MutSβ
TFIIS
RNAPII
mRNA
R-loop
mRNA
R-loop
BRCA1/
BARD1
RNAPII
MutSβ
BRCA1/
BARD1
RNAPII
mRNA
R-loop
mRNA
R-loop
RNAPII
MutSβ
mRNA
R-loop
Apoptosis/
Degeneration
TC-NER/MMR/BER
Differential instability levels

Figure 8. Transcription induced TNR instability. Initiation of the transcription facilitates the formation of stable secondary structures –hairpins and slip-outs- by loosening the DNA structure. Those structures promote arrest of RNAP II processing thus inducing its backtracking by TFIIS, its ubicuitination by BRCA1/BARD1 or recruitment of ATR-dependent apoptosis factors. Stable R-loops form at the repeats and also promote RNAP II arrest. During transcription TOP I and TDP1 are involved in the regulation of the relaxation of supercoilings. R-loops, TFIIS, BRCA1/BARD1, TOP I and TD1 action results in recruitment of repair factors of TC-NER, MMR and BER and leads to differential instability levels. Damages present at the repeats during transcription may exacerbate the repair mechanisms and instability.

2.2.2.4. Double strand break repair, recombination and TNR instability

Double strand breaks repair (DSBR) is another mechanism that has been implicated in TNR instability. Double strand break are repaired through homologous recombination (HR) or non homologous end-joining (NHEJ). Those two mechanisms share some repair proteins; however HR uses the sister chromatide as DNA template, which is available during the S and G2 phases of the cell cycle, whereas NHEJ uses single-stranded overhangs on the ends of double strand breaks (Lieber, 2008; Mao et al., 2008; Thompson and Schild, 2001; Weterings and Chen, 2008).

In *E.coli,* DSBR was shown to increase CAG/CTG repeat instability by intermolecular repair in an orientation-dependent manner (Hebert et al., 2004; Hebert and Wells, 2005). Deficient DSBR at GAA/TTC repeats also led to increase of repeat instability, however in repeat length- and orientation-independent manners (Pollard et al., 2008). Using yeast models, it has been shown that deficiency in recombination factors, including *Rad51* or *Rad52*, suppresses TNR instability (Kerrest et al., 2009). Additionally, overexpression of the *MRE11/RAD50/XRS2* complex or deficiency of *Mre11/Sae2/Tel1* led to fragilization of repeat sites and repeat instability, supporting a role for DSBR in TNR instability in yeast (Richard et al., 2000;

Sundararajan et al., 2010). Double strand breaks were suggested to be generated at TNR in a repeat length-dependent manner during replication, resulting in repeat instability (Freudenreich et al., 1998; Richard et al., 1999; Richard et al., 2000). DSBR during meiosis might also contribute to TNR instability, as shown in yeast (Jankowski et al., 2000). In COS cells, DSBR led to increased instability of CAG/CTG repeats by preferential deletion of repeats. This polarized effect was further enhanced in presence of intrastrand slipped structures (Marcadier and Pearson, 2003). In a similar manner, double strand breaks directed to CAG/CTG repeats led to increased repeat contractions in CHO cells (Mittelman et al., 2009). Whether DSBR is an important player of TNR instability in mammals remains to be demonstrated.

2.2.3. Replication and TNR instability

DNA replication has also been involved in TNR instability (Lopez Castel et al., 2010; McMurray, 2010; Pearson et al., 2005). Several studies in yeast have shown that mutations in replication proteins, including DNA ligase I and FEN1, both implicated in Okasaki fragment processing, affect CAG/CTG instability (Callahan et al., 2003; Freudenreich et al., 1998). Moreover, replication inhibitors modulate the instability of CTG/CAG repeats in cells (Liu et al., 2010a; Yang et al., 2003). The location of replication origins, with respect to distance and orientation of CAG/CTG repeats, also modulates instability in cell model systems (Cleary et al., 2002; Liu et al., 2010a). Replication-dependent instability is believed to arise during lagging strand synthesis via DNA slippage during Okasaki fragments processing. A recent study further showed that replication fork progression towards CTG/CAG repeats varies between tissues in DM1 mice and CTCF sites contribute to pausing, favoring a fork-shift model of replication-mediated instability (Cleary et al., 2010). Interestingly, decreased pausing occurred with aging in animal testes, correlating with increased instability with time. However, the correlation between the rate of replication fork progression and the level of instability in somatic tissues was not obvious, suggesting that other mechanisms are critical (likely transcription and/or DNA repair).

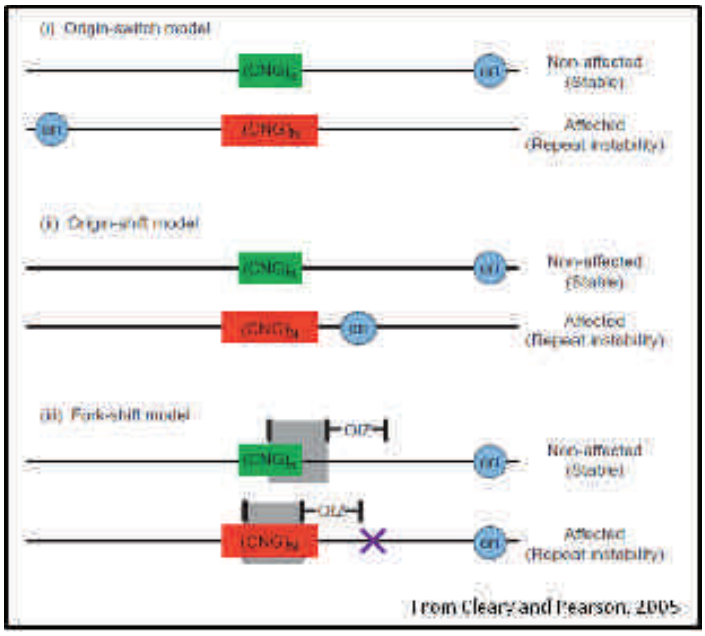


Figure 9: Current models implicating replication origin to the TNR instability. i) Origin-swift model: the replication origin orientation relative to the repeats suggests an influence on the repeat instability, ii) Origin-shift model: the distance of the origin replication relative to the repeats influences the repeats stability, iii) Fork-shift model: the presence of cis-elements (e.g. CTCF binding sites) between the replication origin and the repeats might influence repeat instability. (ori: replication origin, OIZ: Okazaki initiation zone) [Adapted by (Cleary and Pearson, 2005)]

2.2.4 Epigenetics and TNR instability

Chromatin is classically subdivided in the transcriptionally active euchromatin and the transcriptionally silent, condensed heterochromatin. However the chromatin structure at a specific locus is not frozen and changes during development and cell differentiation. TNR instability would result from abnormal processing of stable secondary structures formed by

repeats (*i.e.* hairpins, slip-outs). Whether chromatin structure or chromatin regulators contribute to TNR instability, modulating formation and/or processing of those secondary structures, is an emerging issue.

2.2.4.1 Chromatin structure and remodelers and TNR

2.2.4.1.1 Nucleosome assembly and TNR

Nucleosome assembly is essential for functional organization of chromatin (Kaplan et al., 2009; Segal et al., 2006; Valouev et al., 2011). Euchromatic regions (especially active promoters) show lower nuclosome spacing compared to heterochromatic regions (Valouev et al., 2011). Additionaly, CG rich regions were shown to promote nucleosome occupancy *in vitro*, but *in vivo* CpG islands showed low nucleosome coverage (Hughes and Rando, 2009; Valouev et al., 2011).

Interestingly, the length and methylation state of CGG repeats was shown to regulate nucleosome assembly *in vitro* (Godde et al., 1996). Methylation of non-pathological CGG repeat size promotes nucleosome assembly compared to unmethylated repeats, whereas methylation of longer CGG trackts reduces the efficiency of nucleosome assembly *in vitro* (Godde et al., 1996; Wang and Griffith, 1996). In contrast, the efficiency of nucleosome formation increased with CTG repeats length (Wang et al., 1994). Additionaly, CTG repeats -even as short as found *in vivo*- stabilized the repeats-nucleosome interaction in a repeat length-independent manner *in vitro* (Godde and Wolffe, 1996). Noteworthy, interruption of CAG repeats by CAT repeats decreased nucleosome assembly, whereas AGG interruption of CGG repeats did not affect it (Mulvihill et al., 2005). Hyperacetylated histones led to increased nucleosome assembly in AGG interrupted repeats compared to pure tracts, and the effect was abolished upon CpG methylation. Those results suggest TNR sequences have a direct effect on chromatin structure. Additionnaly, interrupted repeats showed a nucleosome assembly more similar to random sequences when compared to pure tracts, suggesting nucleosome assembly dynamics might play a role in repeat instability (Mulvihill et al., 2005). Mice containing identical CGG repeat number within a *FMR1* transgene presented lower repeat instability in absence of SV40ori/EPR,

a nucleosome excluding sequence, further supporting the role of nucleosome assembly dynamics in repeat instability (Baskaran et al., 2002; Datta et al., 2011).

TNR sequences also directly influence the propensity for methylation. For instance, CGG/CCG repeats, which form CpG islands, appear strongly methylated, and this was associated with heterochromatinization of FRAXA locus and gene silencing (Hornstra et al., 1993; Oberle et al., 1991; Sutcliffe et al., 1992).

2.2.4.1.2 Heterochromatinization at TNR loci

Some TNR diseases have been associated with altered histone modifications. Histone deacetylation has been associated with transcriptional inactivity (Lorincz et al., 2004). In FRAXA the diseased allele is silenced whereas the premutation allele is more expressed than the wild-type allele. Interestingly, lower levels of acetylated H3 and H4 were observed at the *FMR1* 5' region in FRAXA patient cells, the H4 acetylation level being inversely correlated with the CGG length (Coffee et al., 2002; Coffee et al., 1999). The *FMR1* 5' region was also associated with decreased methylation of H3K4 and increased methylation of H3K9, suggesting a switch from euchromatin to heterochromatin (Coffee et al., 2002). Those results suggest that the chromatin condensation level differs according to repeat length. Additionally, in cells from premutation carriers, acetylation of histones was increased at the FMR1 locus compared to control cell lines, in accordance with increased *FRM1* mRNA levels (Todd et al., 2010). In accordance with those results, overexpression of histone deacetylases (HDAC) or treatment with histone acetyltransferase (HAT) inhibitors in a *D. melanogaster* or a cell model for premutation FRAXA induced transcriptional silencing of FMR1 and suppressed CGG repeat-induced neurodegeneration (Todd et al., 2010). Similarly, inhibition of the HDAC SIRT1 reactivated *FMR1* transcription (Biacsi et al., 2008).

Increase of H3K9 di- and tri-methylation and decrease of acetylation of H3 and H4 was also observed downstream of *FXN* GAA repeats in FRDA cell lines, suggesting heterochromatinization of the expanded *FXN* locus (Greene et al., 2007; Herman et al., 2006;

Soragni et al., 2008). The decreased H3k9 acetylation levels and increased H3K9 methylation levels were also found in brain of FRDA patients and in a FRDA mouse model (Al-Mahdawi et al., 2008). Consistent with heterochromatinization of FXN locus, the levels of H3K4me3 and initiating RNA Pol II, which are associated with transcription initiation, and the levels of H3K36me3, which is associated with transcription elongation, were lower in FRDA alleles (Kumari et al., 2011). Similar to FRAXA studies, the use of HDAC inhibitors in patients' lymphocytes reactivated *FXN* transcription, and, when used on a FRDA mouse model, partially corrected FXN deficit (Herman et al., 2006; Rai et al., 2008; Sandi et al., 2011). Heterochromatinization has also been associated with DM1 (Cho et al., 2005). Finally, HDAC inhibitors were shown to prevent repeat instability in a *D. melanogaster* SCA 3 model (Jung and Bonini, 2007). This might suggest that heterochromatinization of TNR loci protects from instability. However, the relationship between heterochromatinization and instability at TNR loci remains elusive and warrants further investigations.

2.2.4.1.3 CTCF sites and TNR instability

The DNA sequences flanking the repeats might also have a role in repeat instability, providing specific chromatin context and binding to specific chromatin modifiers. CTCF is a chromatin insulator zing-finger protein that contains a consensus binding site and associates with CG-rich regions (Filippova et al., 1996). Consensus CTCF binding sites frequently flank TNR loci, including DM1, HD, FRAXA, SCA 2, SCA 7 and DRPLA loci (Filippova et al., 2001; Ladd et al., 2007). The role of CTCF sites on TNR instability was clearly demonstrated using SCA7 mice. Mutation of the CTCF site downstream of the CAG repeats (CTCFI site) resulted in increased somatic and germline instability in SCA7 transgenic mice (see Figure 9. CTCF and TNR instability model.) (Libby et al., 2008). Interestingly, methylation of CTCFI site had a similar effect. Additionally, CTCF was necessary for expression of an antisense transcript of *ataxin-7*, which appears to regulate sense transcription (Sopher et al., 2011), suggesting that CTCF might

indirectly regulate TNR instability by regulation of sense and antisense transcription. Two functional CTCF binding sites flanking the CTG repeats that are methylation-sensitive have been identified at the DM1 locus (Cho et al., 2005; Filippova et al., 2001). In control human cell lines, CTCF binding at the DM1 locus reduced expression of an antisense transcript and H3K9 methylation, whereas in cells from congenital DM1 patients, CTCF binding at the DM1 locus is lost and this is accompanied with spreading of heterochromatin and increased expression of antisense transcript (Cho et al., 2005). Additionally, methylation of CTCF sites at the DM1 locus has also been observed. Interestingly, methylation of DM1 CTCF sites is specific to the expanded allele and is tissue-selective (Lopez Castel et al., 2011). However, there was no clear correlation between somatic CTG instability and CTCF site methylation and more globally methylation of the regions flanking CTG repeats in tissues from DM1 mice (Lopez Castel et al., 2011). Furthermore, it was shown that CTCF sites at the DM1 locus demarcate replication origins and CTCF sites reduced replication progression (Cleary et al., 2010). Moreover, using different tissues from DM1 mice, it was shown that replication progression was dependent on methylation state at the DM1 locus, suggesting a replication-dependent effect of CTCF on CTG instability (Cleary et al., 2010). Finally, depletion of CTCF led to heterochromatinization of the mutant *FXN* locus, increased antisense transcript expression and silencing of FXN in FRDA cells (De Biase et al., 2009). However, the role of CTCF in GAA instability has not yet been investigated. In conclusion, CTCF sites flanking TNR loci might regulate TNR instability through various mechanisms, including regulation of sense/antisense transcription and replication progression. Furthermore, methylation of CTCF sites or, more globally, regions flanking TNR loci appears to be key in modulating TNR instability. This role of DNA methylation in TNR instability is further supported by the study by Dion et al., showing that haploinsufficiency of the DNA methyltransferase *Dnmt1* leads to decreased DNA and histone methylation, and to increased intergenerational CAG repeat instability in SCA 1 mice (Dion et al., 2008).

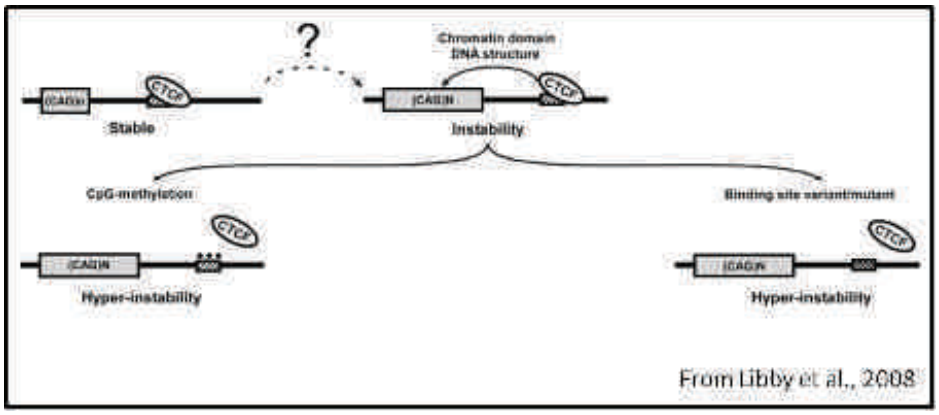


Figure 9. CTCF and TNR instability model. Binding of CTCF in its site protects CAG repeats form expansion. Once repeats expand, they alter the chromatin environment and DNA structure around them leading to instability. Loss of CTCF binding in its site following its mutation or CpG methylation leads to repeat instability, further suggesting that chromatin environment and repeats structures affect repeat instability.

3 Thesis outline

The goal of this thesis was to investigate the implication of different factors in the tissue-selectivity of somatic CAG/CTG instability to improve the understanding of TNR instability and TNR diseases. Several physiological cellular mechanisms, including DNA repair, have been involved in triplet repeat instability. I focused my interest on the role of oxidant lesions and BER in the tissue-selectivity of somatic CAG/CTG instability. For this purpose, I used both *in vitro* and *in vivo* approaches.

1) We first analyzed the role of oxidant lesions and the implication of BER in the tissue selectivity of somatic CAG/CTG instability, by using HD transgenic mice. Specifically, we compared the levels of oxidant lesions and BER in HD mouse striatum and cerebellum, i.e. in two tissues showing opposite levels of repeat instability (CAG/CTG instability is high in the striatum versus low in the cerebellum). We observed that the level of oxidant lesions was increased at repeats both in the striatum and in the cerebellum of HD mice, and therefore did not correlate with the level of somatic instability. We also showed that oxidant lesions within hairpin structures were refractory to repair by BER. In addition, we found that BER protein levels and activities were reduced in the striatum of HD mice when compared to the cerebellum. In particular, FEN1 and LigI, two enzymes involved in LP-BER, were lower in the striatum than in the cerebellum of HD mice. Thus, BER proteins and activities correlated to some extent with the tissue selectivity of CAG/CTG instability in HD mice.

2) We assessed the stoichiometry of major BER proteins at a molar level in the striatum and in the cerebellum of HD mice. By using fully reconstituted BER assays and various oligonucleotide substrates, we then assessed the role of BER stoichiometry on the efficiency and mode of repair. We found that nucleotide sequence, protein stoichiometry and lesion positioning influenced repair. In particular, our data showed

that a mixture of BER proteins reflecting the stoichiometry found in the mouse striatum resulted in a less efficient repair, in comparison to a protein mixture reflecting the stoichiometry seen in the cerebellum. In addition, our results suggested that CAG/CTG-containing substrates are preferentially repaired through LP-BER, and lesion position within repeats modulates the requirement for LP-BER.

Taken together, those results suggest that BER stoichiometry contributes to the tissue selectivity of somatic CAG/CTG instability in HD.

II) Results

Chapter 1: Stoichiometry of Base Excision Repair Proteins Correlates with Increased Somatic CAG Instability in Striatum over Cerebellum in Huntington's Disease Transgenic Mice

Goula AV, Berquist BR, Wilson DM 3rd, Wheeler VC, Trottier Y, Merienne K.
PLoS Genet. 2009 Dec;5(12):e1000749. Epub 2009 Dec 4

Introduction:

Huntington's disease is one of the most common repeat associated neurodegenerative disorders. This dominant inherited disease is caused by abnormal expansion of CAG repeats in the first exon of the coding sequence of Huntingtin (*HTT*). The mutation confers a toxic gain of function to the resulting protein relative to the abnormal repeats length. The expansion is unstable in the germline and in somatic tissues. Somatic instability is thought to contribute to disease progression as the higher instability level and thus the longer and more toxic expansions are present in the striatum, i.e. the tissue that preferentially degenerates. Thus understanding the factors that participate in the repeat instability process is essential to identify disease modifiers.

Several mechanisms have been implicated in the repeat instability process (see § I. Introduction). Recently the idea of the toxic instability has been supported by the observation that in HD mice deficient for *Msh2* or *Msh3*, two MMR genes, somatic instability was reduced and the pathology was delayed (Dragileva et al., 2009; Wheeler et al., 2003). Somatic instability was also inhibited in HD mice deficient for OGG1, a BER initiating glycosylase (Kovtun et al., 2007). Taking into consideration that oxidant damage was reported to increase with age in HD mouse brain but that OGG1 was not age related, it has been hypothesized that the accumulation of oxidant damage with age could lead to the age-dependent increase of somatic CAG instability in HD. In addition, involvement of BER in CAG repeat instability was supported by *in vitro* and yeast studies showing that Polβ promotes and FEN1 prevents repeat expansion, respectively (Freudenreich et al., 1998; Kovtun et al., 2007).

BER is involved in the elimination of oxidant DNA damage (Imam et al., 2006). It is composed of several highly coordinated steps as previously presented (see § 2.2.2.1.1. in the Introduction part). In addition, few studies indicate that BER is regulated in tissues-specific and age-dependent manners (Imam et al., 2006), raising the possibility that BER activity may underlie the tissue-selectivity or age-dependency of CAG instability. In this study we have addressed the role of oxidant DNA damage and BER mechanism in the tissue-selectivity and age-dependence of CAG/CTG repeat associated instability in HD.

Results:

To address those questions we used R6/1 and R6/2 HD transgenic mouse models (Mangiarini et al., 1997) that recapitulate major features of the human pathology, including the somatic CAG instability seen in patients. Those mouse models contain the first exon of human *HTT* including ≈125 or ≈160 CAG repeats respectively, and show a reduced life span associated with motor and cognitive impairments (Mangiarini et al., 1996b). In the study we tested two tissues showing opposite levels of CAG/CTG instability: the striatum, the tissue that preferentially degenerates in HD and shows high instability levels, and the disease spared cerebellum, which shows limited instability levels.

To examine whether accumulation of DNA damage in the aging brain contributes to the somatic CAG instability (Kovtun et al., 2007), we quantified the global genome AP site level in the striatum and cerebellum of R6/1 and R6/2 mice, by using an ELISA-based detection system. Apuric/Apyrimidic (AP) sites, intermediates generated during BER, increased with age, and was higher in the cerebellum compared to the striatum, in both wild-type (WT) and HD mice. Therefore, the global level of AP sites did not correlate with the tissue-specific instability pattern. In order to specify if oxidative DNA damage accumulates at CAG repeats, we measured the oxidant damage level at CAG repeats in the striatum and cerebellum of R6/1 mice. To this purpose, the tissular DNA was digested by the bacterial glycosylase *Fpg* and then submitted to quantitative PCR. Noteworthy, DNA damage level was tissue- and age-independent, but it abnormally accumulated at CAG repeats. In order to examine whether the spontaneously formed CAG hairpin structures (Owen et al., 2005) participates in the oxidant damage accumulation at CAG repeats, we used an *in vitro* assay to test whether hairpin structures would impair the access of BER enzymes. We found that 8-oxoG and AP lesions were refractory to incision by purified OGG1 and APE1, respectively, suggesting that oxidant damage accumulates at CAG repeats due to structural hindrance.

In order to test whether BER could play a role in the somatic CAG repeats instability, we measured main BER enzyme expression and activity. Western Blot analyses revealed that the protein levels of APE1 and FEN1 were increased in the cerebellum compared to the striatum in

both R6/1 and WT mice. In addition, we measured main BER enzymes activities in R6/1 and WT mice at different ages. Specifically modified oligonucleotidic substrates for each activity -OGG1, APE1, DNA Polymerase, and FEN1- were incubated with striatum or cerebellum extracts. Globally BER protein activities were increased in the cerebellum compared to the striatum for all ages, except that OGG1 activity which was similar between the different conditions. Strikingly, FEN1 -a LP-BER protein- activity was 5- to 10-fold increased in the cerebellum compared to the striatum. In addition, relative activities of FEN1/Polβ were lower in the striatum versus the cerebellum, suggesting that coordination between DNA synthesis and 5-flap removal might be reduced in the striatum. Furthermore, chromatin immunoprecipitation experiments showed enrichment of Polβ at the CAG repeats in the striatum but not in the cerebellum of R6 mice. *In vitro* repair assay showed that a FEN1/Polβ ratio reflecting the situation in the striatum is more prone to multi-nucleotide DNA synthesis than a FEN1/Polβ ratio reflecting the situation in the cerebellum.

Conclusions:

Our data suggest that the stable secondary structures formed by CAG sequences result in accumulation of oxidant DNA damage at CAG repeats. We propose that some DNA lesions at repeats become accessible to BER enzymes, due to the dynamic nature of the secondary structures, and the lesions are preferentially processed through LP-BER. Moreover, we propose a model where the inefficient cooperation between FEN1 and Polβ in the striatum contributes to somatic CAG instability. In contrast, in the cerebellum the efficient cooperation between Polβ and FEN1 would protect from CAG instability. Thus, the tissue-specific BER stoichiometry would contribute to the tissue-selective somatic CAG instability in HD mice.

Publication 1

OPEN ACCESS Freely available online PLoS GENETICS

Stoichiometry of Base Excision Repair Proteins Correlates with Increased Somatic CAG Instability in Striatum over Cerebellum in Huntington's Disease Transgenic Mice

Agathi-Vassiliki Goula[1], Brian R. Berquist[2], David M. Wilson, III[2], Vanessa C. Wheeler[3], Yvon Trottier[1]*, Karine Merienne[1]*

1 Department of Neurobiology and Genetics, Institute of Genetics and Molecular and Cellular Biology (IGBMC), UMR 7104-CNRS/INSERM/UdS, Illkirch, France, **2** Laboratory of Molecular Gerontology, National Institute on Aging (NIA)/National Institutes of Health (NIH), Baltimore, Maryland, United States of America, **3** Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, United States of America

Abstract

Huntington's disease (HD) is a progressive neurodegenerative disorder caused by expansion of an unstable CAG repeat in the coding sequence of the Huntingtin (*HTT*) gene. Instability affects both germline and somatic cells. Somatic instability increases with age and is tissue-specific. In particular, the CAG repeat sequence in the striatum, the brain region that preferentially degenerates in HD, is highly unstable, whereas it is rather stable in the disease-spared cerebellum. The mechanisms underlying the age-dependence and tissue-specificity of somatic CAG instability remain obscure. Recent studies have suggested that DNA oxidation and OGG1, a glycosylase involved in the repair of 8-oxoguanine lesions, contribute to this process. We show that in HD mice oxidative DNA damage abnormally accumulates at CAG repeats in a length-dependent, but age- and tissue-independent manner, indicating that oxidative DNA damage alone is not sufficient to trigger somatic instability. Protein levels and activities of major base excision repair (BER) enzymes were compared between striatum and cerebellum of HD mice. Strikingly, 5′-flap endonuclease activity was much lower in the striatum than in the cerebellum of HD mice. Accordingly, Flap Endonuclease-1 (FEN1), the main enzyme responsible for 5′-flap endonuclease activity, and the BER cofactor HMGB1, both of which participate in long-patch BER (LP-BER), were also significantly lower in the striatum compared to the cerebellum. Finally, chromatin immunoprecipitation experiments revealed that POLβ was specifically enriched at CAG expansions in the striatum, but not in the cerebellum of HD mice. These *in vivo* data fit a model in which POLβ strand displacement activity during LP-BER promotes the formation of stable 5′-flap structures at CAG repeats representing pre-expanded intermediate structures, which are not efficiently removed when FEN1 activity is constitutively low. We propose that the stoichiometry of BER enzymes is one critical factor underlying the tissue selectivity of somatic CAG expansion.

Citation: Goula A-V, Berquist BR, Wilson DM, III, Wheeler VC, Trottier Y, et al. (2009) Stoichiometry of Base Excision Repair Proteins Correlates with Increased Somatic CAG Instability in Striatum over Cerebellum in Huntington's Disease Transgenic Mice. PLoS Genet 5(12): e1000749. doi:10.1371/journal.pgen.1000749

Editor: Christopher E. Pearson, The Hospital for Sick Children and University of Toronto, Canada

Received May 19, 2009; **Accepted** November 2, 2009; **Published** December 4, 2009

Funding: This work was supported by funds from CNRS and INSERM, by the Hereditary Disease Foundation (USA), by the NIH (National Institute of Neurological Disease and Stroke) grant NS049200, and by the Intramural Research Program of the NIH, National Institute on Aging. The funders have not been implicated in the design, realization, and analysis of the study. They were not involved in manuscript preparation or the decision to submit.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: merienne@igbmc.fr (KM); yvon@igbmc.fr (YT)

Introduction

Huntington's disease (HD) is a neurodegenerative disorder caused by aberrant expansion of a CAG repeat tract within the coding sequence of the Huntingtin (*HTT*) gene, resulting in the production of a mutant protein with a toxic elongated polyglutamine (polyQ) stretch. This dominantly inherited disease, which shares the same mutation mechanism with eight other neurodegenerative disorders, is characterized by preferential and progressive degeneration of the medium-spiny neurons in the striatum. The HD mutation is unstable in germline and somatic cells, and CAG repeat expansion in both cell types has deleterious clinical consequences. Transmission of the mutation to offspring, in particular by fathers, is characterized by an expansion bias, leading to the phenomenon of anticipation [1–3], whereby the disease tends to worsen over successive generations due to the production of mutant huntingtin protein with an increased glutamine tract length that triggers earlier disease onset and increased disease severity. In addition, somatic CAG repeat expansion occurs in several tissues, including the brain. Interestingly, different brain regions are unequally affected [2,4]: somatic instability is extensive in striatal neurons but very limited in cerebellar neurons, which are largely spared by the disease [5,6]. It is therefore proposed that somatic expansion in the striatum and other target tissues, leading to the production of increasingly toxic mutant huntingtin proteins, accelerates HD pathology and acts as a disease-modifier [5–7]. Consistent with this hypothesis, an early mutant huntingtin phenotype is delayed in HD mice in which somatic instability is prevented by a deficiency in either the Msh2 or Msh3 mismatch repair (MMR) protein [8,9].

Several studies support the notion that DNA repair contributes to somatic instability of CAG repeats, though other DNA-associated processes, such as replication, may also play a role [10]. In particular, at least two repair pathways appear to regulate

Author Summary

Huntington's disease (HD) is a neurodegenerative disorder that belongs to a family of genetic diseases caused by abnormal expansion of CAG/CTG repetitive sequences. The instability of trinucleotide repeat expansions in germline and somatic cells has deleterious clinical consequences in HD. For instance, transmission of longer repeats to offspring results in an earlier onset of disease, where extensive somatic expansion in the striatum, the brain region primarily affected in HD, is proposed to accelerate disease pathology. Thus, understanding the mechanisms of trinucleotide repeat instability is a major interest. We have examined the role of oxidative DNA damage and base excision repair (BER) in somatic instability, which is tissue-selective and age-dependent. We show that oxidative DNA lesions abnormally accumulate at CAG expansions in a length-dependent, yet age- and tissue-independent manners, likely due to the secondary structures formed by CAG repeats that limit access of enzymes initiating BER. In addition, our data indicate that repair by BER enzymes of some of the accessible lesions results in somatic expansion when the ratio of FEN1 to POLβ is low, as found to occur in the striatum. Our results support BER enzyme stoichiometry as a contributor to the tissue selectivity of somatic CAG expansion in HD.

in vivo somatic instability in the brain of HD mouse models. In addition to MMR, whose role is well documented [8,9,11,12], base excision repair (BER), which is specialized in DNA base damage removal, has recently been implicated. Indeed, somatic CAG repeat instability was reduced in HD mice lacking the DNA glycosylase OGG1, an enzyme that initiates BER of 8-oxoguanine (8-oxoG) lesions [13]. Furthermore, the global level of oxidative DNA damage increased with age in the brains of HD mice, while OGG1 activity did not change, suggesting that the level of DNA oxidation at CAG expansions might underlie the age dependence of somatic instability [13]. *In vitro* biochemical and yeast studies have further suggested that other BER enzymes might modulate somatic instability. DNA polymerases, including DNA polymerase beta (POLβ), a central participant in most BER responses, can promote CAG repeat extensions *in vitro* [13]. In contrast, Flap Endonuclease-1 (FEN1), involved in the long-patch BER (LP-BER) subpathway, prevents CAG repeat expansion in yeast [14].

BER is characterized by a sequence of highly coordinated steps [15,16], starting with damaged base recognition and removal by a DNA glycosylase, leading to formation of an apurinic/apyrimidinic (AP) site. In the brain, OGG1 appears to be the major glycosylase involved in removing endogenous lesions such as 8-oxoG from the genome [17]. Cleavage at the AP site is performed by the major AP endonuclease, APE1, which incises 5′ to the damage leaving behind a 3′OH and a 5′deoxyribose phosphate (5′dRP), which is subsequently removed by the dRP lyase activity of POLβ. A DNA polymerase then fills the resulting gap. In single-nucleotide BER (SN-BER), POLβ incorporates the missing nucleotide. However, when the 5′-terminal residue is refractory to POLβ lyase activity, repair proceeds via LP-BER. In this case, multiple nucleotides are incorporated by one of several DNA polymerases, including POLβ, POLδ or POLε, through a strand displacement mechanism. Subsequently, FEN1 is required to remove the 5′-flap structure formed during LP-BER synthesis. Interestingly, a number of studies suggest that BER enzyme activities vary with age and between tissues, indicating that the repair kinetics of oxidative DNA damage may be age- and tissue-dependent [18].

We sought to clarify the role of DNA damage and BER in the tissue specificity and age dependence of somatic CAG expansion in HD. First, we asked whether DNA lesions accumulated at CAG repeats, particularly in the striatum. Second, we asked whether BER proteins and activities displayed a specific or unique pattern of expression in the striatum. We have systematically analyzed DNA damage at CAG repeats, BER protein levels, and BER activities in both the unstable striatum, and in the mostly stable cerebellum of young and aged HD mice. Our results show that the stoichiometry of BER enzymes, rather than DNA damage levels, correlates with the tissue selectivity of somatic CAG expansion.

Results

Normal aging contributes to somatic CAG instability in R6 HD mice

Several studies have suggested that HD pathology contributes to neuronal aging. Specifically, oxidative stress and consequently oxidative DNA damage, which accumulates during normal aging, are elevated in HD neurons [19]. Thus, the "aging process" for HD neurons can potentially be envisioned as two separate components: i) normal and ii) disease-associated. We set out to determine the role of these two components in the process of somatic expansion. To this end, we compared various R6 transgenic mouse models of HD, which recapitulate several features of the human pathology, including somatic instability [6,20]. R6/1 mice express the first exon of the human *HTT* gene with ~125 CAG repeats and develop a progressive disease leading to death within approximately 8 months. The same transgene is expressed in R6/2 mice, yet is inserted in an alternate genomic location. R6/2–160 CAG and R6/2–100 CAG (spontaneously derived from R6/2–160 CAG) mice contain 160 and 100 CAG repeats, respectively. In both R6/2 lines, disease onset and progression is much more severe than in the R6/1 mice. Indeed, R6/2–160 CAG and R6/2–100 CAG mouse lines die at 3 and 4.5 months, respectively [21].

Somatic cell CAG instability was determined in R6/1 mice of various ages using standard PCR amplification and Genescan analysis from two brain regions that exhibit contrasting degrees of CAG instability, *i.e.* the striatum and the cerebellum (Figure 1A, top panel). To obtain a baseline reference for the number of CAG repeats, we analyzed tails from mice at 3 weeks of age, because the CAG repeat region from this tissue is mostly stable when taken from very young animals. As expected, quantification of instability showed that both the median and the amplitude of the Genescan profiles increased upon aging in the R6/1 striatum, but to a much lesser extent in the R6/1 cerebellum (Figure 1B, upper panels). As previously described, the profiles generated from aging striata, which display extensive somatic instability, were bimodal [6].

Comparison of 13 week-old R6/1 mice with R6/2–160 CAG and R6/2–100 CAG mice of similar ages, 12 and 16 weeks, respectively, showed that CAG repeat expansion was significantly higher in the striatum of R6/2 mice (Figure 1A and 1B, lower left panel). Thus, the process of somatic CAG instability was quicker in R6/2 mice than in R6/1 mice, regardless of the repeat length in the R6/2 mice, suggesting that disease-associated mechanisms and/or positional effects are involved in somatic expansion. However, when R6/1 and R6/2–160 CAG were compared at similar pathological stages (*e.g.* end-stages), we found that somatic instability was significantly higher in the striatum and the cerebellum of the R6/1 mice (Figure 1A and 1B, lower right panel). In agreement with previous results [6], these data indicate that normal aging contributes to somatic CAG instability.

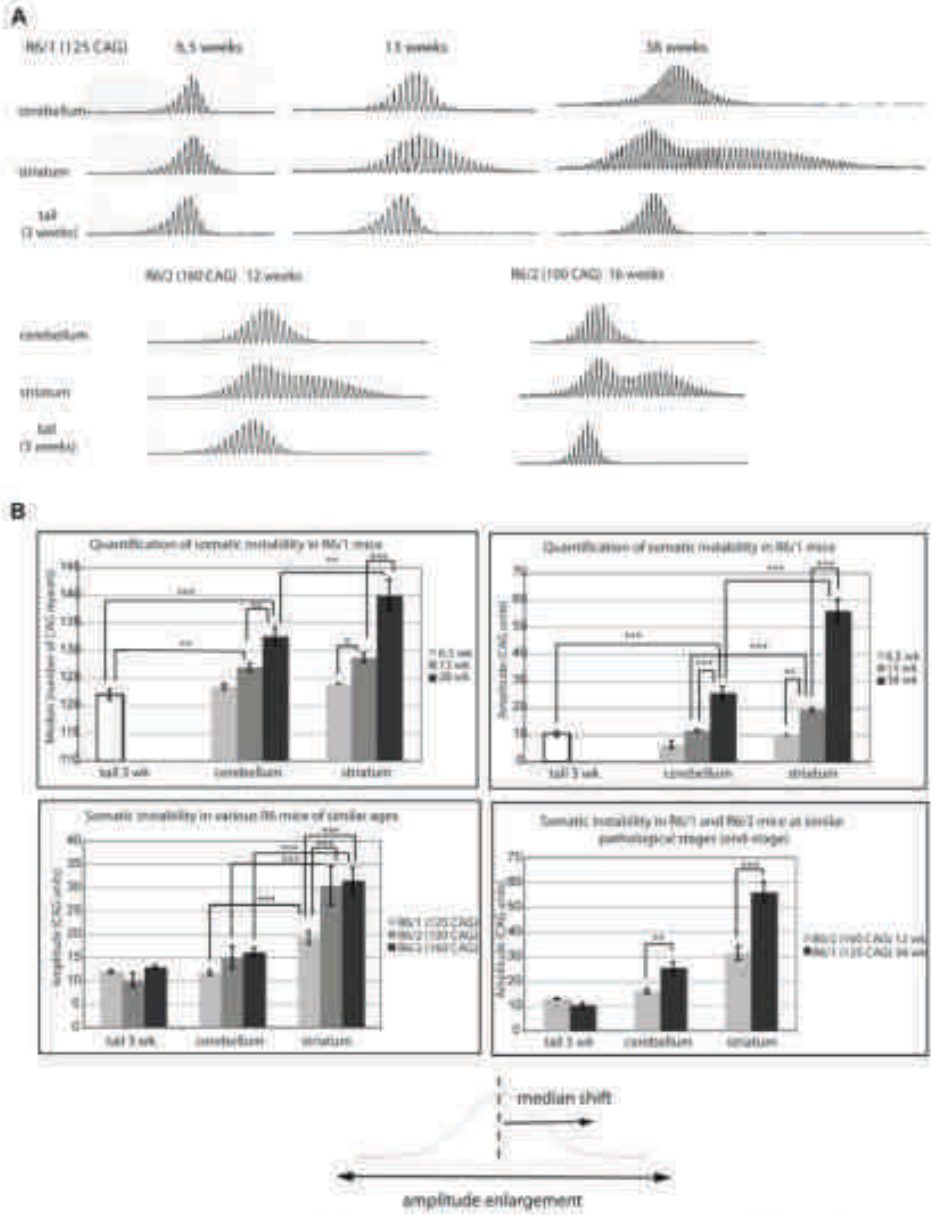


Figure 1. Comparison of various R6 mouse lines shows that natural aging contributes to somatic CAG expansion in HD mouse models. (A) Representative Genescan profiles showing CAG repeat size distribution from genomic DNA prepared from striatum or cerebellum of R6/1, R6/2 (160 CAG) and R6/2 (100 CAG) mice of various ages. The same transgene—except for CAG repeat length—is inserted in the different R6 mouse lines. One transgene copy is integrated into both the R6/1 and R6/2 lines, although integration sites are different [21]. For each mouse analyzed, the Genescan profile of the tail was also determined at 3 weeks to calculate the number of CAG repeats initially transmitted (corresponding to the peak with the highest fluorescent intensity). (B) Quantification of somatic CAG expansion in R6/1, R6/2 (160 CAG) and R6/2 (100 CAG), using either the median or the amplitude of the profiles as quantitative parameters. Comparison of mean medians (top left panel) and mean amplitudes (top right panel) between the striatum and the cerebellum of R6/1 mice with aging. Middle left panel. Comparison of mean amplitudes in striatum and cerebellum between R6/1, R6/2 (100 CAG) and R6/2 (160 CAG) mice at similar ages (12–14 weeks). Middle right panel. Comparison of mean amplitudes in striatum and cerebellum of R6/1 and R6/2 (160 CAG) at similar pathological stages, *i.e.* 38 weeks and 12 weeks, respectively. Mean values were calculated from 4 to 8 independent profiles (obtained from different mice). Error bars, sem; *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$ (ANOVA followed by Newman Keuls test for post-hoc comparisons). Bottom. Somatic CAG instability leads to both amplitude enlargement of repeat size distribution and shift of the median towards longer alleles.
doi:10.1371/journal.pgen.1000749.g001

Global DNA damage does not correlate with tissue-specific propensity of somatic CAG instability

It was recently proposed that DNA oxidation contributes to the age dependence of somatic CAG instability, due to damage accumulation in the brain during aging [13]. To clarify the role of DNA damage (and aging) in somatic expansion, we quantified AP sites within the whole genome in the striatum and cerebellum of R6/1 and R6/2 mice. AP sites are formed either spontaneously or as intermediates during repair of oxidized, deaminated or alkylated bases. AP sites increased with age in the striatum of R6/1 mice, with greater abundance at 8 months than at 3 months (Figure 2A). AP sites were also significantly higher at 8 months in the R6/1 striatum than in the striatum of littermate control animals, potentially reflecting increased oxidative stress or reduced repair in the HD mouse striatum. In contrast, the levels of AP sites detected in the cerebella of R6/1 were similar to those measured from control mice at 8 months (Figure 2A). To our surprise, we found that AP sites were significantly higher in the cerebellum *vs* striatum in age-matched animals for both the R6/1 and R6/2 lines (Figure 2A), indicating the level of DNA damage does not correlate with the tissue selectivity of somatic CAG instability in HD mice.

Accumulation of DNA damage at CAG expansions in HD mice is neither tissue- nor age-dependent

We then asked whether DNA damage specifically accumulates at CAG expansions in the striatum of aging mice, as proposed by Kovtun *et al.* [13]. We evaluated the level of DNA oxidation at CAG expansions based on digestion of genomic DNA by the bacterial DNA glycosylase Fpg followed by real-time quantitative PCR (QPCR) as previously described [22]. The efficacy of PCR amplification was predicted to be reduced if the amplified DNA region contained Fpg sensitive sites, which includes several oxidative base lesions (*i.e.* 8-oxoG, 8-oxoadenine, fapy-guanine, methy-fapy-guanine, fapy-adenine, aflatoxin B_1-fapy-guanine, 5-hydroxy-cytosine, 5-hydroxy-uracil, as well as AP sites, all of which are converted to strand breaks by the enzyme. As a control, we amplified a portion of the orthologous *HTT* gene in mouse (*Hdh*), which encompasses seven interrupted CAG repeats. We chose primers to amplify PCR fragments from the CAG-expanded transgene and the control *Hdh* locus that were equivalent in size and GC content. The sizes of the CAG-expanded and *Hdh* PCR fragments are 440 bp (corresponding to 120 CAG repeats) and 420 bp, respectively, with the GC content reaching 67.2% and 72.2% (Figure S1).

Given that the relative PCR efficiency was >0.8, the number of Fpg sensitive sites at the murine *Hdh* control locus appeared to be negligible in the striata and cerebella of both wild-type and R6/1 mice at 6 and 35 weeks of age, suggesting that spurious oxidation did not significantly affect the assay (Figure 2B, upper panel). In contrast, Fpg sensitive sites accumulated significantly at CAG expansions in the R6/1 mice, as revealed by a reduced relative PCR efficiency of ~0.6 (Figure 2B, lower panel). Surprisingly, damage accumulation at CAG expansions was neither age- nor tissue-dependent, as the levels were similar in the striatum and cerebellum of young (6 weeks) and old (35 weeks) mice (Figure 2B).

We then measured Fpg sensitive sites in the striatum of mice with different CAG repeat lengths. For this purpose, we analyzed HD knockin mice heterozygous for the HD mutation and bearing either 18, 40 or 115 CAG repeats [23]. Accumulation of Fpg sensitive sites was CAG repeat length-dependent, as only mice with 115 CAG repeats displayed significantly reduced PCR amplification (Figure 2C). Thus, DNA damage at CAG expansions is not sufficient to trigger somatic instability, as accumulation of Fpg sensitive sites did not correlate with the age dependence or tissue selectivity seen in HD. The repeat length-dependence of DNA damage accumulation suggests that a build-up of presumably oxidative DNA lesions at CAG repeats might be dictated by the propensity to form alternate structures in DNA.

DNA lesions at hairpin structures are inaccessible to OGG1 and APE1

DNA sequences containing CAG repeats spontaneously form secondary structures such as hairpins [24,25]. To assess the role of DNA structure in oxidative DNA damage accumulation at CAG expansions, we synthesized substrates mimicking hairpin structures formed by CAG sequences and containing either an oxidized base (8-oxoG) or an AP site analog (tetrahydrofuran, THF) at the tip of the hairpin. Two series of hairpin oligonucleotides were generated containing an 8-oxoG or THF modification at (1) the 3rd position within a 3-nucleotide loop (referred to as CAG1oxoG and CAG1THF) or (2) the 2nd position within a 4-nucleotide loop (referred to as CAG2oxoG and CAG2THF) (Figure 3A). Design of hairpin oligonucleotides included sites for *Eco*RI and *Bam*HI restriction enzymes (Figure 3A), which allowed for verification of hairpin structure formation. Digestion with either restriction enzyme revealed that the oligonucleotides formed the expected stable hairpin structure (Figure 3B). As additional controls, we designed linear oligonucleotides containing either an 8-oxoG lesion (referred to as 34-oxoG) or an AP site analog (referred to as 34THF). In contrast to OGG1, APE1 can incise single stranded oligonucleotides with varying efficiency, and thus, 34THF was used directly as a control [26,27]. 34-oxoG, on the other hand, was annealed to a complementary oligonucleotide (34C) to form a control double stranded DNA substrate (34-oxoG/34C). To test whether OGG1 or APE1 were able to incise damage in hairpin substrates, increasing concentrations of human OGG1 or APE1 were incubated with the appropriate 5′-^{32}P labeled oligonucleotide substrates. As shown in Figure 3C, no incision products were detected when using the aforementioned CAG hairpin substrates, even when OGG1 or APE1 concentrations were high (*i.e.* <2.5% product at 30 ng OGG1 or <3.5% product at 1 ng APE1). Conversely, both enzymes, even at low concentrations, efficiently incised their respective control substrates. Thus, OGG1 and APE1 are unable to efficiently incise hairpin substrates *in vitro*, supporting the idea that oxidative DNA lesions accumulate at CAG repeats because they can form secondary structures that are refractory to processing by BER enzymes.

BER stoichiometry is different in the striatum and cerebellum of HD mice

Oxidative DNA damage is repaired by the BER pathway, which appears to be regulated in an age-dependent and tissue-specific manner [18]. We therefore asked whether the activity or stoichiometry of BER enzymes could contribute to the age dependence or tissue selectivity of somatic CAG instability in HD. We first determined mRNA expression levels of several major BER genes in the striatum and cerebellum of R6/1 HD and control mice at 40 weeks of age using quantitative RT-PCR. The BER genes that we tested were *Ogg1*, *Ape1*, *Fen1* and *Polβ*. Expression of most of these genes was similar between the transgenic and control mice, both in the striatum and cerebellum, except for *Ogg1*, which was downregulated 2-fold in the cerebellum of R6/1 mice (Figure 4A).

We next analyzed protein levels of several major BER enzymes by Western blotting (Figure 4B). The results showed that FEN1 was more abundant in the cerebellum than in the striatum of 40 week-old R6/1 mice. FEN1 was also produced at higher levels in R6/1 mice compared to control animals. APE1, POLβ and

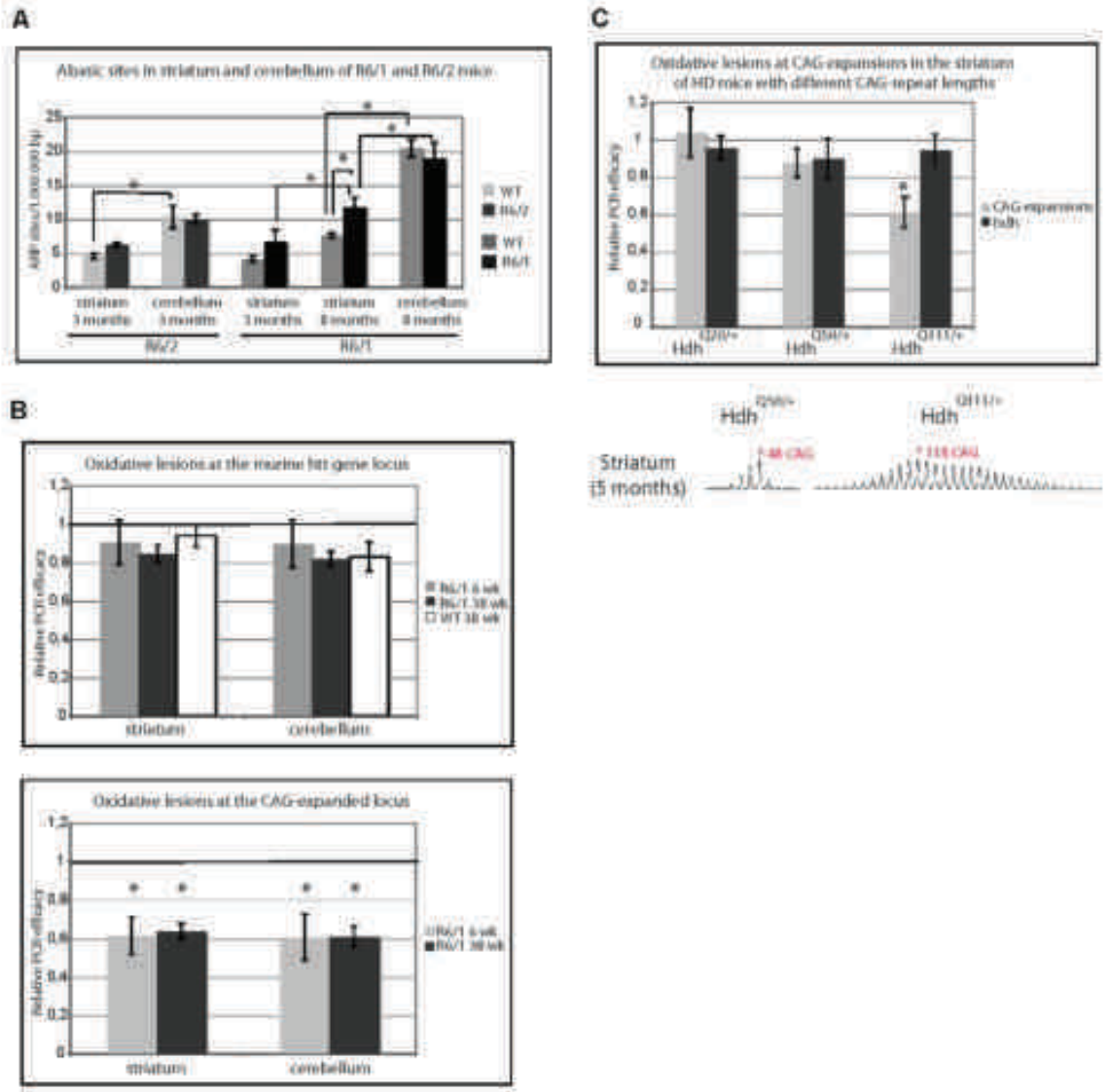


Figure 2. DNA damage accumulates at CAG expansions. (A) Quantification of AP sites from R6/1 or R6/2 mice, as well as their control littermates. AP sites were measured from striatum and cerebellum of 3 month-old R6/2 mice (at end-stage pathology) and from striatum and cerebellum of 3- or 8-month-old R6/1 mice (early- and end-stage pathology, respectively). 3 to 4 mice of each genotype and age were used. Error bars, sem; *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$ (ANOVA followed by Newman Keuls test for post-hoc comparisons). (B) Oxidative DNA damage level at the murine *HTT* gene locus (*Hdh*) (top panel) and at the transgene locus (CAG-expanded locus) (bottom panel) in striatum and cerebellum of R6/1 mice at 6 or 38 weeks. Oxidative DNA damage levels at the murine *Hdh* gene locus were also determined from wild type littermate controls (WT) at 38 weeks. Ratios plotted in the graphs correspond to the Q-PCR–deduced DNA relative concentration of samples after Fpg treatment divided by the Q-PCR relative concentration of untreated samples. A ratio of 1 means that the specific gene locus tested did not accumulate oxidative DNA damage, while a ratio below 1 indicates that oxidative DNA damage accumulated. 3 to 4 mice of each genotype and age were analyzed. Error bars, sem; *, $p<0.01$ (Student's t-test). (C) Top panel. Oxidative DNA damage level at the CAG-expanded and *Hdh* loci in striatum of heterozygous knockin (KI) HD mice at 3 months of age with either 18, 48 or 115 CAG repeats. 3 mice of each genotype were analyzed. *, $p<0.01$ (Student's t-test). Bottom panel. Representative Genescan profiles of CAG repeat size distribution in the striatum of $Hdh^{Q50/+}$ or $Hdh^{Q111/+}$ mice.
doi:10.1371/journal.pgen.1000749.g002

OGG1 levels were not significantly changed when comparing striatum *vs* cerebellum or R6/1 *vs* control animals. However, OGG1 tended to be lower in the cerebellum of R6/1 mice relative to the striatum, whereas APE1 appeared to be higher in the cerebellum of R6/1 mice. As a consequence, the stoichiometry of BER proteins was different between the cerebellum and striatum of R6/1 mice (Figure S3A). Using recombinant FEN1 and POLβ proteins as standards and Western blotting, we more precisely

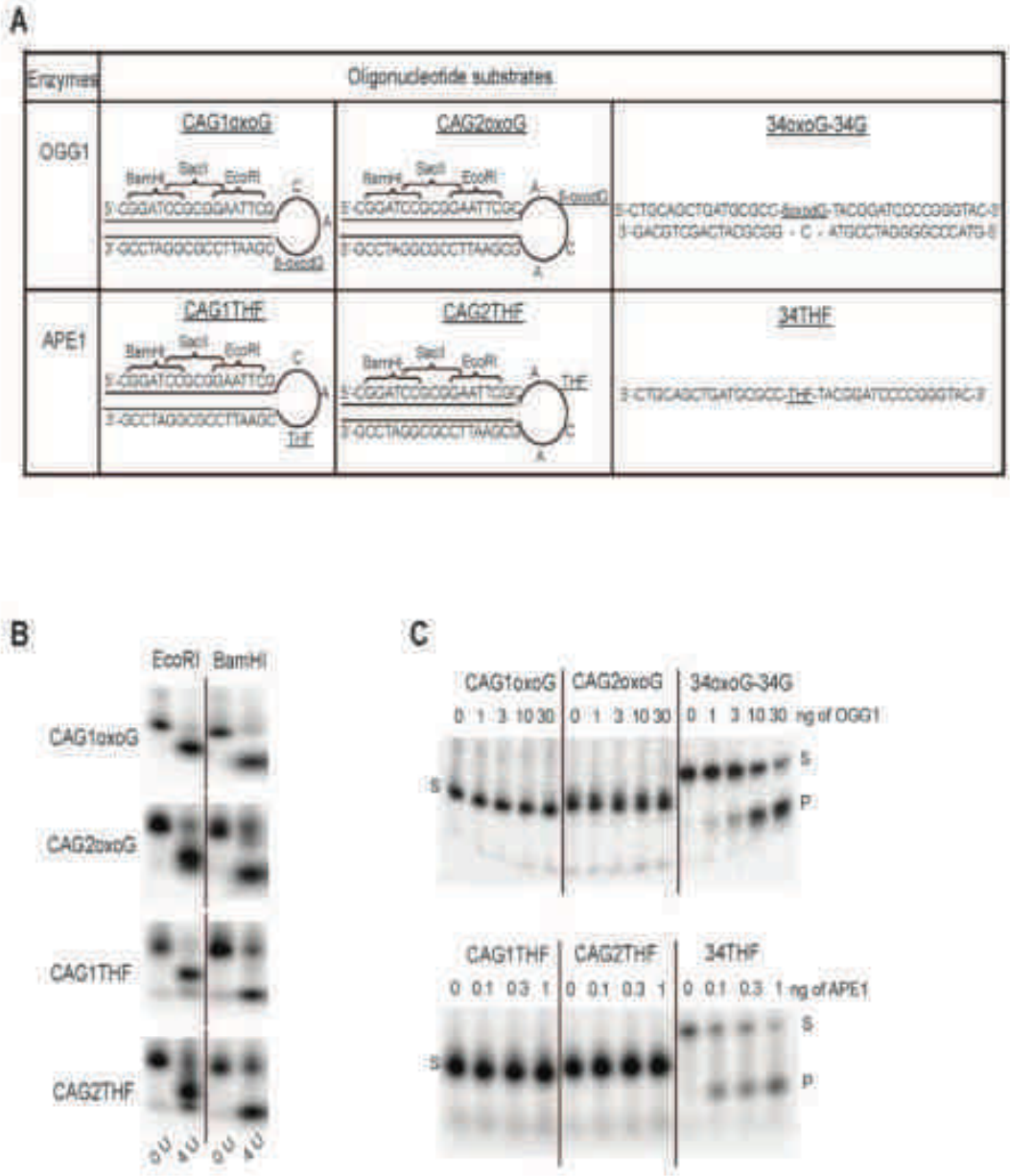


Figure 3. OGG1 and APE1 are unable to incise lesions at hairpin DNA structures. (A) Design of hairpin and control oligonucleotides. Hairpin oligonucleotides contain a 8-oxoG lesion or a THF lesion in 3-nt (CAG1oxoG or CAG1THF) or 4-nt (CAG2oxoG or CAG2THF) loop. The hairpin structured oligonucleotides contained *BamHI*, *SacII* and *EcoRI* restriction sites on the linear double stranded portion of the hairpin. (B) Digestion of hairpin oligonucleotides with *EcoR1* or *BamH1* (4 U). Efficient digestion shows that the designed oligonucleotides adopt a stable hairpin structure. (C) Hairpin and control oligonucleotides were incubated with increasing amounts of OGG1 (top panel) or APE1 (bottom panel). S: substrate, P: product. Recombinant enzymes efficiently incised control, but not hairpin, oligonucleotides.
doi:10.1371/journal.pgen.1000749.g003

determined the stoichiometric ratio between FEN1 and POLβ in the striatum and cerebellum of R6/1 mice (Figure S3B). The results revealed that the FEN1/POLβ molar ratio was 1.4+/−0.3 in the striatum and 6+/−1.2 in the cerebellum corresponding to a 4-fold difference, thus supporting that the stoichiometry of BER enzymes is different in the striatum vs cerebellum of R6/1 mice, largely due to a significant disparity in FEN1 protein levels. We also analyzed the high-mobility group box 1 (HMGB1) protein, which was recently identified as a cofactor of BER that stimulates the activities of APE1 and FEN1 [28]. Interestingly, we found that

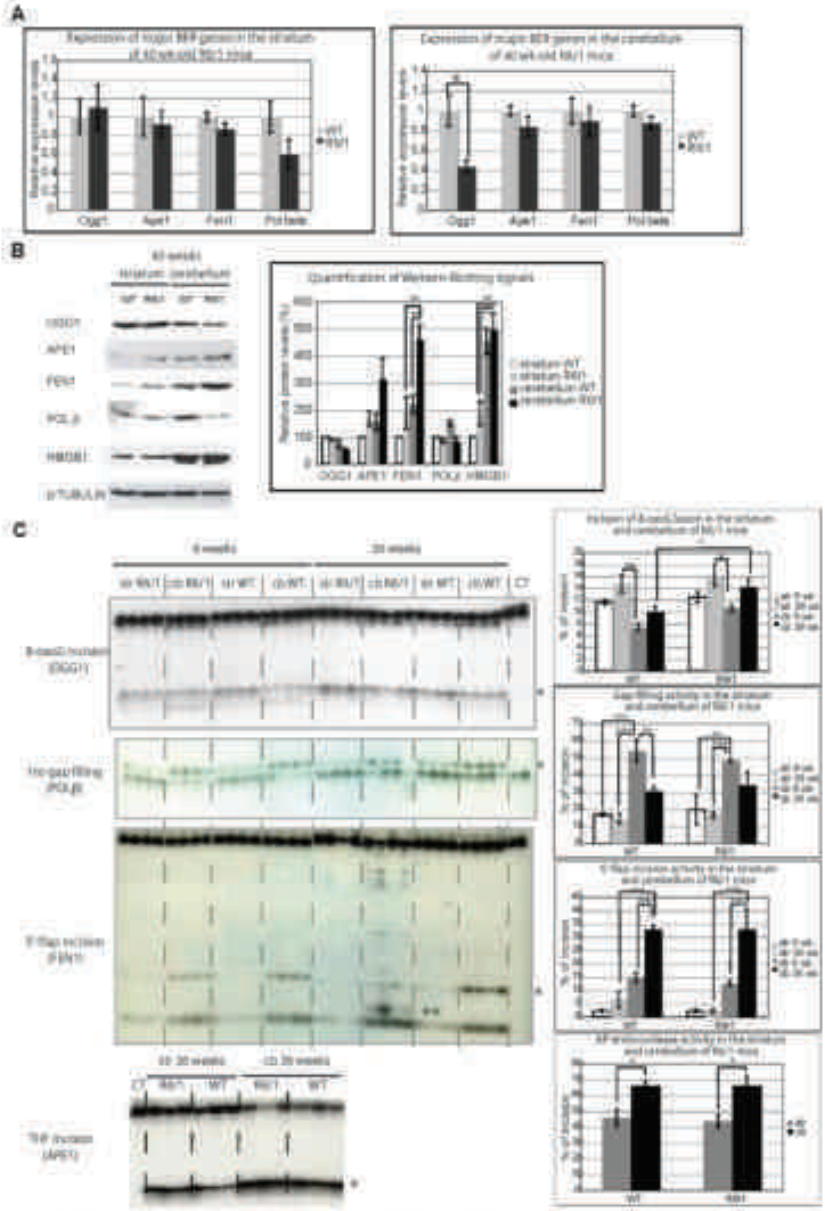


Figure 4. Analysis of BER mRNA, protein, and activity levels in striatum and cerebellum of R6/1 mice. (A) Real-time quantitative RT–PCR analysis of total RNA extracted from striatum (left panel) and cerebellum (right panel) of R6/1 mice and littermate controls (WT) at 40 weeks of age using primers specific for *Ogg1*, *Ape1*, *Fen1*, or *Pol*β. Error bars, sem. *, $p<0.05$ (Student's *t*-test). 3 to 4 mice of each genotype were used. (B) Left panel. Western blotting analysis of whole cell protein extracts from 40 week-old R6/1 or WT striata or cerebella. Cell extracts were prepared by pooling tissues from 4 animals. Right panel. Quantification of western blotting data. **, $p<0.01$ (ANOVA followed by Newman Keuls test for post-hoc comparisons). (C) Left panels. Activity analyses using a specific assay for each tested enzymatic activity (see Materials and Methods). Specific labeled substrates were incubated with striatal or cerebellar nuclear extracts prepared from R6/1 mice and control littermates, at both 6 and 36 weeks of age. *, normal incision or incorporation products; **, an additional incision product observed in FEN1 assay, when substrate was incubated with extracts from 36-week-old R6/1 cerebella, and included in quantification of the assay (see below). Right panels. Quantification of tested enzymatic activities in the striatal and cerebellar mouse extracts. The percentages of incision (likely OGG1, APE1, and FEN1) or incorporation (DNA polymerase), calculated by dividing the quantity of incised or elongated substrate by the total quantity of substrate, are represented on the graphs. *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$ (ANOVA followed by Newman Keuls test for post-hoc comparisons).
doi:10.1371/journal.pgen.1000749.g004

HMGB1 protein was 2–3 fold higher in the cerebellum than in the striatum of R6/1 animals, suggesting that APE1 and FEN1 activities may be further activated by HMGB1 in the cerebellum of R6/1 mice. Finally, the level of BER enzymes and HMGB1 were similar in the cortex, another brain region characterized by substantial somatic expansion, to those in the striatum of R6/1 mice (data not shown), indicating that the stoichiometry of BER proteins might modulate the propensity of a given tissue for somatic CAG expansion.

BER activities are less efficient in the striatum than in the cerebellum of HD mice

To evaluate further the role of BER in the tissue selectivity or age dependence of somatic CAG instability, we compared the enzymatic activities of the major steps of BER in extracts prepared from the striatum and cerebellum of R6/1 mice at both early- (6 weeks) and late- (36 weeks) stages. In addition, BER activities were compared between R6/1 and control littermate animals. Using standard oligonucleotide substrates (Figure S2), we performed *in vitro* biochemical assays allowing for evaluation of 8-oxoG lesion removal, gap filling, 5′-flap excision and AP endonuclease activities (Figure 4C). OGG1, POLβ, FEN1 and APE1 are most likely the main enzymes responsible for these activities, respectively, but other enzymes might contribute to the various activities as well. Strikingly, we did not observe significant differences between R6/1 mice and control animals regardless of the activity examined. Our results revealed that 8-oxoG incision tended to increase with age, both in the striatum and cerebellum of R6/1 and control mice, and to be higher in the striatum than in the cerebellum. In young animals (both R6/1 and controls), gap filling activity was significantly higher in the cerebellum compared to the striatum (by 2- to 3-fold). This difference disappeared at 36 weeks, due to an overall decrease in gap filling activity in the cerebellum (by 1.5- to 2-fold). AP endonuclease activity was slightly elevated in the cerebellum of both 36 week-old R6/1 and control mice (1.5 fold higher) when compared to the striatum. Most remarkably, 5′-flap incision activity was increased 5- to 10-fold in the cerebellum when compared to the striatum for both control and R6/1 mice. Furthermore, 5′-flap incision dramatically increased in the cerebellum upon aging, while it remained very low in the striatum. Finally, additional incision products were detected only in the cerebellum of 36 week old R6/1 mice, indicating that other exo- or endo-nuclease activities may be able to remove the 5′-flap structure in the cerebellum of aging transgenic animals.

Our results indicate that key steps implicated in BER, specifically downstream of 8-oxoG incision (OGG1) activity, are less efficient in the striatum than in the cerebellum of R6/1 mice. They also show that aging subtly modifies the stoichiometry of BER enzymes in a tissue-specific manner. Indeed, upon aging, 8-oxoG incision activity tends to increase both in the striatum and cerebellum, gap-filling activity decreases in the cerebellum and remains stable in the striatum, and flap endonuclease activity dramatically increases in the cerebellum, but remains stable, and low, in the striatum. These findings suggest that the LP-BER subpathway, which requires FEN1, is much less efficient in the striatum than in the cerebellum of R6/1 mice.

Enrichment of POLβ at CAG expansions is tissue-specific

To further interrogate the role of BER in somatic CAG instability *in vivo*, we performed chromatin-immunoprecipitation (ChIP) experiments. To set up conditions, we first used an antibody against acetylated K9/14 histone H3 (AcH3K9/14) that immunoprecipitates transcribed regions. ChIPed DNA was amplified using two primer sets encompassing the *CAG*-expanded transgenic region and a comparable portion of the mouse *Hd* gene as a control. The results show that AcH3K9/14 is specifically and highly enriched at CAG expansions and at the *Hd* locus in both the striatum and the cerebellum of R6/1 and R6/2 HD mice (Figure 5A and 5B, lower panels). This study indicates that our

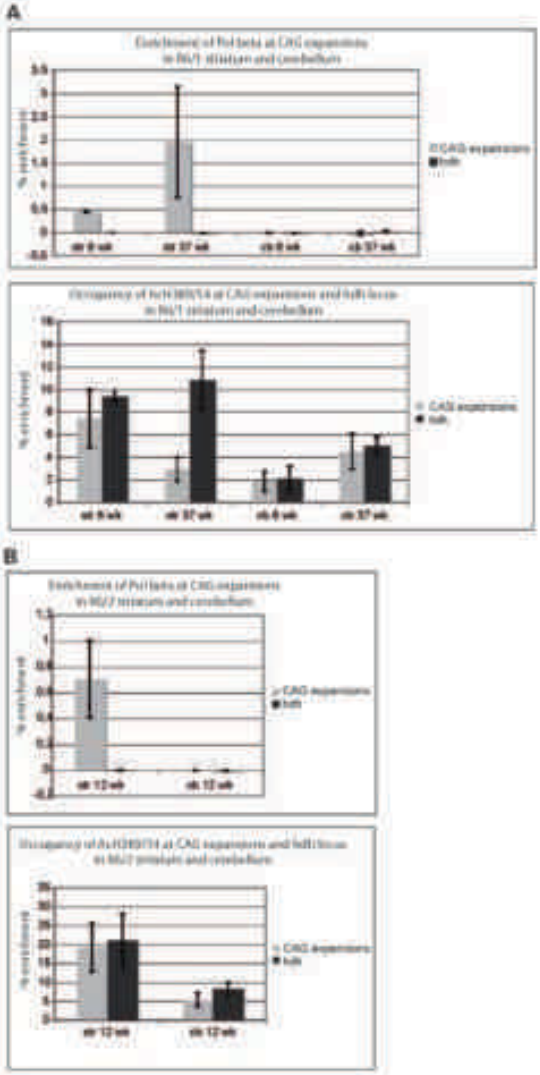


Figure 3. POLβ is specifically enriched at CAG expansions in the striatum but not in the cerebellum of HD mice. ChIP of *CAG*-expanded and *Hdh* loci from striatum and cerebellum of R6/1 (A) and R6/2 (B) mice using α-POLβ antibody (upper panels) and as control, α-AcH3K9/14 antibody (lower panels). (A) R6/1 mice at both 6 and 37 weeks of age and R6/2 mice at 12 weeks of age were analyzed. Each ChIP experiment was performed by pooling striata and cerebella from 2 to 4 mice. The values plotted on the graphs represent the mean values obtained from 3 to 4 independent experiments. The values correspond to percentage of enrichment, calculated as follows: % enrichment = (relative DNA concentration after ChIP with POLβ or AcH3—relative DNA concentration after ChIP with no antibody/Input. Relative DNA concentrations were measured using quantitative PCR. Error bars are sem.
doi:10.1371/journal.pgen.1000749.g003

experimental conditions allowed for significant and specific detection of protein bound to both CAG expansions and the *Hdh* gene. We then employed several antibodies against BER proteins. α-OGG1 and α-APE1 antibodies did not reveal any specific enrichment at either CAG expansions or at the *Hdh* locus (data not shown). However, we did find that POLβ was specifically enriched at CAG expansions in the striatum, but not in the cerebellum of R6/1 mice. Interestingly, this enrichment tended to increase with age (Figure 5A). POLβ was also specifically enriched at CAG expansions in the striatum of R6/2 mice (Figure 5B). Thus, enrichment of POLβ at CAG expansions was tissue-specific. These *in vivo* data support the concept that POLβ promotes somatic expansion [13].

POLβ multi-nucleotide gap filling synthesis is modulated by the stoichiometry of POLβ and FEN1 during DNA damage repair of CAG repeat substrates

Having shown that the FEN1/POLβ molar ratio is 1.4+/−0.3 in the striatum and 6+/−1.2 in the cerebellum of R6/1 mice (Figure S3B), that FEN1-like nuclease activity is 5- to 10-fold more elevated in the cerebellum (Figure 4C), and that POLβ is specifically enriched at CAG expansions in the striatum of HD mice (Figure 5), we suspected that the FEN1/POLβ ratio might be critical in determining the propensity of a given tissue to experience somatic CAG instability. To explore this possibility, we examined the effects of varying FEN1:POLβ stoichiometry on POLβ DNA synthesis using a 100-mer substrate consisting of a 23-mer primer followed by a single nucleotide gap containing a 5′-abasic site analog (THF) and nineteen CAG repeats (Figure 6A). The effect of FEN1, at different molar ratios, on POLβ DNA synthesis activity was evaluated as follows. We employed purified POLβ at a single concentration, determined to give predominantly single nucleotide incorporation, and varied FEN1 to give 1:2, 1:1, 2:1, 4:1 and 8:1 FEN1:POLβ stoichiometries. We found that enhancement of POLβ multiple nucleotide incorporation, presumably due to strand displacement synthesis, was achieved at FEN1:POLβ ratios of 1:1 to 2:1 and, to a lesser extent, 4:1. Interestingly, at the 8:1 ratio, POLβ multiple nucleotide synthesis was inhibited (Figure 6B). Under identical conditions, we observed no significant change in FEN1 flap endonuclease activity, although we did observe increasing 5′-flap endonuclease activity with increasing FEN1 protein (data not shown). These results support a model whereby BER enzyme stoichiometry affects the outcome of DNA damage processing, and as a consequence, potentially influences trinucleotide repeat instability.

Discussion

We have shown that oxidative DNA damage accumulates at CAG expansions in a CAG-repeat length-dependent, but age- and tissue-independent manner. Our data indicate that the DNA structure of CAG repeats likely contributes to this accumulation, as BER enzymes such as OGG1 and APE1 were unable to process lesions within hairpin configurations. In addition, comparing the striatum and cerebellum of HD mice, we have discovered a correlation between the propensity for somatic CAG expansion and BER protein stoichiometry and enzymatic activities. In particular, somatic instability correlates mainly with low flap endonuclease activity, as seen in the striatum. Our results also suggest that POLβ contributes to *in vivo* somatic expansion, as it is specifically enriched at CAG repeats in the striatum of aging animals. We thus propose a model whereby BER contributes to somatic CAG instability (Figure 7). In this model, accumulation of DNA damage at CAG repeats results from low accessibility of BER enzymes to lesions in pre-existing secondary DNA structures formed within CAG repeats. Repair of accessible lesions, which would arise due to the dynamic nature of trinucleotide repeat sequences, would be initiated by a protein such as OGG1 as proposed by Kovtun *et al.* [13] and potentially lead to somatic CAG expansion. This outcome would be particularly likely where there is poor cooperation between the strand displacement activity of POLβ and the 5′-flap excision activity of FEN1 during LP-BER.

The models proposed to explain disease-associated trinucleotide repeat instability involve specific stable secondary DNA structures. *In vitro* approaches based on substrates containing CAG/CTG

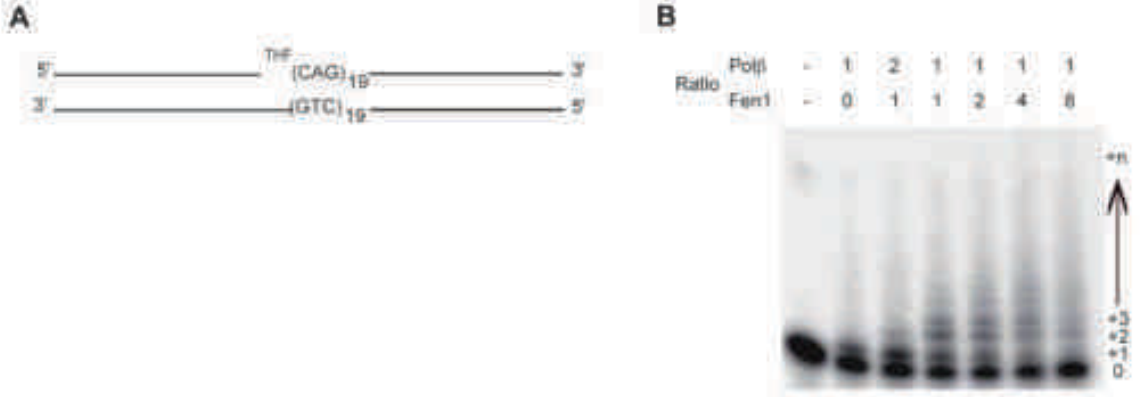


Figure 6. FEN1:POLβ stoichiometry modulates multi-nucleotide gap filling synthesis of POLβ during DNA damage repair of CAG repeat substrates. (A) CAG repeat substrate used for POLβ DNA synthesis and FEN1 flap endonuclease reactions. The substrate consists of a 23mer DNA primer strand, a gap, and a 70mer containing a 5′ tetrahydrofuran molecule (THF), 19 CAG repeats and 19 additional nucleotides (Top strand), both were annealed to a 100mer of complementary sequence. (B) Differential stoichiometric ratios of FEN1 modulate POLβ dependent synthesis of CAG repeat tracts. Reactions contained 2 nM POLβ and 1, 2, 4, 8, or 16 nM FEN1, respectively. Shown on top are POLβ/FEN1 ratios designed to recapitulate the ratios found in the specified HD mouse brain tissues. The position of the 23mer radiolabeled primer (0) or various extension products is indicated to the right.
doi:10.1371/journal.pgen.1000749.g006

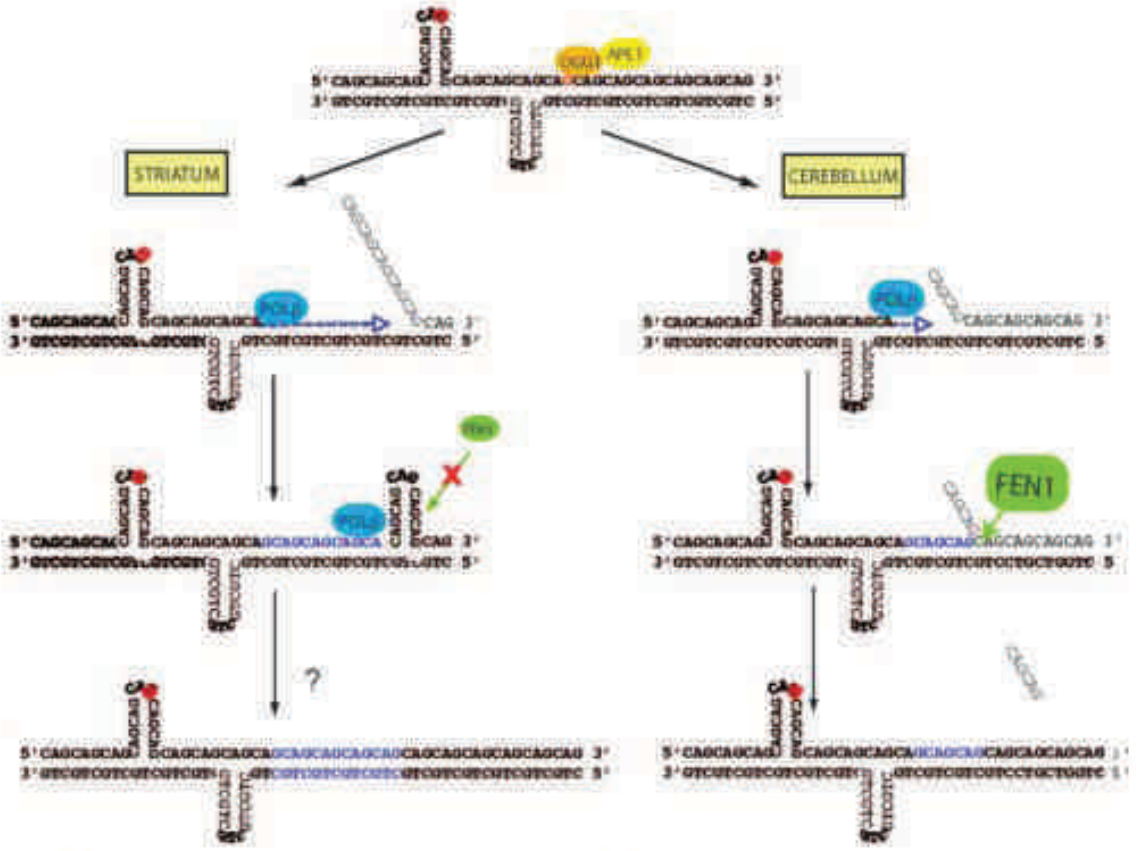


Figure 7. Model for somatic CAG expansion in HD. Pre-existing secondary DNA structures formed within CAG repeats prevent the accessibility of BER enzymes to oxidative DNA damage (red letters). When lesions become accessible, which occurs due to the dynamic nature of CAG repeats, repair can be initiated by OGG1. In the striatum, low flap endonuclease (FEN1) activity prevents cleavage of the 5′-flap structure generated by POLβ strand displacement activity, thereby leading to formation of an intermediate slipped strand structure. Processing of this intermediate structure, via an unknown mechanism, leads to somatic expansion. In contrast, in the cerebellum, high flap endonuclease (FEN1) activity permits efficient removal of the 5′-flap structure generated by POLβ strand displacement activity. Thus, no intermediate slipped strand structures are formed. Aging potentially worsens the situation because OGG1 activity tends to increase with age.
doi:10.1371/journal.pgen.1000749.g007

repeats have shown that stable slipped strand DNA or hairpin structures form in a repeat length-dependent manner [24,25]. Furthermore, three repair outcomes, including correct repair, escaped repair and error prone repair, have been observed for plasmid-based substrates that mimic slipped strand structures formed by CAG/CTG expansions, supporting the idea that processing of such structures by neuronal proteins can result in somatic expansion [29]. Although direct evidence is still lacking, several observations support that CAG/CTG expansions form secondary structures *in vivo* and the stability of these structures underlies the process of CAG/CTG repeat instability. First, longer repeats are more unstable in humans and mouse models of CAG/CTG repeat diseases [1,23]. Second, CAG/CTG repeat sequence interruptions, which reduce the propensity to form slipped strand DNA structures, also prevent instability in mice and humans [30,31]. Third, our data showing an extensive and repeat length-dependent accumulation of oxidative DNA damage at CAG expansions in HD mice imply that secondary structures form that are resistant to repair enzymes at the repeat locus. We in fact demonstrate that 8-oxoG and AP lesions, when located within the hairpin loop, cannot be efficiently incised *in vitro* by OGG1 and APE1 respectively. These results are in agreement with other studies showing that excision of 8-oxoG by OGG1 is paired-base-dependent and that the repair efficiency of some BER enzymes is affected by surrounding sequence context [32]. Thus, the probability that 8-oxoG lesions, an abundant endogenous DNA damage, are detected at CAG/CTG expansions by OGG1 should be significantly reduced if repeat expansions form secondary structures such as hairpins, as we indirectly demonstrate in Figure 2B and 2C).

Our results showing that POLβ is enriched at CAG expansions, in combination with the results from Kovtun *et al.* [13] that provide evidence that OGG1 modulates the extent of somatic CAG instability, indicate that at least some of the oxidative lesions present at CAG/CTG repeats are (or become) accessible and processed by the BER pathway. This would imply that the secondary structures formed by CAG/CTG repeats are dynamic in nature. Several studies support that MMR, replication and

transcription contribute to instability of CAG/CTG repeats, although the underlying mechanisms remain elusive [10]. These physiological processes induce chromatin remodeling and strand unpairing and, therefore, could provide a window for remodeling of secondary structures. While replication is unlikely to be involved in somatic instability in neurons, MMR and transcription could possibly interplay with BER at repeat expansions. During transcription, chromatin opens and the secondary structures in the transcribed strand have to be disrupted to allow RNA polymerase to proceed, only to re-form at a slightly different place. Thus, transcription could indirectly increase accessibility and the probability of repair of some oxidative DNA lesions by BER. The mechanism by which MMR contributes to somatic instability of CAG/CTG repeats is yet unknown. However, recent data indicate that functional MMR is required [33,34], suggesting that MMR proteins not only bind, but also disrupt loops or secondary structures at CAG/CTG repeats. Whether BER cooperates with other DNA-associated mechanisms to promote somatic CAG instability is an intriguing possibility.

Using HD mice, we have found that the global level of AP sites was elevated in the mouse cerebellum, which displays little instability relative to the striatum (Figure 2A). In addition, abnormal accumulation of DNA damage at CAG expansions was similar in cerebellum and striatum and did not increase with age (Figure 2B). Thus, the propensity for somatic instability, which increases with age and varies between tissues, does not strictly correlate with levels of DNA lesions at CAG expansions, unlike the toxic oxidation cycle model proposed by Kovtun *et al.* [13]. We have found that the corresponding activity and relative stoichiometry of major BER enzymes, including POLβ, APE1 and FEN1 varied between the striatum and cerebellum and to a lesser extent with age (Figure 4 and Figure S3). In particular, FEN1 protein level and 5′-flap endonuclease activity were much lower in the striatum than in the cerebellum. We therefore propose that stoichiometry and the relevant activity levels of the corresponding BER enzymes, rather than the level of oxidative DNA damage at repeats, is crucial in determining the probability that repair of lesions at CAG expansions drives somatic expansion.

A 5′-flap endonuclease activity, such as that of FEN1, has for some time been implicated in models of instability for CAG/CTG repeats. In yeast, CAG/CTG instability is enhanced in strains defective for rad27 (FEN1) in a repeat-length dependent manner [14,35]. It has been hypothesized that flap structures formed at CAG/CTG repeats during replication or repair inhibit FEN1, because they form complex secondary structures. In agreement, *in vitro* results have shown that secondary structures, such as hairpins, reduce processing by FEN1 at CAG/CTG repeats in a length-dependent manner [36]. Ligation of the unprocessed flap to produce an expansion mutation is supported by a study in yeast showing that overexpression of *cdc9* (Ligase I) increases the rates of trinucleotide repeat expansion [37]. However, FEN1 haploinsufficiency did not change the extent of somatic instability in HD and DM1 mice, though intergenerational expansion tended to increase in the HD background [30,39]. As complete inactivation of *Fen1* is lethal, HD and DM1 mouse models could only be examined in the *Fen1* heterozygous state. One cannot therefore exclude that a compensatory activity prevents exacerbation of repeat instability in a $Fen1^{+/-}$ background. By showing that FEN1 protein and 5′-flap endonuclease activity levels in the striatum and cerebellum are inversely correlated with the propensity for somatic expansion, we revive a role for FEN1 in preventing CAG repeat instability.

In BER, FEN1 is required to remove the 5′-flap structure generated during LP-BER strand displacement DNA synthesis [40,41]. Interestingly, studies support that the strand displacement activity of POLβ is also crucial in LP-BER in neuronal cells [42]. Furthermore, *in vitro* experiments have shown that FEN1 strongly stimulates the strand displacement activity of POLβ and, reciprocally, POLβ stimulates FEN1 [43,44]. It has been suggested that the functional interaction between POLβ and FEN1 controls incision product size in LP-BER: the tighter the cooperation, the shorter the product. Thus, cooperation between POLβ and FEN1 is believed to be essential to allow efficient repair of an oxidative lesion via LP-BER [43,44]. Our results show that POLβ is enriched at CAG repeats in the striatum of R6/1 mice (Figure 5). Additionally, gap filling and flap endonuclease activities are much lower in the striatum than in the cerebellum (Figure 4), and the molar ratio of FEN1/POLβ proteins is decreased by about 4-fold in the striatum when compared to the cerebellum (Figure S3B). Additionally, we show that differential molar ratios of FEN1 influence DNA synthesis length within CAG repeat-containing substrates by POLβ *in vitro*, with protein stoichiometries consistent with those found in the striatum leading to increased synthesis and stoichiometries like those found in the cerebellum causing reduced nucleotide incorporation (Figure 6). Altogether, the findings suggest that poor cooperation between POLβ and FEN1, as apparently present in the striatum of R6/1 animals, would result in incorporation of long stretches of nucleotides by POLβ strand displacement activity and the formation of long 5′-flap structures that are generally resistant to FEN1 cleavage activity and precursors to somatic expansion (see Figure 7).

In vitro studies have shown that FEN1 is also stimulated by other BER partners, including APE1 [44–46], and by cofactors such as HMGB1 [28], which are both higher in the cerebellum when compared to the striatum (Figure 4). Thus, low APE1 and HMGB1 levels in the striatum might further contribute to the intrinsically low levels of flap endonuclease activity, thereby reducing the probability that structured 5′-flaps at CAG/CTG repeats are processed by FEN1. Alternatively, studies have shown that FEN1 possesses, in addition to its flap endonuclease activity, exonuclease and gap-dependent endonuclease activities, which can help process unusual 5′-flap structures, such as those generated by triplet repeat sequences during maturation of Okazaki fragments [47]. Interestingly, when assaying 5′-flap endonuclease activity (Figure 4C), we identified an alternative product of shorter size, present only in the cerebellum of 36 week-old R6/1 mice, suggesting that additional exonuclease or endonuclease activities exist in the cerebellum, but not in the striatum of old transgenic animals. These activities might help cerebellar neurons to remove the structured 5′-flaps formed at CAG repeats. Conversely, the low intrinsic endonuclease and exonuclease activities of FEN1 (or FEN1-related enzymes) in the striatum would lead to the persistence of structured flaps, thereby resulting in the generation of pre-expanded alleles, which could be processed through an error prone mode to facilitate somatic expansion as shown by Panigrahi *et al.* [29] (see Figure 7).

In conclusion, our results agree with a model that incorporates a role for oxidative DNA damage and BER in somatic CAG instability, at least in the context of HD. Our results suggest that, in the cerebellum, optimal cooperation between gap filling by POLβ and 5′-flap excision by FEN1 during the repair of oxidative lesions via LP-BER prevents (or at least limits) formation of slipped strand structures at CAG repeats. Conversely, in the striatum, poor cooperation between these two enzymes likely leads to the formation of complex intermediate structures, which, as shown by Panigrahi *et al.* [29], can then be processed through an error prone mechanism to foster somatic expansion.

Note. While this manuscript was under revision, two papers came out from Liu *et al.* [48] and Lopez Castel *et al.* [49] that

support a contribution of poor coordination between specific enzymatic steps during DNA damage repair to somatic expansion of CAG/CTG repeats. The study of Liu *et al.* [48], based on *in vitro* experiments, indicates that dysfunctional coordination between POLβ and FEN1 during LP-BER triggers somatic expansion of substrates containing CAG repeats. Our study nicely complements these data by proposing that the lack of coordination between the two corresponding enzymatic steps of BER is correlated with the stoichiometric ratios of POLβ and FEN1 and is tissue-dependent. The study of Lopez Castel *et al.* [49], which is based on the use of mammalian cell lines impaired for Ligase I, extends this view by suggesting that coordination with the downstream ligation step is also crucial.

Materials and Methods

Mouse lines and breeding

Hemizygous R6/1 and R6/2 (160 CAG) and R6/2 (100 CAG) mice from the Jackson Laboratory were maintained on a mixed CBAxC57BL/6 genetic background [21]. Hdh^{Q111}, Hdh^{Q150} and Hdh^{Q250} heterozygotes were maintained on a CD1 outbred genetic background [23]. The experiments were approved by the ethical committee C.R.E.M.E.A.S (Comite Regional d'Ethique en Matiere d'Experimentation Animale de Strasbourg).

DNA and RNA extraction

The isolation of genomic DNA from mouse striatum and cerebellum for analysis of oxidative DNA damage was performed under conditions that minimize *ex vivo* oxidation artifacts according to the protocol developed by Lu *et al.* [22]. To this end, the silica-gel-membrane based DNeasy Tissue Kit (Qiagen) was used. Importantly, 50 μM of the free radical spin trap phenyl-tert-butyl nitrone (PBN, Sigma) was included in all buffers. High temperature and phenol use were avoided. DNA extracts were treated with RNase I. Striatum and cerebellum DNA extracted under these conditions was also used for CAG repeat sizing. Tail DNA for CAG repeat sizing was isolated using a standard protocol. Total RNA for CAG sizing and quantitative RT-PCR analysis was prepared using the RNeasy Mini Kit (Qiagen).

CAG repeat sizing and analysis

CAG repeat size was determined by PCR amplification using the HEX labeled primer 31329 and primer 33934 previously described [20]. PCR reactions were performed using the expand high fidelity DNA polymerase (Roche), according to manufacturer's instructions. PCR reaction products were subsequently purified using the Nucleospin Extraction II Kit (Macherey-Nagel). Products were then analyzed using the ABI Prism 3100 DNA analyzer instrument and GeneScan and Genotyper softwares. Size calibration was performed by including ROX 500 or ROX 1000 (Applied Biosystems) with the analyzed PCR products. Amplitude of Genescan profile was determined by calculating the number of peaks above 10% of the maximum fluorescent peak intensity. From the amplitude value, we deduced the median peak.

Quantitative amplification of CAG repeats

To quantitatively amplify CAG repeats from R6/1 and R6/2 DNA or RNA, we used the following protocol. DNA was amplified with primers 31329 and 33934, previously described [20] using the Herculase Hotstart DNA Polymerase (Stratagene). Concentrations of dNTP and primers were those recommended by the manufacturer. 8% DMSO was included in the reaction, as well as Sybr green (Molecular probe). The PCR reactions were performed and analyzed on a Light Cycler instrument (Roche). PCR cycling conditions were as follows: DNA was first inactivated for 3 min at 90°C, followed by 45 cycles consisting in 10 seconds at 90°C, 30 seconds at 60°C and 2 min at 72°C.

Real-time RT-PCR analysis

Reverse transcription was performed on 1 μg of total RNA using SuperScriptII (Invitrogen) and random hexamers according to the manufacturer instructions. We performed PCR amplification of cDNA on a Light-Cycler instrument (Roche). PCR primers for detection of *Ogg1*, *Ape1*, *Polβ*, *Fen1*, *Hprt*, *Gapdh* and *36B4* are available upon request.

Detection of ABASIC (AP) sites

Detection of AP sites was performed using the DNA damage quantification kit (Biovision), according to the manufacturer's instructions. The method is based on specific reaction of the Aldehyde Reactive Probe (ARP) reagent with the open ring form of AP sites. Briefly, genomic DNA was treated with ARP tagged with biotin residues. AP sites in the DNA were then quantified using an avidin-biotin assay followed by a colorimetric reaction [50].

DNA damage analysis

DNA damage at specific gene loci was assayed by cleavage of genomic DNA with bacterial formamidopyrimidine glycosylase (Fpg). Fpg is an AP-lyase that specifically excises 8-oxoG among other oxidized bases, and then creates a single strand break at the site of the abasic product. Quantitative real time PCR was used to determine the level of intact DNA at CAG repeats before and after DNA cleavage by Fpg. As a control, a portion of the murine *Hb* gene, which is similar in size and GC content to the CAG-expanded locus, was amplified. The ratio of PCR products after Fpg cleavage to those present in untreated DNA was used to determine the level of intact DNA. We followed the protocol described by Lu *et al.* [22], with the following modifications: Fpg from Sigma was used and incubated with genomic DNA for 1h at 37°C at a concentration of 0.2 μg Fpg/μg of DNA. These conditions, which were predetermined by an Fpg dose response curve and time course, allowed the reaction to reach a steady state. After inactivation at 60°C for 5 min, the Fpg reaction underwent an ethanol precipitation step. DNA was recovered in water and then submitted for quantitative PCR. Conditions for quantitative amplification of the CAG-expanded locus are described above. Quantitative amplification of the *Hdh* locus was performed using a commercial PCR master mix (Qiagen) with forward (5'-TCGAGTCGCTCAAGTCGTTT-3') and reverse (5'-ACTTG-GCAAACTGGGAACGG-3') primers and PCR conditions were as follows: a first step of denaturation was performed at 95°C for 15 min, followed by 45 cycles consisting in a denaturation at 95°C for 30 seconds, hybridization for 30 seconds at 55°C and elongation for 1 min at 72°C.

Western blotting

Striata and cerebella were dissected and homogenized in lysis buffer containing 50 mM Tris-HCl pH 8.0, 10% glycerol, 5 mM EDTA, 150 mM KCl, a cocktail of protease inhibitors (Roche) and 1% NP-40. The extracts were incubated for 15 min on ice and centrifuged for 20 min at 13000 rpm and 4°C. Supernatants were collected and analyzed on SDS-PAGE gels. Rabbit α-OGG1 (Abcam), mouse α-POLβ (Biovision), rabbit α-APE1 (Abcam), rabbit α-FEN1 (Santa Cruz or Abcam), rabbit α-HMGB1 (Abcam) and mouse α-β-tubulin (Chemicon) were used at 1:1000 dilutions and revealed with appropriate α-rabbit or α-mouse peroxidase-conjugated secondary antibodies (Jackson immunoResearch

Laboratories) and the ECL chemiluminescence reaction (Pierce or Millipore).

OGG1 and APE1 activity at lesions located at hairpin structures

Hairpin DNA substrates with specific modifications (Figure 3A) were purchased from Midland Certified Reagent Company, Inc. (Midland, TX). [γ^{32}P]ATP 5′ radiolabeled oligonucleotides were generated as described [51]. After annealing, the labeled oligonucleotides were purified from unincorporated [γ^{32}P]ATP by using a Bio-Rad Micro Bio-Spin P30 column. Briefly, after spinning the column at 1000 g for 1 min to remove packing buffer, the columns used for purification of the 5′-^{32}P THF substrates were washed twice with OPT buffer (25 mM Mops, pH 7.2, 100 mM KCl, 1 mM $MgCl_2$) and the columns used for 5′-^{32}P 8-oxoG substrates were washed twice with NEB2 buffer (New England Biolabs). Columns were centrifuged as above for each washing step. The labeled oligonucleotides were then eluted through the appropriate column by centrifugation at 1000 g for 4 min. The hairpin structure was verified by incubation of the 5′-^{32}P labeled oligonucleotides (0.2 pmol) with *Bam*HI or *Eco*RI (0.1 to 4U; New England Biolabs) and subsequent electrophoresis on a 15% polyacrylamide urea denaturing gel. OGG1 incision assays were performed by incubating 5′-^{32}P labeled 8-oxoG substrate (0.2 pmol) with hOGG1 protein (0 to 30 ng; New England Biolabs) in NEB2 buffer at 37°C for 15 min. APE1 incision assays were performed by incubating 5′-^{32}P labeled THF substrate (0.2 pmol) with hAPE1 protein (0 to 1 ng) in OPT buffer at 37°C for 5 min. Reactions were inhibited by the addition of stop buffer (95% formamide, 20 mM ethylenediaminetetraacetic acid [EDTA], 0.5% bromophenol blue and 0.5% xylene cyanol), and then heated at 95°C for 5 min. Reaction products were resolved by 15% polyacrylamide urea denaturing gel electrophoresis and imaged using a Typhoon phosphoimager.

DNA repair assays

Cell extracts were prepared essentially as described [52]. Briefly, frozen striata or cerebella were homogenized in 10 mM HEPES-KOH, pH 7.7, 0.5 mM $MgCl_2$, 10 mM KCl, 1 mM DTT buffer and then centrifuged at 2000 g at 4°C for 10 min. The pellet was resuspended in 20 mM HEPES-KOH, pH 7.7, 0.5 mM $MgCl_2$, 420 mM NaCl, 0.2 mM EDTA, 25% glycerol, PIC (Complete Protease Inhibitor Cocktail EDTA-free, ROCHE), 1 mM DTT and gently stirred at 4°C for 20–30 min to allow for efficient nuclear lysis. The suspension was centrifuged at 14000 g at 4°C for 15 min. The supernatant was dialyzed against 40 mM HEPES-KOH, pH 7.7, 50 mM KCl and 2 mM DTT buffer overnight at 4°C. Cell extract concentrations were determined using a Bradford assay. OGG1, APE1, POLβ and FEN1 assays were performed using oligonucleotide substrates described by [26], see Figure S2. 0.1 pmol of [γ^{32}P]ATP 5′-radiolabeled oligonucleotides were incubated with 10 μg of cell extract and assay buffer (75 mM KCl, 25 mM $MgCl_2$, 7.5 mM dNTPs, 1.125 mM HEPES-KOH, pH 7.7, 1% glycerol, 0.25 μM EDTA) at 37°C. Incubation times were optimized for each repair activity: 3 h, 1 h, 2 h and 30 min were used for 8-oxoG incision (OGG1), AP-endonuclease (APE1), gap-filling (POLβ) and flap endonuclease (FEN1) assays, respectively. The samples were then treated with 0.5% SDS and 0.6 mg/mL of Proteinase K, heated at 55°C for 15 min, purified by phenol-chloroform extraction and resuspended in solution by adding equal volume of 98% formamide, 10 mM EDTA, bromophenol blue and xylene cyanol buffer. Reaction products were resolved by 20% polyacrylamide urea denaturing gel electrophoresis and imaged on radiographic films. Images were captured with GeneSnap and quantified with GeneTool software on Syngene Chemigenius XE machine.

POLβ DNA synthesis reactions in the presence of FEN1

2 nM of purified human DNA POLβ [33] was incubated with increasing concentrations of purified human FEN1 [34] (1 nM, 2 nM, 4 nM, 8 nM and 16 nM) and 200 fmol of substrate for 30 min at 37°C in 50 mM HEPES-KOH pH 7.5, 10 mM $MgCl_2$, 0.5 mM EDTA, 2 mM DTT, 2 mM ATP, and 20 μM each dATP, dCTP, dGTP, and dTTP. Reactions were stopped by the addition of stop buffer consisting of 90% formamide and 20 mM EDTA, boiled for 3 min and then loaded onto a 15% polyacrylamide-urea gel and run at 300 V for 5 hrs.

Chromatin immunoprecipitation (ChIP)

ChIP experiments were performed essentially as described [55]. For each ChIP experiment, striata and cerebella from 2 to 4 transgenic mice were pooled, cut into small fragments, fixed by adding 37% formaldehyde to a final concentration of 1% and incubated for 10 min at room temperature. Cross-linking was stopped by addition of glycine to 0.125 M. Tissue fragments were washed three times with cold phosphate-buffered saline and treated with sonication buffer (50 mM HEPES pH 7.9, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium dodecyl sulfate, 0.1% Na-deoxycholate) containing protease and phosphatase inhibitors. Tissue was then homogenized, and lysates were sonicated to obtain DNA fragments of 200 to 1000 bp, as revealed by ethidium bromide staining of aliquots separated on agarose gels (Figure S4). Samples were centrifuged to pellet debris and an aliquot was taken for gel analysis and input. The soluble chromatin fraction was pretreated for 1 h at 4°C with protein A Agarose/Salmon Sperm DNA (50% slurry- Millipore). Samples were then incubated overnight at 4°C with α-DNA POLβ ab194 (ChIP grade, Abcam) or α-AcH3K9/14 (Upstate) antibodies. Protein A Agarose/Salmon Sperm DNA was then added, and the mixture was incubated for 2 h at 4°C. Agarose beads were washed twice for 10 min with sonication buffer, twice for 10 min with wash buffer A (sonication buffer with 500 mM NaCl), twice for 10 min with wash buffer B (20 mM Tris-HCl, pH 8.0, 1 mM EDTA, 0.25 M LiCl, 0.5% NP-40, 0.5% Na-deoxycholate), and finally with Tris-EDTA (TE, pH 8.0). Immune complexes were eluted from the beads with 1% SDS in TE (pH 8.0) and protein-DNA cross-links were reversed by adding 200 mM NaCl and heating overnight at 65°C. After treatment with proteinase K for 2 h at 42°C, the samples were purified by phenol-chloroform-isoamyl alcohol extraction and precipitated with ethanol. One-sixth (for amplification of CAG expansions) to one-fifteen (for amplification of a fragment of the *Hdh* gene) of the immunoprecipitated DNA and 1% of the input DNA were quantified by real-time quantitative PCR (see above). Results are expressed relative to the amount of input DNA per ChIP.

Supporting Information

Figure S1 Quantitative PCR amplification of CAG repeat locus from R6/1 mice. (A) Schematic representation of the HD transgene and the region surrounding exon 1 at the *Hdh* locus. Location of the primers used to amplify the CAG-expanded fragment and the *Hdh* control region is denoted by arrows. (B) Analysis of the GC content of the PCR fragment containing CAG expansions and of the PCR fragment located at the *Hdh* locus using DNA strider software. Both fragments are similar in size and GC content (around 70%). (C) Real time quantitative PCR amplification of CAG expansion from R6/1 striatum and cerebellum. Top left. Fusion profiles showing

that primers 31329 and 33934 allow for amplification of a product specific to R6/1 mice. Bottom. Representative analysis of PCR amplification of CAG expansion from the striatum of R6/1 showing PCR is relative to the quantity of DNA doubling at each cycle (slope close to -3,3). Top right. Histogram showing that the relative DNA concentration calculated by the Light Cycler software is proportional to the initial quantity of DNA and similar between striatum and cerebellum.

Found at: doi:10.1371/journal.pgen.1000749.s001 (0.74 MB TIF)

Figure S2 Oligonucleotide substrates used to assess BER activities from mouse tissues. (A) Table showing the modified oligonucleotides used to assess the different steps involved in BER. Oligonucleotides containing an 8-oxodG and a tetrahydrofuran modification (THF) are used to assess glycosylase and AP-endonuclease activities, respectively. The gap-filling activity was assessed with two adjacent oligonucleotides producing a 1-nucleotide gap. 5′-flap excision activity was evaluated using an oligonucleotide with a 10 nt flap. The enzymes that mainly carry out the corresponding reactions, i.e. OGG1, APE1, POLβ and FEN1, respectively, are shown on the left of the table. Due to functional redundancy other enzymes may contribute to the reactions. (B) Sequences of the oligonucleotides described above.

Found at: doi:10.1371/journal.pgen.1000749.s002 (0.24 MB TIF)

Figure S3 Stoichiometry of BER proteins is different in striatum and cerebellum of R6/1 mice. (A) Table showing the relative protein levels and activities of the designated BER proteins including OGG1, APE1, FEN1 and POLβ, in the striatum and cerebellum of R6/1 and control (WT) mice at 40 weeks of age. (B) Steady state levels of FEN1 and POLβ in the striatum and cerebellum of R6/1 mice were evaluated by western blotting using purified recombinant proteins corresponding to human FEN1 (42 kDa) and human POLβ (39 kDa), respectively. Top, 100 μg of whole cell extract (WCE) from the cerebellum of an R6/1 mouse were run on an SDS-polyacrylamide gel together with 15 ng and 150 ng recombinant FEN1 (left) or 15 ng and 150 ng recombinant POLβ (right) and detected with either α-FEN1 or α-POLβ antibodies. Band intensities were quantified and expressed as relative fold changes, which allowed calculation of FEN1:POLβ molar ratio. Bottom. WCE extracts from the striatum and cerebellum of two different 40 week-old R6/1 mice (numbered 1 and 2) were run on a gel and analyzed with α-FEN1 (rabbit), α-POLβ (mouse) and α-β-Tubulin (mouse). α-β-Tubulin was used to control sample loading. The same membrane was sequentially probed. Band intensities were quantified and the FEN1/POLβ molar ratio was estimated in the striatum and cerebellum of R6/1 mice. One representative set of detection with α-FEN1 and α-POLβ antibodies is shown. The extracts were loaded on gels and the antibody signal quantified 3 times independently. The mean and sem of the FEN1:POLβ molar ratios obtained after quantification of the 3 experiments are reported in the table.

Found at: doi:10.1371/journal.pgen.1000749.s003 (0.47 MB TIF)

Figure S4 Sonication of striatum and cerebellum extracts from R6 mice generates DNA fragments between 100 and 1,000 bp. DNA from striatum and cerebellum of R6/1 mice at 6 and 37 weeks of age was sonicated and analyzed by running aliquots on ethidium bromide stained agarose gels. DNA is sonicated to fragments below 1,000 bp.

Found at: doi:10.1371/journal.pgen.1000749.s004 (0.64 MB TIF)

Acknowledgments

We thank F. Coin, E. Compe, V. Schreiber, F. Klein, A. Martelli, and J-L. Mandel for helpful discussions; C. Weber, E. Metzger, and staff at the IGBMC animal facility for mouse care and genotyping.

Author Contributions

Conceived and designed the experiments: DMW KM. Performed the experiments: AVG BRB KM. Analyzed the data: AVG BRB DMW YT KM. Contributed reagents/materials/analysis tools: DMW VCW YT KM. Wrote the paper: AVG BRB DMW KM.

References

1. Duyao M, Ambrose C, Myers R, Novelletto A, Persichetti F, et al. (1993) Trinucleotide repeat length instability and age of onset in Huntington's disease. Nat Genet 4: 387–392.
2. Telenius H, Kremer B, Goldberg YP, Theilmann J, Andrew SE, et al. (1994) Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm. Nat Genet 6: 409–414.
3. Wheeler VC, Persichetti F, McNeil SM, Mysore JS, Mysore SS, et al. (2007) Factors associated with HD CAG repeat instability in Huntington disease. J Med Genet 44: 695–701.
4. Kennedy L, Evans E, Chen CM, Craven L, Detloff PJ, et al. (2003) Dramatic tissue-specific mutation length increases are an early molecular event in Huntington disease pathogenesis. Hum Mol Genet 12: 3359–3367.
5. Shelbourne PF, Keller-McGandy C, Bi WL, Yoon SR, Dubeau L, et al. (2007) Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain. Hum Mol Genet 16: 1133–1142.
6. Gonitel R, Moffitt H, Sathasivam K, Woodman B, Detloff PJ, et al. (2008) DNA instability in postmitotic neurons. Proc Natl Acad Sci U S A 105: 3467–3472.
7. Swami M, Hendricks AE, Gillis T, Massood T, Mysore J, et al. (2009) Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. Hum Mol Genet 18: 3039–3047.
8. Wheeler VC, Lebel LA, Vrbanac V, Teed A, te Riele H, et al. (2003) Mismatch repair gene Msh2 modifies the timing of early disease in Hdh(Q111) striatum. Hum Mol Genet 12: 273–281.
9. Dragileva E, Hendricks A, Teed A, Gillis T, Lopez ET, et al. (2009) Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. Neurobiol Dis 33: 37–47.
10. Pearson CE, Nichol Edamura K, Cleary JD (2005) Repeat instability: mechanisms of dynamic mutations. Nat Rev Genet 6: 729–742.
11. Manley K, Shirley TL, Flaherty L, Messer A (1999) Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. Nat Genet 23: 471–473.
12. Kovtun IV, McMurray CT (2001) Trinucleotide expansion in haploid germ cells by gap repair. Nat Genet 27: 407–411.
13. Kovtun IV, Liu Y, Bjoras M, Klungland A, Wilson SH, et al. (2007) OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. Nature 447: 447–452.
14. Freudenreich CH, Kantrow SM, Zakian VA (1998) Expansion and length-dependent fragility of CTG repeats in yeast. Science 279: 853–856.
15. Fortini P, Dogliotti E (2007) Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways. DNA Repair (Amst) 6: 398–409.
16. Robertson AB, Klungland A, Rognes T, Leiros I (2009) DNA repair in mammalian cells: Base excision repair: the long and short of it. Cell Mol Life Sci 66: 981–993.
17. Larsen E, Reite K, Nesse G, Gran C, Seeberg E, et al. (2006) Repair and mutagenesis at oxidized DNA lesions in the developing brain of wild-type and Ogg1-/- mice. Oncogene 25: 2425–2432.
18. Imam SZ, Karahalil B, Hogue BA, Souza-Pinto NC, Bohr VA (2006) Mitochondrial and nuclear DNA-repair capacity of various brain regions in mouse is altered in an age-dependent manner. Neurobiol Aging 27: 1129–1136.
19. Browne SE, Beal MF (2006) Oxidative damage in Huntington's disease pathogenesis. Antioxid Redox Signal 8: 2061–2073.
20. Mangiarini L, Sathasivam K, Mahal A, Mott R, Seller M, et al. (1997) Instability of highly expanded CAG repeats in mice transgenic for the Huntington's disease mutation [see comments]. Nat Genet 15: 197–200.
21. Mangiarini L, Sathasivam K, Seller M, Cozens B, Harper A, et al. (1996) Exon 1 of the *HD* gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. Cell 87: 493–506.
22. Lu T, Pan Y, Kao SY, Li C, Kohane I, et al. (2004) Gene regulation and DNA damage in the ageing human brain. Nature 429: 883–891.
23. Wheeler VC, Auerbach W, White JK, Srinidhi J, Auerbach A, et al. (1999) Length-dependent gametic CAG repeat instability in the Huntington's disease knock-in mouse. Hum Mol Genet 8: 115–122.
24. Gacy AM, Goellner G, Juranic N, Macura S, McMurray CT (1995) Trinucleotide repeats that expand in human disease form hairpin structures in vitro. Cell 81: 533–540.

25. Pearson CE, Sinden RR (1996) Alternative structures in duplex DNA formed within the trinucleotide repeats of the myotonic dystrophy and fragile X loci. Biochemistry 35: 5041–5053.
26. Wilson DM 3rd (2005) Ape1 abasic endonuclease activity is regulated by magnesium and potassium concentrations and is robust on alternative DNA structures. J Mol Biol 345: 1003–1014.
27. McNeill DR, Narayana A, Wong HK, Wilson DM 3rd (2004) Inhibition of Ape1 nuclease activity by lead, iron, and cadmium. Environ Health Perspect 112: 799–804.
28. Prasad R, Liu Y, Deterding LJ, Poltoratsky VP, Kedar PS, et al. (2007) HMGB1 is a cofactor in mammalian base excision repair. Mol Cell 27: 829–841.
29. Panigrahi GB, Lau R, Montgomery SE, Leonard MR, Pearson CE (2005) Slipped (CTG)*(CAG) repeats can be correctly repaired, escape repair or undergo error-prone repair. Nat Struct Mol Biol 12: 654–662.
30. Pearson CE, Wang YH, Griffith JD, Sinden RR (1998) Structural analysis of slipped-strand DNA (S-DNA) formed in (CTG)n. (CAG)n repeats from the myotonic dystrophy locus. Nucleic Acids Res 26: 816–823.
31. Gray M, Shirasaki DI, Cepeda C, André VM, Wilburn B, et al. (2008) Full-length human mutant huntingtin with a stable polyglutamine repeat can elicit progressive and selective neuropathogenesis in BACHD mice. J Neurosci 28: 6182–6195.
32. Ide H (2001) DNA substrates containing defined oxidative base lesions and their application to study substrate specificities of base excision repair enzymes. Prog Nucleic Acid Res Mol Biol 68: 207–221.
33. Gomes-Pereira M, Fortune MT, Ingram L, McAbney JP, Monckton DG (2004) Pms2 is a genetic enhancer of trinucleotide CAG·CTG repeat somatic mosaicism: implications for the mechanism of triplet repeat expansion. Hum Mol Genet 13: 1815–1825.
34. Tome S, Holt I, Edelmann W, Morris GE, Munnich A, et al. (2009) MSH2 ATPase domain mutation affects CTG*CAG repeat instability in transgenic mice. PLoS Genet 5: e1000482. doi:10.1371/journal.pgen.1000482.
35. Yang J, Freudenreich CH (2007) Haploinsufficiency of yeast FEN1 causes instability of expanded CAG/CTG tracts in a length-dependent manner. Gene 393: 110–115.
36. Spiro C, Pelletier R, Rolfsmeier ML, Dixon MJ, Lahue RS, et al. (1999) Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats. Mol Cell 4: 1079–1085.
37. Subramanian J, Vijayakumar S, Tomkinson AE, Arnheim N (2005) Genetic instability induced by overexpression of DNA ligase I in budding yeast. Genetics 171: 427–441.
38. Spiro C, McMurray CT (2003) Nuclease-deficient FEN-1 blocks Rad51/BRCA1-mediated repair and causes trinucleotide repeat instability. Mol Cell Biol 23: 6063–6074.
39. van den Broek WJ, Nelen MR, van der Heijden GW, Wansink DG, Wieringa B (2006) Fen1 does not control somatic hypermutability of the (CTG)n*(CAG)n repeat in a knock-in mouse model for DM1. FEBS Lett 580: 5208–5214.
40. Liu Y, Kao HI, Bambara RA (2004) Flap endonuclease 1: a central component of DNA metabolism. Annu Rev Biochem 73: 589–615.
41. Kovtun IV, McMurray CT (2008) Features of trinucleotide repeat instability in vivo. Cell Res 18: 198–213.
42. Wei W, Englander EW (2008) DNA polymerase beta-catalyzed-PCNA independent long patch base excision repair synthesis: a mechanism for repair of oxidatively damaged DNA ends in post-mitotic brain. J Neurochem 107: 734–744.
43. Prasad R, Dianov GL, Bohr VA, Wilson SH (2000) FEN1 stimulation of DNA polymerase beta mediates an excision step in mammalian long patch base excision repair. J Biol Chem 275: 4460–4466.
44. Balakrishnan L, Brandt PD, Lindsey-Boltz LA, Sancar A, Bambara RA (2009) Long patch base excision repair proceeds via coordinated stimulation of the multi-enzyme DNA repair complex. J Biol Chem 284: 15158–15172.
45. Wong D, Demple B (2004) Modulation of the 5'-deoxyribose-5-phosphate lyase and DNA synthesis activities of mammalian DNA polymerase beta by apurinic/apyrimidinic endonuclease 1. J Biol Chem 279: 25268–25275.
46. Liu Y, Prasad R, Beard WA, Kedar PS, Hou EW, et al. (2007) Coordination of steps in single-nucleotide base excision repair mediated by apurinic/apyrimidinic endonuclease 1 and DNA polymerase beta. J Biol Chem 282: 13532–13541.
47. Singh P, Zheng L, Chavez V, Qiu J, Shen B (2007) Concerted action of exonuclease and Gap-dependent endonuclease activities of FEN-1 contributes to the resolution of triplet repeat sequences (CTG)n- and (GAA)n-derived secondary structures formed during maturation of Okazaki fragments. J Biol Chem 282: 3465–3477.
48. Liu Y, Prasad R, Beard WA, Hou EW, Horton JK, et al. (2009) Coordination between Pol (beta) and FEN1 can modulate CAG repeat expansion. J Biol Chem 284: 28352–28366.
49. Lopez Castel A, Tomkinson AE, Pearson CE (2009) CTG/CAG repeat instability is modulated by the levels of human DNA ligase I and its interaction with PCNA: A distinction between replication and slipped-DNA repair. J Biol Chem 284: 26631–26645.
50. Atamna H, Cheung I, Ames BN (2000) A method for detecting abasic sites in living cells: age-dependent changes in base excision repair. Proc Natl Acad Sci U S A 97: 686–691.
51. Berquist BR, McNeill DR, Wilson DM 3rd (2008) Characterization of abasic endonuclease activity of human Ape1 on alternative substrates, as well as effects of ATP and sequence context on AP site incision. J Mol Biol 379: 17–27.
52. Takao M, Kanno S, Shiromoto T, Hasegawa R, Ide H, et al. (2002) Novel nuclear and mitochondrial glycosylases revealed by disruption of the mouse Nth1 gene encoding an endonuclease III homolog for repair of thymine glycols. Embo J 21: 3486–3493.
53. Lee BI, Wilson DM 3rd (1999) The RAD2 domain of human exonuclease 1 exhibits 5' to 3' exonuclease and flap structure-specific endonuclease activities. J Biol Chem 274: 37763–37769.
54. Nguyen LH, Barsky D, Erzberger JP, Wilson DM 3rd (2000) Mapping the protein-DNA interface and the metal-binding site of the major human apurinic/apyrimidinic endonuclease. J Mol Biol 298: 447–459.
55. Compe E, Drane P, Laurent C, Diderich K, Braun C, et al. (2005) Dysregulation of the peroxisome proliferator-activated receptor target genes by XPD mutations. Mol Cell Biol 25: 6065–6076.

Figure S1.

Quantitative PCR amplification of CAG repeat locus from R6/1 mice. (A) Schematic representation of the HD transgene and the region surrounding exon 1 at the *Hdh* locus. Location of the primers used to amplify the CAG-expanded fragment and the *Hdh* control region is denoted by arrows. (B) Analysis of the GC content of the PCR fragment containing CAG expansions and of the PCR fragment located at the *Hdh* locus using DNA strider software. Both fragments are similar in size and GC content (around 70%). (C) Real time quantitative PCR amplification of CAG expansion from R6/1 striatum and cerebellum. Top left. Fusion profiles showing that primers 31329 and 33934 allow for amplification of a product specific to R6/1 mice. Bottom. Representative analysis of PCR amplification of CAG expansion from the striatum of R6/1 showing PCR is relative to the quantity of DNA doubling at each cycle (slope close to -3,3). Top right. Histogram showing that the relative DNA concentration calculated by the Light Cycler software is proportional to the initial quantity of DNA and similar between striatum and cerebellum.

Figure S2.

Oligonucleotide substrates used to assess BER activities from mouse tissues. (A) Table showing the modified oligonucleotides used to assess the different steps involved in BER. Oligonucleotides containing an 8-oxodG and a tetrahydrofuran modification (THF) are used to assess glycosylase and AP-endonuclease activities, respectively. The gap-filling activity was assessed with two adjacent olionucleotides producing a 1-nucleotide gap. 5′-flap excision activity was evaluated using an oligonucleotide with a 10 nt flap. The enzymes that mainly carry out the corresponding reactions, i.e. OGG1, APE1, POLβ and FEN1, respectively, are shown on the left of the table. Due to functional redundancy other enzymes may contribute to the reactions. (B) Sequences of the oligonucleotides described above.

Figure S3.

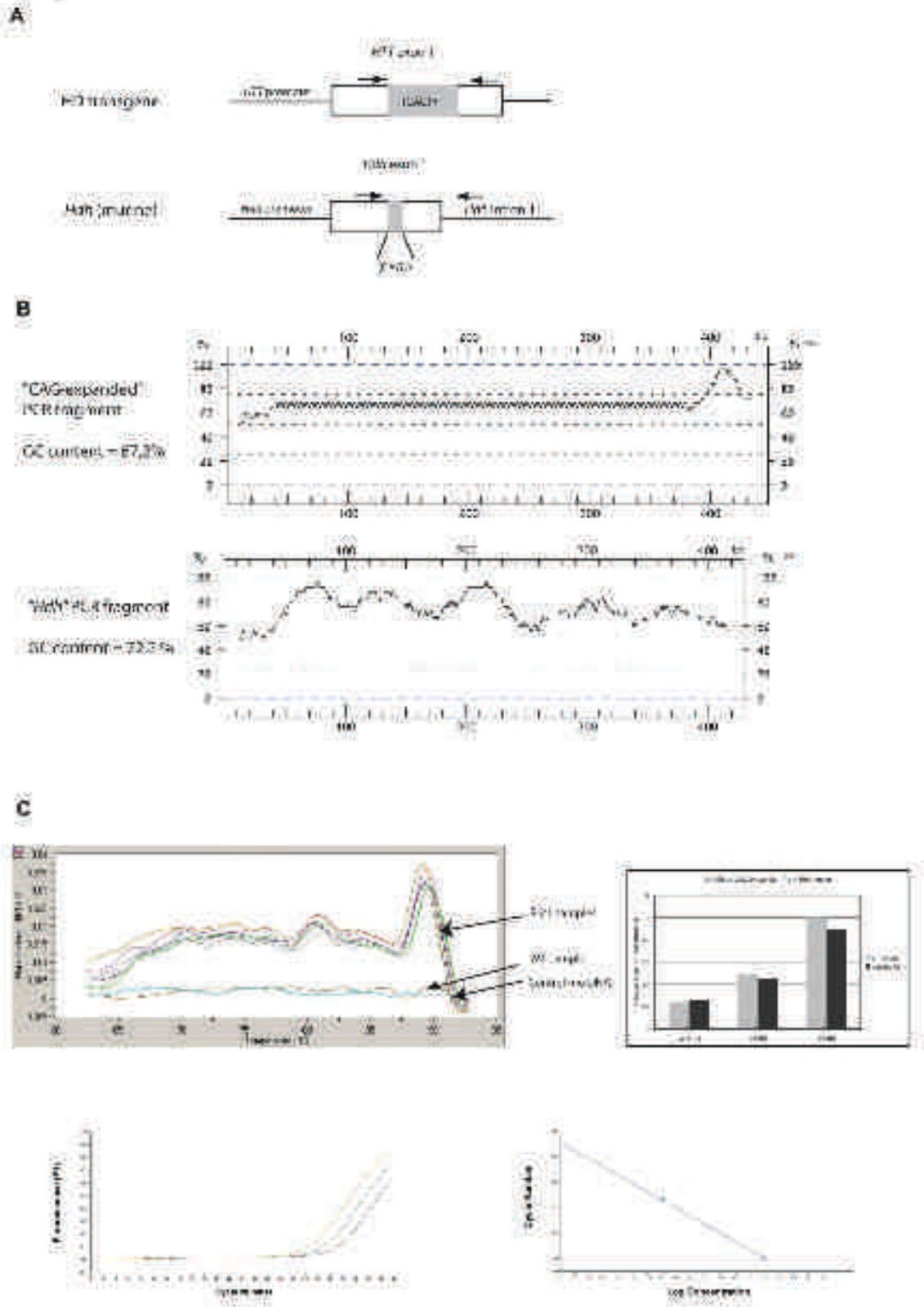
Stoichiometry of BER proteins is different in striatum and cerebellum of R6/1 mice. (A) Table showing the relative protein levels and activities of the designated BER proteins including

OGG1, APE1, FEN1 and POLβ, in the striatum and cerebellum of R6/1 and control (WT) mice at 40 weeks of age. (B) Steady state levels of FEN1 and POLβ in the striatum and cerebellum of R6/1 mice were evaluated by western blotting using purified recombinant proteins corresponding to human FEN1 (42 kDa) and human POLβ (39 kDa), respectively. Top. 100 µg of whole cell extract (WCE) from the cerebellum of an R6/1 mouse were run on an SDS-polyacrylamide gel together with 15 ng and 150 ng recombinant FEN1 (left) or 15 ng and 150 ng recombinant POLβ (right) and detected with either α-FEN1 or α-POLβ antibodies. Band intensities were quantified and expressed as relative fold changes, which allowed calculation of FEN1:POLβ molar ratio. Bottom. WCE extracts from the striatum and cerebellum of two different 40 week-old R6/1 mice (numbered 1 and 2) were run on a gel and analyzed with α-FEN1 (rabbit), α-POLβ (mouse) and α-β-Tubulin (mouse). α-β-Tubulin was used to control sample loading. The same membrane was sequentially probed. Band intensities were quantified and the FEN1/POLβ molar ratio was estimated in the striatum and cerebellum of R6/1 mice. One representative set of detection with α-FEN1 and α-POLβ antibodies is shown. The extracts were loaded on gels and the antibody signal quantified 3 times independently. The mean and sem of the FEN1:POLβ molar ratios obtained after quantification of the 3 experiments are reported in the table.

Figure S4.

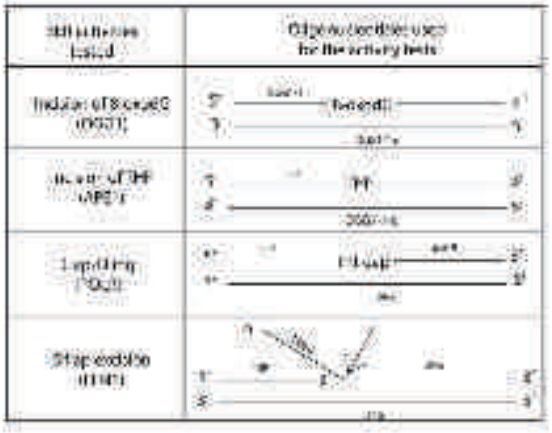
Sonication of striatum and cerebellum extracts from R6 mice generates DNA fragments between 100 and 1,000 bp. DNA from striatum and cerebellum of R6/1 mice at 6 and 37 weeks of age was sonicated and analyzed by running aliquots on ethidium bromide stained agarose gels. DNA is sonicated to fragments below 1,000 bp.

Sup Figure 1

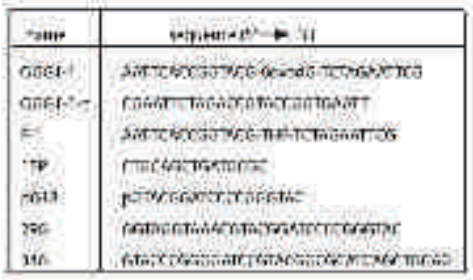


Sup Figure 2

A



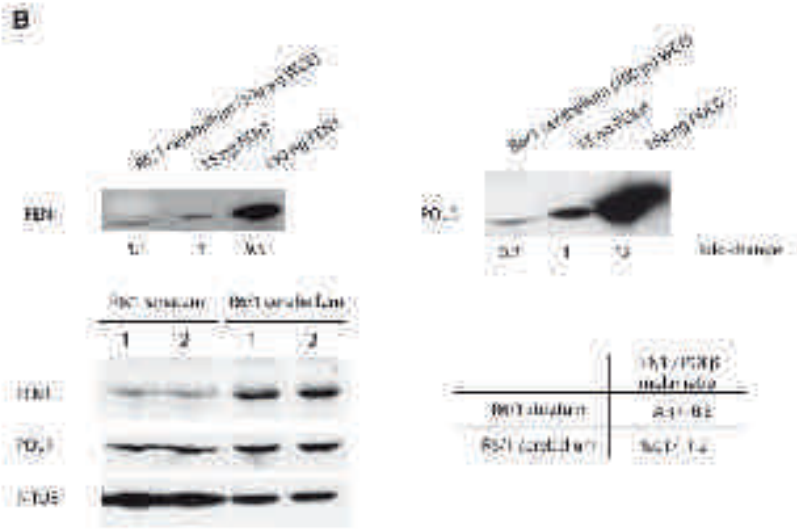
B



Sup Figure 3

A

B



Sup Figure 1

Chapter 2: Nucleotide sequence, DNA damage position and protein stoichiometry influence base excision repair outcome at CAG/CTG repeats

Introduction:

In the previous chapter, we proposed that the tissue-specific stoichiometry of BER proteins contributes to the tissue-selective instability of CAG/CTG repeats in HD (Goula et al., 2009). Briefly, our data showed that in wild-type and HD transgenic mice, BER protein levels and associated activities are lower in the striatum relative to the cerebellum. In particular, FEN1 protein and corresponding activity were decreased in the striatum when compared to the cerebellum, resulting in a lower FEN1:Polβ ratio. We hypothesized that oxidant lesions at CAG/CTG repeats are preferentially processed via LP-BER, and due to impaired coordination between DNA synthesis and 5'-flap removal, the striatum would be more prone to somatic instability. Accordingly, in vitro reconstituted repair assays suggest that the impairment of the steps mediated by Polβ and FEN1 during LP-BER results in CAG expansion (Liu et al., 2009).

LigI is the DNA ligase involved in sealing the repaired DNA during LP-BER, while LigIII is implicated in SN-BER (Cappelli et al., 1997; Sleeth et al., 2004). Coordination of the steps mediated by FEN1 and LigI is essential to insure proper repair, and requires PCNA, a common interactor for both proteins (Gary et al., 1999; Klungland and Lindahl, 1997; Levin et al., 2000; Wu et al., 1996). Interestingly, CAG/CTG instability was modulated by the levels of LigI and its interaction with PCNA in yeast and in a mammalian cell model (Lopez Castel et al., 2009; Subramanian et al., 2005). However, somatic CAG/CTG instability was not changed in DM1 mice expressing a mutated version of Lig1 that results in low residual activity (Tome et al., 2011). Thus, the role of Lig1 in CAG/CTG instability may be complex.

Repair outcome at CAG/CTG repeats has been previously investigated by incubating cell extracts with circular substrates containing CAG/CTG repeats and a nick located at various positions with respect to the CAG/CTG stretch (Hou et al., 2009; Panigrahi et al., 2005). Though the precise enzymes involved in repair at CAG/CTG repeats still need to be elucidated, these studies indicate that repair outcome is influenced by the location of the nick, suggesting the the position of a lesion within a CAG/CTG repeat stretch may be critical with respect to repair.

Here, we investigate the hypothesis that repair outcome at CAG/CTG repeats may be influenced by the stoichiometry of BER proteins, including Lig1 and Lig3, and lesion position

within the repeats. To this end, we first determined the stoichiometry of main BER proteins in the striatum and in the cerebellum of HD transgenic mice. Second, we used a fully reconstituted repair assay, based on CAG/CTG-containing and control substrates containing an abasic site lesion, to evaluate repair outcomes at BER protein stoichiometries reflecting the situation in the striatum or in the cerebellum. In addition, we evaluated the effect of repositioning the DNA lesion within the CAG/CTG stretch.

Results:

To address those questions we first determined by Western Blot the main BER proteins - APE1, Polβ, FEN1, LigI, LigIII, XRCC1 and PCNA- molar ratio in the cerebellum and striatum of wild-type (WT) and R6/1 HD transgenic mouse model which recapitulates major features of the human pathology (Mangiarini et al., 1997). BER Proteins, in particular the LP-BER proteins LigI and FEN1, were increased in the cerebellum compared to the striatum in both HD and WT mice.

In order to examine whether the tissue-specific BER protein stoichiometry could lead to differential DNA repair on repeat substrates, we used in vitro reconstitution repair assay. To this purpose, we generated CAG/CTG- or control-substrates with an AP site lesion, which represent an obligatory intermediate during repair via BER. The AP-substrates were then incubated with BER protein cocktails reflecting the stoichiometry of the striatum or cerebellum. We used radioincorporation assays to assess repair outcomes over time. Our results show that repair efficiency, is influenced by the stoichiometry of BER proteins, the sequence of the substrate and the position of the lesion. More specifically, our data show that the production of full-length repaired products is reduced and the production of intermediate products is increased when using the striatal BER stoichiometry compared to the cerebellar BER stoichiometry, indicating that the striatal BER stoichiometry leads to poor repair efficiency. Moreover, our data show that the level of intermediate repair products resulting from the incorporation of multiple nucleotides by Polβ was increased for the CAG/CTG-substrates relative to the control substrates, indicating that LP-BER is preferentially used to process the

CAG/CTG substrates. Moreover, the data indicate that the level of Lig1 is essential in triggering efficient repair of CAG/CTG substrates. Finally, our data show that relocating the AP site lesion within the CAG/CTG-substrates influenced repair efficiency. Lesions on the CAG strand were repaired more efficiently than lesions on the CTG strand, and repositioning the lesion from the 5' to the 3' region of the repeat stretch led to more efficient repair and fewer intermediate products.

Conclusion:

We show that repair is dependent on BER protein stoichiometry, and on the DNA sequence surrounding the lesion. We propose a model where the cerebellar BER stoichiometry leads to more efficient repair than the striatal BER stoichiometry. CAG/CTG substrates are less efficiently repaired than random sequence substrates, suggesting that the CAG/CTG repeat tract impedes the repair. Globally CAG/CTG substrates are preferentially repaired through LP-BER. However, moving the lesion from the 5' to the 3' region of the CAG/CTG repeat progressively shifts the repair pathway from LP- to SN-BER. The less efficient repair of 5' located lesions on the CAG/CTG repeats could result from the higher propensity for downstream CAG/CTG repeats to adopt secondary structures, thereby impeding the repair. Similarly, the less efficient repair of lesion on the CTG strand relative to the CAG strand could result from the increased propensity for CTG sequences compared to CAG sequences to form stable hairpin structures (Hou et al., 2009; Mitas et al., 1995; Panigrahi et al., 2005). Together, our data are consistent with a model where repair via BER at CAG/CTG repeats is inhibited by DNA structural impediments, which may ultimately lead to repeat instability. We propose that the tissue-specific stoichiometry of BER proteins may exacerbate (as in the striatum) or compensate (as in the cerebellum) the intrinsic inefficiency of repair at CAG/CTG repeats.

Publication 2

(Submitted)

Nucleotide sequence, DNA damage location and protein stoichiometry influence base excision repair outcome at disease-associated CAG/CTG repeats

Agathi-Vasiliki Goula [1], Christopher E. Pearson [2,3] Julie Della Maria[4], Alan E. Tomkinson [4], David M. Wilson III [5], Karine Merienne [1*]

1 Department of Neurogenetics and Translational Medicine, Institute of Genetics and Molecular and Cellular Biology (IGBMC), UMR 7104-CNRS/INSERM/UdS, Illkirch, France, **2** Genetics and Genome Biology, The Hospital for Sick Children, TMDT Building 101 College St., 15th Floor, Room 15-312 East Tower, Toronto, ON, M5G 1L7 **3** Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada **4** Department of Radiation Oncology and the Marlene and Stewart Greenebaum Cancer Center, University of Maryland School of Medicine, Baltimore, Maryland, United States of America **5** National Institute on Aging (NIA)/ National Institutes of Health (NIH), Baltimore, Maryland, United States of America

* corresponding author:
Email: merienne@igbmc.fr;
Phone/Fax: +33 3 88 65 34 06/+33 3 88 65 32 46

Keywords: Trinucleotide repeats, instability, Huntington's disease, Base Excision repair

ABSTRACT

Expansion of CAG/CTG repeats is the underlying cause of at least fourteen genetic disorders, including Huntington's disease (HD) and a series of spinocerebellar ataxias. The mutational process is ongoing, with an increase in repeat size enhancing the toxicity of the expansion in specific tissues. In many repeat diseases the repeats exhibit high instability in the striatum, whereas instability is minimal in the cerebellum. Towards understanding human genetic diseases, we explore the hypothesis that base excision repair (BER) protein stoichiometry contributes to the tissue-selective instability of CAG/CTG repeats by using a reconstituted repair assay. These experiments specifically employed oligonucleotide substrates with an abasic site and BER protein stoichiometries that mimic the levels present in HD mouse striatum or cerebellum. The efficiency of BER at CAG/CTG repeats and at a control DNA sequence was markedly reduced under the striatal stoichiometry, likely due to the lower level of APE1, FEN1 and LIG1. Furthermore, lesions located on the CTG strand were poorly repaired in comparison to the CAG strand. Also, lesions located towards the 5' end of the repeat tract were poorly repaired accumulating incompletely repaired intermediates within the repeat tract, compared to lesions at the central or 3' end which progressed to full length products. Lesions at CAG/CTG repeats were processed by long patch-BER, particularly when 5'-located, suggesting that susceptibility to hairpin formation via strand displacement contributes to pathway selection. Together, our results suggest that BER stoichiometry, nucleotide sequence and DNA damage position modulate repair outcome, contributing to CAG/CTG repeat instability.

INTRODUCTION

Trinucleotide repeat (TNR) expansions are responsible for more than fifteen neurological, neuromuscular and neurodegenerative genetic disorders (Lopez Castel et al., 2010; Tome et al., 2011). This family of disorders includes CAG/CTG repeat-associated diseases, such as Huntington's disease (HD) and nine other polyglutamine (polyQ) disorders, as well as myotonic dystrophy type 1 (DM1). The TNR tracts need to reach a threshold length of 30 to 50 units to become genetically unstable and trigger pathogenesis. At this size, the length of the mutant allele continuously changes, in both germline and somatic tissues, most often incurring expansions. However, all tissues do not undergo repeat instability to an equal degree, and this tissue selectivity varies between diseases (Lopez Castel et al., 2010; Lopez Castel et al., 2009). For instance, in HD and other polyQ disorders, both the striatum and the cortex exhibit increased CAG/CTG instability and the largest degree of expansion, while the same repeats remain stable and shorter in length in the cerebellum (Chong et al., 1995a; Hashida et al., 2001; Kennedy et al., 2003; Lopes-Cendes et al., 1996; Shelbourne et al., 2007; Telenius et al., 1994). In DM1, continuous expansion of the mutated allele is observed in the heart, skeletal muscle and cortex, whereas TNR expansion is limited in the cerebellum (Anvret et al., 1993; Ishii et al., 1996; Lopez Castel et al., 2011; Thornton et al., 1994; Wong et al., 1995). In DM1 and HD, instability is most prevalent in affected tissues, presumably accelerating disease progression (Groh et al., 2011; Groh et al., 2002; Shelbourne et al., 2007; Swami et al., 2009). It is therefore important to delineate the molecular mechanisms of tissue-specific TNR instability to better understand the pathogenesis of the different disorders.

The mechanisms underlying the tissue selectivity of TNR instability remain unclear. Gene-specific *cis*-elements and tissue-specific epigenetic modifications have been implicated (Cleary et al., 2010). In addition, tissue-specific *trans*-factors may contribute to the process. For instance, MSH2 and

MSH3, which participate in the mismatch repair (MMR) pathway, have been implicated in CAG/CTG instability. Notably, mouse models for DM1 or HD that are deficient in *Msh2* or *Msh3* exhibit reduced CAG/CTG instability (Dragileva et al., 2009; Manley et al., 1999; Savouret et al., 2004; Tome et al., 2009; van den Broek et al., 2002; Wheeler et al., 2003). Since CAG/CTG repeats have a high propensity to form stable secondary DNA structures such as hairpins *in vitro* (Gacy et al., 1995; Pearson and Sinden, 1996), it has been hypothesized that the aberrant processing of these structures by MMR promotes instability. Interestingly, recent studies showed that downregulation of MMR genes upon differentiation of DM1-derived human embryonic stem cells correlates with decreased CAG/CTG instability (Seriola et al., 2011), suggesting that the contribution of MMR to CAG/CTG instability is regulated in a tissue-specific manner.

Somatic CAG/CTG instability is reduced in HD mice deficient for the DNA glycosylase *Ogg1*, indicating that base excision repair (BER) also contributes to CAG/CTG instability (Kovtun et al., 2007). In yeast, the flap-endonuclease (FEN1) and DNA ligase I (LIG1), two proteins involved in long-patch BER (LP-BER), modulate the instability of CAG/CTG repeats (Freudenreich et al., 1998; Subramanian et al., 2005). Human cells and DM1 transgenic mice also revealed a role for LIG1 in CAG/CTG instability (Lopez Castel et al., 2009; Tome et al., 2011). In classic BER, removal of an oxidative base lesion by a DNA glycosylase results in the formation of an abasic (AP) site, which is then cleaved by an AP endonuclease (APE1 in mammals) (Fortini and Dogliotti, 2007; Robertson et al., 2009). This DNA strand break is subsequently processed by either single-nucleotide BER (SN-BER) or LP-BER. In SN-BER, POLβ incorporates a single nucleotide and excises the remaining 5'-abasic fragment, prior to ligation by DNA ligase III α (LIG3α) in complex with x-ray cross-complementing 1 (XRCC1) protein. In LP-BER, FEN1 removes the 5'-flap structure generated during the multi-nucleotide synthesis step mediated by POLβ or a replicative DNA polymerase prior to ligation by LIG1.

We previously found that in wild-type and HD transgenic mice, BER proteins and associated enzymatic activities are reduced in the striatum in comparison to the cerebellum, though to different levels (Goula et al., 2009). In particular, FEN1 protein is greatly decreased in the striatum, leading to a lower FEN1:POLβ ratio compared to the cerebellum. We therefore proposed that due to impaired coordination between DNA synthesis and 5'-flap removal in the striatum, LP-BER at CAG/CTG repeats would contribute to the high level of TNR instability seen in the striatum of HD animals. This hypothesis is consistent with a previous study, which used *in vitro* repair assays with oligonucleotide substrates containing a tetrahydrofuran (THF) abasic site analog that can only be processed by LP-BER, showing that disruption of POLβ and FEN1 coordination results in CAG repeat expansion (Liu et al., 2009). Two distinct models could explain these results. In one model, both SN- and LP-BER are competent to process DNA damage at CAG/CTG repeats, but BER stoichiometry promotes CAG/CTG instability by influencing BER subpathway selection (SN- *versus* LP-BER). The second model proposes that only LP-BER would be competent to process DNA damage at CAG/CTG repeats, and that the relative stoichiometry of LP-BER proteins defines the risk of CAG/CTG instability by modulating repair outcome.

Repair outcomes at nicked CAG/CTG substrates with slipped-out repeats have been reported using mammalian cell extracts (Hou et al., 2009; Hou et al., 2011; Panigrahi et al., 2005; Panigrahi et al., 2010). In these assays, repair outcome and efficiencies clearly depended upon nick location and slip-out sequence (CAG *versus* CTG). Interestingly, repair efficiency was significantly increased when the slip-out was located on the CAG strand in comparison to the CTG strand. Furthermore, the nick location (in the slipped- *versus* continuous-strand or 5' *versus* 3' of the slip-out) dramatically affected repair outcome. Whether the position of an oxidative DNA lesion within a CAG/CTG repeat sequence also influences repair is unknown. In all reports examining BER processing of trinucleotide repeats to date, the DNA lesion has been placed in the CAG strand at the 5' end (the first repeat unit) (Kovtun et al., 2007; Liu et al., 2009) or within a CAG hairpin (Jarem et al., 2009; Jarem et al., 2011). Here we use a reconstituted

repair assay that employs oligonucleotide substrates harboring an abasic lesion positioned upstream (5'-oriented), downstream (3'-oriented), or centrally located within the CAG or CTG strand. Repair assays were carried out with BER protein mixtures that reiterate the stoichiometry in the striatum or cerebellum of HD mice. Our data demonstrate that nucleotide sequence, DNA damage location and the stoichiometry of BER factors influence repair at disease-associated CAG/CTG repeats. In addition, our studies support a model in which LP-BER is required to process DNA damage at CAG/CTG repeats, and BER protein stoichiometry modulates repair efficacy.

MATERIALS AND METHODS

Materials

The human recombinant proteins LIG1, LIG3/XRCC1, PCNA, APE1, POLβ and FEN1 were purified as previously described (Chen et al., 2006; Della-Maria et al., 2011; Erzberger et al., 1998; Lee and Wilson, 1999; Levin et al., 1997; Nguyen et al., 2000). Uracil-DNA glycosylase (UNG) from *E.coli* was a generous gift from the late Dale Mosbaugh (Oregon State University, USA). Primary rabbit and mouse antibodies were purchased from MBL (mouse α-LIG1, K 0190-3), BD transduction Labs (mouse α-LIG3, 611876), Abcam (rabbit α-APE1, ab92744), Sigma (mouse α-PCNA, P 8825) and Chemicon (mouse α–β-tubulin). The rabbit α-XRCC1 was a kind gift from P.J. McKinnon (St. Jude Children's Research Hospital, USA). DNA oligonucleotides were purchased from Eurogentec. The radionucleotides, [γ-^{32}P] ATP (7000μCi/mmol), [α-^{32}P]-dCTP and [α-^{32}P]-dGTP (3000μCi/mmol), were from Perkin Elmer.

Mice

Hemizygous R6/1 HD transgenic mice from the Jackson Laboratory were maintained on a mixed CBAxC57BL/6 genetic background, and were genotyped as described (Mangiarini et al., 1996a). The experiments were approved by the ethical committee C.R.E.M.E.A.S (Comite Regional d'Ethique en Matiere d'Experimentation Animale de Strasbourg).

Western blotting

Whole cell extracts from mouse cerebellum and striatum of R6/1 mice were prepared as previously described (Goula et al., 2009). For western blot analysis, rabbit and mouse antibodies against mouse endogenous and human recombinant BER proteins (see above) were used at 1:500 (α-LIG1), 1:1000 (α-LIG3, α-XRCC1, α-PCNA), or 1: 100 000 (α-APE1) dilutions, and were detected with appropriate α-rabbit or α-mouse peroxidase-conjugated secondary antibodies (Jackson immunoResearch Laboratories) and the ECL chemiluminescence kit (Pierce or Millipore). Signals were imaged on radiographic films, captured with GeneSnap and quantified with GeneTool softwares on Syngene Chemigenus XE machine.

Preparation of AP-substrates

To create the AP-DNA duplex substrates, the target oligonucleotide strand harbored a uracil (U) modification that replaced a cytosine or a guanine residue, and this strand was hybridized with the complementary oligonucleotide at a molar ratio of 1:1 (see Table 1). We note that the sequences of the CAG/CTG substrates and the control substrates were based upon the sequences described by Liu et al. (Liu et al., 2009) and Petermann et al. (Petermann et al., 2003), respectively. Subsequently, the optimal conditions for complete removal of uracil from the various DNA substrates by uracil-DNA glycosylase (UNG) were determined (Fig. S2). In brief, 0.5 pmol of double strand substrate was incubated with 2.9 units of UNG in 75 mM KCl, 25 mM MgCl2, 3.125 mM HEPES–KOH, pH 7.7, 1% glycerol, and 0.25 mM EDTA for one hour at 37°C, and the reaction was stored on ice for 10 min until further needed. For 5' [γ-^{32}P]-labeled experiments, U-containing substrates were 5'-radiolabeled using [γ-^{32}P]ATP and T4 polynucleotide kinase (New England Biolabs) for 40 min at 37°C , prior to annealing with the template strand.

Reconstitution repair assays

The AP-DNA substrates were incubated with a mixture of BER proteins as described throughout. The repair reactions were performed by incubating 0.5 pmol of AP-DNA substrate with 0.1 nmol of POLβ, 0.5 nmol of APE1, 0.75 nmol of FEN1, 0.1 nmol of PCNA, 0.05 nmol of LIG1, and 0.75 nmol each of LIG3 and XRCC1 (striatal stoichiometry), or 0.1 nmol of POLβ, 1.5 nmol of APE1, 0.6 nmol of FEN1, 0.2 nmol of PCNA, 0.2 nmol of LIG1, and 0.2 nmol each of LIG3 and XRCC1 (cerebellar stoichiometry). The reactions were carried out at 37°C for the indicated period of time in 75 mM KCl, 25 mM MgCl2, 3.125 mM HEPES–KOH, pH 7.7, 1% glycerol, 0.25 mM EDTA, as previously described (Goula et al., 2009). For radioincorporation repair assays, the reaction buffer was supplemented with 20 μM of dNTPs and 125 nM of [α-^{32}P]-dCTP or [α-^{32}P]-dGTP, depending on the substituted base (a cytosine or guanine). For 5' [γ-^{32}P]-labeled experiments, the reaction buffer was supplemented with 20 μM dNTPs only. The reactions were stopped by addition of stop buffer (98% formamide, 20 mM EDTA, bromophenol blue and xylene cyanol) and heating at 95 °C for 5 min. The reaction products were resolved by 15% polyacrylamide urea denaturing gel electrophoresis, imaged on a Typhoon phosphoimager, and quantified with the ImageQuant TL software.

RESULTS

LIG1 and FEN1 protein levels are increased in the cerebellum compared to the striatum

To determine the potential role of BER protein stoichiometry in the tissue-selective instability of CAG/CTG repeats in HD, the levels of key BER proteins, APE1, LIG1, LIG3, XRCC1, and PCNA, were determined in the cerebellum and striatum of HD transgenic mice by Western blot analysis (Fig. 1a). β-tubulin was used as a loading control (Fig. 1a). The molar levels of POLβ, the main BER polymerase in the brain (Rao et al., 2001; Wei and Englander, 2008), and FEN1 were determined previously ((Goula et al., 2009) and data not shown). The amount of each BER protein (see above) was measured relative to a

known quantity of the corresponding purified recombinant protein, thereby permitting determination of the molar ratio of BER proteins in mouse striatum and cerebellum (Fig. 1b). BER protein levels were globally higher in the cerebellum in comparison to the striatum of HD mice (Fig. 1), with wild-type and HD mice exhibiting similar tissue-specific protein ratios (data not shown). LIG1, FEN1, and APE1, which showed the greatest difference in expression between tissues, were 5-fold, 4-fold and 3-fold more elevated in the cerebellum relative to the striatum (Fig. 1B and (Goula et al., 2009)). XRCC1 and PCNA were increased by ~2-fold in the cerebellum, while POLβ and LIG3 levels were similar between the two tissues (Fig. 1 and (Goula et al., 2009)). Thus, the concentration of several key BER proteins was different between the striatum and cerebellum.

Tissue-specific BER stoichiometry affects repair of AP site-containing CAG repeat substrates

Tissue-specific differences in BER protein stoichiometries could lead to different DNA repair outcomes on CAG repeat substrates. To test this hypothesis, we used a CAG-containing oligonucleotide duplex (D1D1c) and, as a control, a random sequence oligonucleotide duplex (U16G16), which were treated with the monofunctional DNA glycosylase UNG to excise a strategically positioned uracil residue and generate an AP site (Table 1; see Materials and Methods). We introduced a natural AP site, as opposed to the synthetic abasic site analog THF, to allow for processing by either SN-BER or LP-BER. These AP-DNAs were then incubated with a mixture of recombinant BER proteins, which included APE1, POLβ, FEN1, PCNA, the XRCC1/LIG3α complex, and LIG1, at either the cerebellum- or striatum-specific stoichiometry. Repair was monitored by radionucleotide incorporation over time.

Repair of the CAG (D1D1c) and the random sequence (U16G16) substrates yielded full-length repair products, as well as BER intermediate products that involved incorporation of one or several (up to +9nt) nucleotides (Fig. 2a and 2b). The production of full-length repair product,

an indication of complete repair efficiency, was quantified over time in the presence or absence of ATP (Fig. 2a and 2b, right panels). Since a significant fraction of the protein molecules in the purified preparations of LIG1 and LIG3 were pre-adenylated, full-length repair products were detected in a somewhat time-dependent manner in the absence of ATP. Supplementation with ATP, which enables the DNA ligases to catalyze more than one ligation event, substantially increased repair efficiency over time under both tissue-specific BER stoichiometries, indicating that ligation is a major determinant of repair progression.

Interestingly, in the absence of ATP, the production of full-length repair product was lowest under the striatal BER protein ratios, and was inversely correlated with the level of intermediate products (Fig. 2a and 2b). A similar effect was observed for both D1D1c and U16G16, indicating that the striatal stoichiometry results in less efficient repair irrespective of the DNA substrate. In addition, repair of the CAG substrate (D1D1c) under the striatal or cerebellar stoichiometry involved multi-nucleotide synthesis as revealed by the presence of ≥2nt intermediate products, indicating that the AP lesion was preferentially processed by LP-BER (Fig. 2a). In contrast, repair of the AP lesion in the random sequence (U16G16) led predominantly to synthesis of +1nt intermediate products, indicative of SN-BER processing (Fig. 2b). Furthermore, during repair of the CAG D1D1c substrate, both the amount of +n (+>1) products and the number of nucleotides incorporated (up to +9nt) was greatest under the striatal conditions; though the level of ≥3nt remained low, and the relative intensities of these intermediates mildly increased over time, suggesting stalled repair progression (Fig. 2a). In contrast, under the cerebellar BER protein ratio, the intensity of the +1nt intermediate product was progressively converted to +2, +3, +4, +5nt products over time and ultimately to full-length repair product, indicative of more efficient repair. Moreover, repair of 5'-labeled CAG and

control DNA substrates agreed with the radioincorporation experiments, indicating that more intermediate products and fewer full-length repair products were generated with the CAG substrate under the striatal BER protein stoichiometry (Sup. Fig. 1). Taken together, the results suggest that repair efficiency of an AP lesion is sensitive to the stoichiometry of BER proteins, while the DNA sequence influences BER subpathway selection, where CAG repeats follow a LP-BER and non-repetitive sequences follow a SN-BER path.

Position of the AP site within CAG- or CTG-substrates influences repair outcome

The CAG substrate was processed primarily by LP-BER (see above), suggesting that the propensity for forming hairpin structures influences subpathway choice. The stability of the hairpin structures formed at CAG/CTG repeats is greater on the CTG strand than on the CAG strand and increases with repeat length (Gacy et al., 1995; Hou et al., 2009; Panigrahi et al., 2005; Pearson and Sinden, 1996). Therefore, we reasoned that changing the location of the lesion between repeat strands might influence repair outcome. Similarly the ability to form a hairpin during repair may be influenced by the polarized location of the lesion. To test these hypotheses, we synthesized a series of DNA substrates designed to harbor an AP site that (i) replaces a *cytosine* (D1D1c, M1M1c) or a *guanine* (M2M2c, M5M5c) near the 5' end of a CTG or CAG repeat sequence (5'-oriented), (ii) is embedded within the middle of the repeat stretch (M3M2c, M6M5c), or (iii) is located towards the 3' end of the repeats (3'-oriented: M4M2c, M7M5c) (Table 1 and Fig. 3a). Two AP site-containing random sequence substrates were used as controls (U16G16, U14G14).

All DNA substrates tested yielded intermediate and full-length repair products following incubation with cerebellar or striatal BER protein stoichiometries (Fig. 3a). However, the repair

efficiency and the intermediate product pattern were different depending on the substrate sequence, the position of the lesion, and the BER protein ratio. The cerebellar protein stoichiometry led to more efficient repair for all substrates compared to the striatal conditions, with the ratio of full-length repair products to intermediate products higher in the cerebellar conditions (Fig. 3b). This was consistent with the results of Fig. 2. Placing the AP site within the 3' portion of the repeat tract led to more efficient repair and shorter intermediate products, indicative of preferential SN-BER, in comparison to when the lesion was placed towards the 5' end, which leads to an accumulation of longer +n products (Fig. 3a and 3b). The efficiency and pattern of repair were comparable between the 3'-oriented substrates and the control substrates (Fig. 3a and 3b). This polar position effect was seen with both the CAG and CTG strand and under both tissue-specific stoichiometries, suggesting more efficient repair when the lesion is 3'-oriented than 5'-oriented, likely due to the decreased propensity of the 3' lesions to allow the formation of a stable hairpin structure. Finally, although the CAG- and CTG-substrates showed a similar repair pattern when the damage was located in the same position (*i.e.* 5', middle, or 3'), substrates with lesions in the CAG strand exhibited higher repair efficiency than lesions in the CTG strand [compare D1D1c, D2D2c, M2M2c, M3M2c and M4M2c (CAG substrates) with M1M1c, M5M5c, M6M5c and M7M5c (CTG substrates), Fig. 3b]. Taken together, our results show that both the strand (CAG *versus* CTG) and the location of the lesion within the repeat tract affect repair outcome, supporting the view that the propensity to form a hairpin structure influences BER subpathway selection and repair efficiency.

LIG1 contributes to the differential repair efficiency of cerebellar and striatal BER stoichiometries

A key difference between the BER protein stoichiometry in the mouse striatum and cerebellum is the lower levels of LIG1 in the striatum, while LIG3 is similar in the two tissues (Fig. 1). In particular, LIG1 is

5-fold lower in the striatum relative to the cerebellum. This suggested that LIG1 levels may contribute to the differential repair between the striatal and cerebellar BER protein ratios. To examine this hypothesis, we incubated the AP site-containing CAG substrate (D1D1C) with a BER protein cocktail consisting of various ligase protein combinations (Fig 4a). AP-DNA repair was then assessed over time. Repair under the strict cerebellar (cb, lane 1 throughout) or striatal (st, lane 3 throughout) conditions were consistent with the results seen above (Fig. 2 and Fig. 3). Interestingly, repair efficiency was highest at the cerebellar ligase stoichiometry, regardless of the other BER protein concentrations (cb, lanes 1 and 4). The presence of intermediate products was generally more pronounced under conditions mimicking the striatal situation for ligases (st, lanes 2 and 3). Most strikingly, reducing the ligase levels to that of the striatum, where all other BER proteins were at the cerebellar ratios led to a markedly increased intensity and patch-length for the intermediate products (Fig. 4a, lane 2 throughout). Similar results were obtained using the U16G16 control substrate, except that the difference in repair efficiency for the various BER protein cocktails was not as pronounced when changing the ligase concentration (Fig. 4b). The enhanced stalling of repair progression with the reduced ligase levels of the striatum confirmed the suspicion that limited ligase activity in the absence of ATP, may limit repair progression (Fig. 2).

The LIG1:LIG3 ratio and the total amount of DNA ligase are both increased in the cerebellum environment in comparison to the striatum (Fig. 1). LIG1:LIG3 ratio is 4 in the cerebellum and 1 in the striatum, and the total ligase amount is 2-fold higher in the cerebellum than the striatum. To clarify whether the total amount of ligase protein or the relative ratio of the two ligases determines repair efficacy, we incubated the AP site-containing CAG-substrate (D1D1C) with different LIG1:LIG3 molar ratios, while maintaining the other BER proteins at either the cerebellar or striatum stoichiometry (Fig. 4c). Repair efficiency was influenced mainly by the LIG1 levels, whereas the amount of LIG3 had a lesser

effect on repair outcome. These results indicate that the lower LIG1 level primarily contributes to the reduced repair efficiency seen with the striatal protein stoichiometry.

DISCUSSION

Several studies support a role for BER in disease-associated CAG/CTG instability; yet the underlying mechanisms remain elusive. It has been hypothesized that LP-BER is specifically involved in TNR instability. However, it is unclear whether LP-BER is necessary for processing of a lesion at CAG/CTG repeats and contributes to the tissue selectivity of TNR instability. The results herein demonstrate that tissue-specific BER stoichiometry affects DNA repair efficiency, while nucleotide sequence (*e.g.* CAG *versus* CTG repeats) and DNA damage location along the repeat tract (5' *versus* 3') influence LP-BER choice. In a reconstituted system, repair of AP substrates was less efficient when the BER protein stoichiometry reflected that of HD mouse striatum, the tissue showing the highest CAG/CTG instability, as compared to the stoichiometry of the cerebellum, a tissue exhibiting minimal TNR instability. This reduced repair was observed irrespective of the substrate (*e.g.* CAG/CTG or control substrate); however, repair of the CAG/CTG substrates primarily involved LP-BER, while control substrates were predominantly processed by SN-BER. Furthermore, repair of CAG or CTG substrates containing a 5'-located lesion was less efficient than 3'-located lesions and involved mainly LP-BER, presumably due to the increased propensity to form a hairpin structure. Finally, repair was less efficient when the lesion was located on the CTG strand relative to the CAG strand, likely due to the increased stability of CTG hairpins. We suggest that inefficient BER progression at CAG/CTG repeats, resulting from both the tissue-specific BER protein stoichiometry and the position of the lesion within the repeat tract, contributes to TNR instability.

BER is central to the maintenance of genomic stability. The biochemistry of BER, including the nature, kinetics and coordination of the different enzymes, has been extensively studied (Wilson et al.,

2010). However, how BER is regulated at a cell- or tissue-specific level throughout life and how this impacts on disease remains largely unknown. Studies have shown that the pattern of expression of BER genes is regulated throughout development and post-natal life in a tissue-specific manner (Wilson and McNeill, 2007). Furthermore, tissue- and age-dependent variations of BER protein levels and activities have been reported (Intano et al., 2003; Karahalil et al., 2002; Szczesny et al., 2010; Wilson and McNeill, 2007). For example, the levels of different DNA glycosylases, including OGG1, were assessed in different mouse tissues, and revealed variations between tissues (Karahalil et al., 2002). In addition, the proportional abundance of other BER proteins, such as APE1, POLβ, XRCC1, LIG1 and LIG3, was found to differ between the mouse liver and brain (Intano et al., 2003). BER activity was also impaired in terminally differentiated cells (Narciso et al., 2007), and was significantly lower in skeletal muscle when compared to liver or kidney (Szczesny et al., 2010). We determined the molar ratio of several major BER proteins, including APE1, POLβ, XRCC1, LIG1, LIG3 and PCNA, in mouse striatum and cerebellum, and found a significant difference between brain regions, consistent with the idea that BER activity is highly tissue-specific.

Several BER proteins have previously been implicated in regulating the instability of CAG/CTG repeats. For instance, the DNA glycosylase OGG1, an enzyme that initiates BER, is necessary to increase somatic instability of CAG/CTG repeats in HD mice (Kovtun et al., 2007), although excision of an 8-oxoG lesion within a CAG hairpin is rate-limiting (Goula et al., 2009; Jarem et al., 2009; Jarem et al., 2011). In addition, yeast studies support a role for FEN1 and LIG1 in the instability of CAG/CTG tracts (Freudenreich et al., 1998; Lopez Castel et al., 2009; Subramanian et al., 2005), and LIG1 is implicated in human cell and transgenic mouse studies (Lopez Castel et al., 2009). Furthermore, FEN1 cleaves 5' flap-bearing structures formed by CTG repeats less efficiently than unstructured flaps (Vallur and Maizels, 2010) and is inhibited by the secondary structures formed at these repeats in a length-dependent manner (Spiro et al., 1999). Lastly, reconstitution experiments indicate that coordination between POLβ

and FEN1 modulates CAG/CTG repeat expansion during LP-BER (Goula et al., 2009; Liu et al., 2009). These results suggest that lesions at CAG/CTG repeats are poorly processed by BER, likely due to structural impediments, and optimal coordination of the BER enzymatic steps is essential to insure correct repair and prevent TNR instability. The results herein support this view, showing that repair outcome at CAG/CTG repeats is dependent upon the stoichiometry of BER proteins, which dictates coordination of repair, and the location of the lesion, which influences secondary structure formation. Specifically, repair of CAG/CTG substrates was inefficient when using the BER protein stoichiometry reflecting the situation in the striatum, in comparison with the cerebellar stoichiometry (Fig. 2 and 3). The low repair efficiency under the striatal BER stoichiometries correlated with reduced APE1, FEN1 and LIG1 relative to POLβ, as compared to the ratio in the cerebellum, suggesting that poor coordination of DNA synthesis by POLβ with the upstream and/or downstream BER enzymatic steps led to inefficient repair. In accordance, decreasing the level of DNA ligase, and in particular of LIG1, under the cerebellar stoichiometry for other BER proteins led to less efficient repair of CAG/CTG substrates and a concomitant accumulation of stalled repair intermediates (Fig. 4), consistent with previous studies showing that LIG1 controls repair patch length (Levin et al., 2000; Lopez Castel et al., 2009; Pascucci et al., 1999). In HD and other CAG/CTG repeat-associated diseases, somatic instability is elevated in the striatum and minimal in the cerebellum, suggesting that CAG/CTG instability correlates with inefficient BER progression.

FEN1 and LIG1 have been implicated in CAG/CTG instability, suggesting that LP-BER is specifically required for processing of oxidative DNA damage at CAG/CTG repeats (Freudenreich et al., 1998; Goula et al., 2009; Liu et al., 2009; Lopez Castel et al., 2009; Subramanian et al., 2005). Our data indicate that repair of CAG/CTG substrates is strongly dependent upon LIG1 and not LIG3 (Fig. 4). In addition, our results support the view that an AP lesion located in a CAG/CTG tract is preferentially processed by LP-BER, as revealed by the formation of ≥2nt intermediate products, whereas a lesion in a random DNA

sequence is predominantly repaired via SN-BER (Fig. 2). LP-BER was executed on CAG/CTG substrates under both the striatal and cerebellar BER protein stoichiometries, despite the fact that the levels of FEN1 and LIG1 varied greatly (relative to POLβ) between the two situations. These results indicate that selection of LP-BER at CAG/CTG repeats is strongly dependent upon the DNA sequence and that BER stoichiometry plays a minor role. Furthermore, a 5'–located AP site resulted in the production of ≥2nt intermediate products, whereas +1nt intermediate products predominated when the lesion was positioned in the 3'-location (Fig. 3). This finding suggests that the involvement of LP-BER at CAG/CTG repeats increases when the AP site is located 5' within the CAG or CTG repeat tract, presumably due to the increased propensity to form a hairpin structure via strand displacement toward the 3' end (Hartenstine et al., 2002). In addition, the production of longer +n intermediates was associated with a decrease in repair efficiency, as less full-length repair product was observed with lesions located at the 5' end of the repeat tract (Fig. 3). Together, these results indicate that LP-BER operates more frequently when processing a lesion within a repetitive DNA sequence prone to forming a structural impediment, such as a hairpin.

SN-BER plays a prominent role in cells (Sobol et al., 1996), yet the factors contributing to LP-BER selection *in vivo* remain elusive. Several studies have provided evidence that LP-BER is functional in vertebrate cells, including brain cells where it is catalyzed by POLβ (Asagoshi et al., 2010a; Asagoshi et al., 2010b; Wei and Englander, 2008). It has been reported that reduction or oxidation of AP sites, as well as ATP cellular concentration control BER subpathway selection (Klungland and Lindahl, 1997; Petermann et al., 2003). We show herein that the DNA sequence that surrounds the lesion is critical for LP-BER selection. Together, our results indicate that BER outcome at CAG/CTG repeats is influenced by both the propensity to form stable secondary structures (determining pathway selection) and protein stoichiometry (modulating repair efficacy). We propose that tissue-specific stoichiometry of BER

enzymes may exacerbate (as in the striatum) or compensate for (as in the cerebellum) the intrinsic difficulty in repairing a lesion at CAG/CTG repeats, thereby modulating the risk of CAG/CTG instability.

FIGURE LEGENDS

Fig. 1. BER protein stoichiometry in the cerebellum and striatum of HD mice. (a) Steady state levels of XRCC1, LIG1, LIG3, APE1 and PCNA proteins in the striatum and cerebellum of HD mice were determined by Western blot analysis, using purified human recombinant proteins as the reference. 100 μg of whole cell extracts prepared from the striatum or the cerebellum of the same HD mice were run on an SDS-polyacrylamide gel (3 mice were used per gel) with 10 ng of recombinant XRCC1, LIG1, LIG3, APE1 or PCNA and detected with α-XRCC1, α-LIG1, α-LIG3, α-APE1 and α-PCNA antibodies. α–β-tubulin antibody was used as a loading control. Representative images are shown. Band intensities were quantified relative to corresponding recombinant protein. The levels of FEN1 and POLβ in HD mouse striatum (St) and cerebellum (Cb) were previously determined (Goula et al., 2009), thereby allowing determination of the XRCC1:LIG1:LIG3:APE1:PCNA:FEN1:POLβ molar ratios in the two tissues. **(b)** The stoichiometry of XRCC1:LIG1:LIG3:APE1:PCNA:FEN1:POLβ in the striatum and cerebellum of HD mice. The level of POLβ, which is similar in the two tissues of HD mice (Goula et al., 2009), was set as 1, and the molar ratio of the other BER proteins was calculated relative to POLβ.

Table 1. Sequence of oligonucleotide substrates. The name, size and sequence of the target oligonucleotide are shown, as well as the location of the AP site within the sequences (designated as U, which is the uracil residue excised by UNG treatment). The complementary (template) oligonucleotides used for double-stranded DNA preparation are also listed.

Fig. 2. Repair of DNA substrates is modulated by BER stoichiometry. (a) Left panel. Radioincorporation experiment showing the time course of repair of the CAG-substrate D1D1C under striatal (st) and cerebellar (cb) repair conditions, with or without 1 mM ATP supplementation**.** A representative polyacrylamide denaturing gel of the repair reaction is shown. The full-length repaired products (FL) are indicated with bold arrows (➡), and +1, +2 and +3nt intermediate products are shown. Control

reactions include the 5'-labeled U-containing oligonucleotide substrate alone (lane 4), or treated with the cerebellar BER protein ratio in absence of UNG (lane 3), with the striatal BER protein ratio in absence of UNG (lane 2), or with UNG and APE1 only (lane 1). Right panel. Graph representing repair efficiencies of the CAG D1D1C susbstrate. Relative repair efficiencies correspond to the ratios: (full-length repair product at a specific time point and BER stoichiometry) ÷ (full-length repair product at 40 min). Black bars: repair efficiency for the cerebellar stoichiometry, gray bars: repair efficiency for the striatal stoichiometry. **(b)** Left panel. Radioincorporation experiment showing the time course of repair of the control U16G16 substrate under striatal (st) and cerebellar (cb) repair conditions, with or without 1 mM ATP supplementation. The full-length repaired products (FL) are indicated with bold arrows (➡) and +1nt intermediate products are shown. Right panel. Graph representing repair efficiencies of the control U16G16 susbstrate.

Fig. 3. Lesion position within the CAG/CTG repeat stretch modulates repair. (a) Radioincorporation experiment showing repair of CAG or CTG substrates with an AP site located at various positions within the repeat sequence. The scheme above represents the substrates before BER. The CAG and CTG strands are shown in blue and red, respectively. The dark square represents the AP lesion, and the C or G letters indicate that the AP lesion replaces a cytosine or a guanine, respectively. Control substrates are denoted CT1 and CT2. Reactions were performed under striatal (st) or cerebellar (cb) BER protein stoichiometries for 40 min without ATP supplementation. Bold arrows (➡) indicate the full-length repaired product; asterisks (*) indicate the intermediate products; the nature of the modified base is indicated. Bottom. Schematic representation of the different substrates during BER. The CAG and CTG strands are shown in blue and red, respectively. The dark circles represent DNA synthesis by POLβ and the arrows show the directionality of synthesis; Long CAG or CTG tracts form more stable hairpins than shorter tracts; CTG tracts form more stable hairpins than CAG tracts. **(b)** Top panel. Graph representing the relative levels of intermediate products (black bars) and full-length repaired (grey bars) products

under the striatal BER stoichiometry. For a given substrate, the level of intermediate products correspond to the sum of all the intermediate products (*i.e.* +1, +2 and +3 nt products, when detected). The level of the full-length repaired products corresponding to M4M2C substrate under the striatal BER stoichiometry was arbitrarily set to 1. The levels of the full-length and intermediate products for all substrates were calculated relative to M4M2C full-length product. Middle panel. Graph representing the relative levels of intermediate (+1, +2, +3 nt; black bars) and full-length repaired (grey bars) products under the cerebellar BER stoichiometry. The level of the full-length repaired products measured for M4M2C substrate under the cerebellar stoichiometry was arbitrarily set to 1. Bottom panel. Graph representing repair index for all the substrates. The repair index corresponds to the ratio: (full-length repair product) ÷ (sum of all the intermediate products). Black bars: repair index for the cerebellar stoichiometry, grey bars: repair index for the striatal stoichiometry.

Fig. 4. LIG1 concentration is critical in modulating repair of CAG substrates. **(a, b)** Radioincorporation experiments showing the time course of repair of CAG or control substrates when changing the stoichiometries of LIG1 and LIG3. The substrates were incubated with a mixture of BER proteins reflecting the stoichiometry in the cerebellum of HD mice, except that the concentrations of LIG1 and LIG3 were that measured in the striatum (lanes 2), or, conversely, with a mixture of BER proteins reflecting the stoichiometry in the striatum, except that the concentrations of LIG1 and LIG3 were that found in the cerebellum (lanes 4). As controls, the substrates were also incubated with mixtures of BER proteins reflecting the levels in the cerebellum (lanes 1) and striatum (lanes 3). **(c)** Radioincorporation experiment showing repair of the CAG substrate D1D1C when removing either LIG1 or LIG3 from the reaction. The substrates were incubated with a mix of BER proteins according to the stoichiometry in the striatum or cerebellum, except that the LIG1:LIG3 molar ratio was varied as follows: 2:0; 4:0; 0:2 and 0:4. As controls, the striatal ratio (0.5:1.5) and the cerebellar ratio (2:2) were included. Full-length (FL) repair

products are indicated with bold arrows (➡). The experiments were performed without ATP supplementation.

Supplementary fig. 1. Reconstitution repair assay using 5'-labeled CAG substrate. Time course of repair of the CAG substrate D1D1C under cerebellar or striatal stoichiometries using 5'-radiolabeled oligonucleotide substrate. The full-length substrates and repair products are indicated with bold arrows (➡). For control reactions, the substrate was incubated alone (lane 4), or with BER proteins but without UNG (lane 2 and 3), or with UNG and APE1 only (lane 1).

Supplementary fig. 2: Removal of the uracil by UNG. For optimization of UNG treatment, 5' [γ-^{32}P]-labeled U-substrates (U16G16, D1D1C and D2D2C) were incubated with various amounts of UNG and 1 ng of human recombinant APE1 protein. The endonucleolytic cleavage product was then monitored on a polyacrylamide urea denaturing gel. Complete removal of uracil from the various DNA substrates was achieved using ≥2,9 units of UNG (corresponding to lane 4 for each substrate). Arrows show the full-length substrate; arrowheads show the incision product.

ACKNOWLEDGMENTS

We thank Y. Trottier and F. Klein for discussions, C. Weber for technical assistance, and J.L. Mandel and H. Puccio for constant support. We are grateful to P.J. McKinnon (St. Jude Children's Research Hospital, USA) for the kind gift of the antibody to XRCC1. This research was supported by the Centre National de la Recherche Scientifique (CNRS), the Institut National de la Santé et de le Recherche Médicale (INSERM), and the University of Strasbourg (KM), by the Intramural Research Program of the NIH, National Institute on Aging (DMWIII), by the Muscular Dystrophy Association, the Canadian Institutes of Health Research (MOP-94966) and the Paul Wellstone Muscular Dystrophy Cooperative Research Center (CEP), and by Grants from the National Institutes of Health (U54NS48843 to CEP and GM57479 and ES012512 to AET). A.G. was supported by the French Ministery of Research and the Fondation de la Recherche Medicale (FRM).

The authors declare that they have no conflict of interest

REFERENCES

1. Lopez Castel, A., Cleary, J. D., and Pearson, C. E. (2010) Repeat instability as the basis for human diseases and as a potential target for therapy, *Nat Rev Mol Cell Biol 11*, 165-170.
2. Tome, S., Panigrahi, G. B., Lopez Castel, A., Foiry, L., Melton, D. W., Gourdon, G., and Pearson, C. E. (2011) Maternal germline-specific effect of DNA ligase I on CTG/CAG instability, *Hum Mol Genet 20*, 2131-2143.
3. Lopez Castel, A., Tomkinson, A. E., and Pearson, C. E. (2009) CTG/CAG repeat instability is modulated by the levels of human DNA ligase i and its interaction with proliferating cell nuclear antigen: a distinction between replication and slipped-DNA repair, *J Biol Chem 284*, 26631-26645.
4. Telenius, H., Kremer, B., Goldberg, Y. P., Theilmann, J., Andrew, S. E., Zeisler, J., Adam, S., Greenberg, C., Ives, E. J., Clarke, L. A., and et al. (1994) Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm, *Nat Genet 6*, 409-414.
5. Shelbourne, P. F., Keller-McGandy, C., Bi, W. L., Yoon, S. R., Dubeau, L., Veitch, N. J., Vonsattel, J. P., Wexler, N. S., Arnheim, N., and Augood, S. J. (2007) Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain, *Hum Mol Genet 16*, 1133-1142.
6. Chong, S. S., McCall, A. E., Cota, J., Subramony, S. H., Orr, H. T., Hughes, M. R., and Zoghbi, H. Y. (1995) Gametic and somatic tissue-specific heterogeneity of the expanded SCA1 CAG repeat in spinocerebellar ataxia type 1, *Nature Genet. 10*, 344-350.
7. Hashida, H., Goto, J., Suzuki, T., Jeong, S., Masuda, N., Ooie, T., Tachiiri, Y., Tsuchiya, H., and Kanazawa, I. (2001) Single cell analysis of CAG repeat in brains of dentatorubral-pallidoluysian atrophy (DRPLA), *J Neurol Sci 190*, 87-93.
8. Lopes-Cendes, I., Maciel, P., Kish, S., Gaspar, C., Robitaille, Y., Clark, H. B., Koeppen, A. H., Nance, M., Schut, L., Silveira, I., Coutinho, P., Sequeiros, J., and Rouleau, G. A. (1996) Somatic mosaicism in the central nervous system in spinocerebellar ataxia type 1 and Machado-Joseph disease, *Ann Neurol 40*, 199-206.
9. Kennedy, L., Evans, E., Chen, C. M., Craven, L., Detloff, P. J., Ennis, M., and Shelbourne, P. F. (2003) Dramatic tissue-specific mutation length increases are an early molecular event in Huntington disease pathogenesis, *Hum Mol Genet 12*, 3359-3367.
10. Ishii, S., Nishio, T., Sunohara, N., Yoshihara, T., Takemura, K., Hikiji, K., Tsujino, S., and Sakuragawa, N. (1996) Small increase in triplet repeat length of cerebellum from patients with myotonic dystrophy, *Hum Genet 98*, 138-140.
11. Wong, L. J., Ashizawa, T., Monckton, D. G., Caskey, C. T., and Richards, C. S. (1995) Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent, *Am J Hum Genet 56*, 114-122.
12. Thornton, C. A., Johnson, K., and Moxley, R. T., 3rd. (1994) Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes, *Ann Neurol 35*, 104-107.
13. Anvret, M., Ahlberg, G., Grandell, U., Hedberg, B., Johnson, K., and Edstrom, L. (1993) Larger expansions of the CTG repeat in muscle compared to lymphocytes from patients with myotonic dystrophy, *Hum Mol Genet 2*, 1397-1400.
14. Lopez Castel, A., Nakamori, M., Tome, S., Chitayat, D., Gourdon, G., Thornton, C. A., and Pearson, C. E. (2011) Expanded CTG repeat demarcates a boundary for abnormal CpG methylation in myotonic dystrophy patient tissues, *Hum Mol Genet 20*, 1-15.

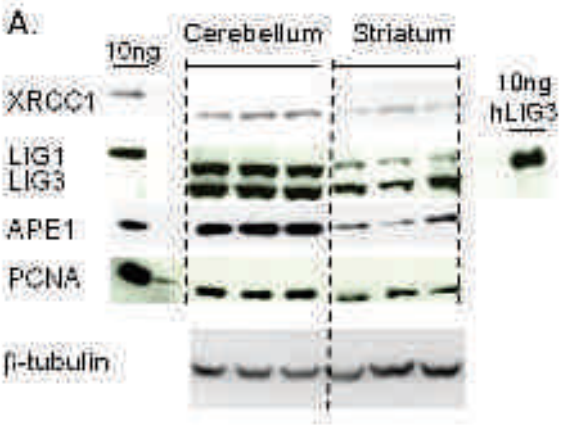
15. Groh, W. J., Groh, M. R., Shen, C., Monckton, D. G., Bodkin, C. L., and Pascuzzi, R. M. (2011) Survival and CTG repeat expansion in adults with myotonic dystrophy type 1, *Muscle Nerve 43*, 648-651.
16. Groh, W. J., Lowe, M. R., and Zipes, D. P. (2002) Severity of cardiac conduction involvement and arrhythmias in myotonic dystrophy type 1 correlates with age and CTG repeat length, *J Cardiovasc Electrophysiol 13*, 444-448.
17. Swami, M., Hendricks, A. E., Gillis, T., Massood, T., Mysore, J., Myers, R. H., and Wheeler, V. C. (2009) Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset, *Hum Mol Genet 18*, 3039-3047.
18. Cleary, J. D., Tome, S., Lopez Castel, A., Panigrahi, G. B., Foiry, L., Hagerman, K. A., Sroka, H., Chitayat, D., Gourdon, G., and Pearson, C. E. (2010) Tissue- and age-specific DNA replication patterns at the CTG/CAG-expanded human myotonic dystrophy type 1 locus, *Nat Struct Mol Biol 17*, 1079-1087.
19. van den Broek, W. J., Nelen, M. R., Wansink, D. G., Coerwinkel, M. M., te Riele, H., Groenen, P. J., and Wieringa, B. (2002) Somatic expansion behaviour of the (CTG)n repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins, *Hum Mol Genet 11*, 191-198.
20. Tome, S., Holt, I., Edelmann, W., Morris, G. E., Munnich, A., Pearson, C. E., and Gourdon, G. (2009) MSH2 ATPase domain mutation affects CTG*CAG repeat instability in transgenic mice, *PLoS Genet 5*, e1000482.
21. Savouret, C., Garcia-Cordier, C., Megret, J., te Riele, H., Junien, C., and Gourdon, G. (2004) MSH2-dependent germinal CTG repeat expansions are produced continuously in spermatogonia from DM1 transgenic mice, *Mol Cell Biol 24*, 629-637.
22. Dragileva, E., Hendricks, A., Teed, A., Gillis, T., Lopez, E. T., Friedberg, E. C., Kucherlapati, R., Edelmann, W., Lunetta, K. L., MacDonald, M. E., and Wheeler, V. C. (2009) Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes, *Neurobiol Dis 33*, 37-47.
23. Manley, K., Shirley, T. L., Flaherty, L., and Messer, A. (1999) Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice, *Nat Genet 23*, 471-473.
24. Wheeler, V. C., Lebel, L. A., Vrbanac, V., Teed, A., te Riele, H., and MacDonald, M. E. (2003) Mismatch repair gene Msh2 modifies the timing of early disease in Hdh(Q111) striatum, *Hum Mol Genet 12*, 273-281.
25. Pearson, C. E., and Sinden, R. R. (1996) Alternative structures in duplex DNA formed within the trinucleotide repeats of the myotonic dystrophy and fragile X loci, *Biochemistry 35*, 5041-5053.
26. Gacy, A. M., Goellner, G., Juranic, N., Macura, S., and Mcmurray, C. T. (1995) Trinucleotide repeats that expand in human disease form hairpin structures in vitro, *Cell 81*, 533-540.
27. Seriola, A., Spits, C., Simard, J. P., Hilven, P., Haentjens, P., Pearson, C. E., and Sermon, K. (2011) Huntington's and myotonic dystrophy hESCs: down-regulated trinucleotide repeat instability and mismatch repair machinery expression upon differentiation, *Hum Mol Genet 20*, 176-185.
28. Kovtun, I. V., Liu, Y., Bjoras, M., Klungland, A., Wilson, S. H., and McMurray, C. T. (2007) OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells, *Nature 447*, 447-452.
29. Subramanian, J., Vijayakumar, S., Tomkinson, A. E., and Arnheim, N. (2005) Genetic instability induced by overexpression of DNA ligase I in budding yeast, *Genetics 171*, 427-441.
30. Freudenreich, C. H., Kantrow, S. M., and Zakian, V. A. (1998) Expansion and length-dependent fragility of CTG repeats in yeast, *Science 279*, 853-856.
31. Robertson, A. B., Klungland, A., Rognes, T., and Leiros, I. (2009) DNA repair in mammalian cells: Base excision repair: the long and short of it, *Cell Mol Life Sci 66*, 981-993.

32. Fortini, P., and Dogliotti, E. (2007) Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways, *DNA Repair (Amst) 6*, 398-409.
33. Goula, A. V., Berquist, B. R., Wilson, D. M., 3rd, Wheeler, V. C., Trottier, Y., and Merienne, K. (2009) Stoichiometry of base excision repair proteins correlates with increased somatic CAG instability in striatum over cerebellum In Huntington's disease transgenic mice, *PLoS Genet 5*, e1000749.
34. Liu, Y., Prasad, R., Beard, W. A., Hou, E. W., Horton, J. K., McMurray, C. T., and Wilson, S. H. (2009) Coordination between polymerase beta and FEN1 can modulate CAG repeat expansion, *J Biol Chem 284*, 28352-28366.
35. Hou, C., Zhang, T., Tian, L., Huang, J., Gu, L., and Li, G. M. (2011) The Role of XPG in Processing (CAG)n/(CTG)n DNA Hairpins, *Cell Biosci 1*, 11.
36. Hou, C., Chan, N. L., Gu, L., and Li, G. M. (2009) Incision-dependent and error-free repair of (CAG)(n)/(CTG)(n) hairpins in human cell extracts, *Nat Struct Mol Biol 16*, 869-875.
37. Panigrahi, G. B., Slean, M. M., Simard, J. P., Gileadi, O., and Pearson, C. E. (2010) Isolated short CTG/CAG DNA slip-outs are repaired efficiently by hMutSbeta, but clustered slip-outs are poorly repaired, *Proc Natl Acad Sci U S A 107*, 12593-12598.
38. Panigrahi, G. B., Lau, R., Montgomery, S. E., Leonard, M. R., and Pearson, C. E. (2005) Slipped (CTG)*(CAG) repeats can be correctly repaired, escape repair or undergo error-prone repair, *Nat Struct Mol Biol 12*, 654-662.
39. Jarem, D. A., Wilson, N. R., and Delaney, S. (2009) Structure-dependent DNA damage and repair in a trinucleotide repeat sequence, *Biochemistry 48*, 6655-6663.
40. Jarem, D. A., Wilson, N. R., Schermerhorn, K. M., and Delaney, S. (2011) Incidence and persistence of 8-oxo-7,8-dihydroguanine within a hairpin intermediate exacerbates a toxic oxidation cycle associated with trinucleotide repeat expansion, *DNA Repair (Amst) 10*, 887-896.
41. Lee, B. I., and Wilson, D. M., 3rd. (1999) The RAD2 domain of human exonuclease 1 exhibits 5' to 3' exonuclease and flap structure-specific endonuclease activities, *J Biol Chem 274*, 37763-37769.
42. Nguyen, L. H., Barsky, D., Erzberger, J. P., and Wilson, D. M., 3rd. (2000) Mapping the protein-DNA interface and the metal-binding site of the major human apurinic/apyrimidinic endonuclease, *J Mol Biol 298*, 447-459.
43. Erzberger, J. P., Barsky, D., Scharer, O. D., Colvin, M. E., and Wilson, D. M., 3rd. (1998) Elements in abasic site recognition by the major human and Escherichia coli apurinic/apyrimidinic endonucleases, *Nucleic Acids Res 26*, 2771-2778.
44. Chen, X., Pascal, J., Vijayakumar, S., Wilson, G. M., Ellenberger, T., and Tomkinson, A. E. (2006) Human DNA ligases I, III, and IV-purification and new specific assays for these enzymes, *Methods Enzymol 409*, 39-52.
45. Della-Maria, J., Zhou, Y., Tsai, M. S., Kuhnlein, J., Carney, J. P., Paull, T. T., and Tomkinson, A. E. (2011) Human Mre11/Human Rad50/Nbs1 and DNA Ligase III{alpha}/XRCC1 Protein Complexes Act Together in an Alternative Nonhomologous End Joining Pathway, *J Biol Chem 286*, 33845-33853.
46. Levin, D. S., Bai, W., Yao, N., O'Donnell, M., and Tomkinson, A. E. (1997) An interaction between DNA ligase I and proliferating cell nuclear antigen: implications for Okazaki fragment synthesis and joining, *Proc Natl Acad Sci U S A 94*, 12863-12868.
47. Mangiarini, L., Sathasivam, K., Seller, M., Cozens, B., Harper, A., Hetherington, C., Lawton, M., Trottier, Y., Lehrach, H., Davies, S. W., and Bates, G. P. (1996) Exon 1 of the *HD* gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice, *Cell 87*, 493-506.

48. Petermann, E., Ziegler, M., and Oei, S. L. (2003) ATP-dependent selection between single nucleotide and long patch base excision repair, *DNA Repair (Amst) 2*, 1101-1114.
49. Wei, W., and Englander, E. W. (2008) DNA polymerase beta-catalyzed-PCNA independent long patch base excision repair synthesis: a mechanism for repair of oxidatively damaged DNA ends in post-mitotic brain, *J Neurochem 107*, 734-744.
50. Rao, K. S., Annapurna, V. V., and Raji, N. S. (2001) DNA polymerase-beta may be the main player for defective DNA repair in aging rat neurons, *Ann N Y Acad Sci 928*, 113-120.
51. Wilson, S. H., Beard, W. A., Shock, D. D., Batra, V. K., Cavanaugh, N. A., Prasad, R., Hou, E. W., Liu, Y., Asagoshi, K., Horton, J. K., Stefanick, D. F., Kedar, P. S., Carrozza, M. J., Masaoka, A., and Heacock, M. L. (2010) Base excision repair and design of small molecule inhibitors of human DNA polymerase beta, *Cell Mol Life Sci 67*, 3633-3647.
52. Wilson, D. M., 3rd, and McNeill, D. R. (2007) Base excision repair and the central nervous system, *Neuroscience 145*, 1187-1200.
53. Karahalil, B., Hogue, B. A., de Souza-Pinto, N. C., and Bohr, V. A. (2002) Base excision repair capacity in mitochondria and nuclei: tissue-specific variations, *Faseb J 16*, 1895-1902.
54. Szczesny, B., Tann, A. W., and Mitra, S. (2010) Age- and tissue-specific changes in mitochondrial and nuclear DNA base excision repair activity in mice: Susceptibility of skeletal muscles to oxidative injury, *Mech Ageing Dev 131*, 330-337.
55. Intano, G. W., Cho, E. J., McMahan, C. A., and Walter, C. A. (2003) Age-related base excision repair activity in mouse brain and liver nuclear extracts, *J Gerontol A Biol Sci Med Sci 58*, 205-211.
56. Narciso, L., Fortini, P., Pajalunga, D., Franchitto, A., Liu, P., Degan, P., Frechet, M., Demple, B., Crescenzi, M., and Dogliotti, E. (2007) Terminally differentiated muscle cells are defective in base excision DNA repair and hypersensitive to oxygen injury, *Proc Natl Acad Sci U S A 104*, 17010-17015.
57. Vallur, A. C., and Maizels, N. (2010) Complementary roles for exonuclease 1 and Flap endonuclease 1 in maintenance of triplet repeats, *J Biol Chem 285*, 28514-28519.
58. Spiro, C., Pelletier, R., Rolfsmeier, M. L., Dixon, M. J., Lahue, R. S., Gupta, G., Park, M. S., Chen, X., Mariappan, S. V., and McMurray, C. T. (1999) Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats, *Mol Cell 4*, 1079-1085.
59. Pascucci, B., Stucki, M., Jonsson, Z. O., Dogliotti, E., and Hubscher, U. (1999) Long patch base excision repair with purified human proteins. DNA ligase I as patch size mediator for DNA polymerases delta and epsilon, *J Biol Chem 274*, 33696-33702.
60. Levin, D. S., McKenna, A. E., Motycka, T. A., Matsumoto, Y., and Tomkinson, A. E. (2000) Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair, *Curr Biol 10*, 919-922.
61. Hartenstine, M. J., Goodman, M. F., and Petruska, J. (2002) Weak strand displacement activity enables human DNA polymerase beta to expand CAG/CTG triplet repeats at strand breaks, *J Biol Chem 277*, 41379-41389.
62. Sobol, R. W., Horton, J. K., Kuhn, R., Gu, H., Singhal, R. K., Prasad, R., Rajewsky, K., and Wilson, S. H. (1996) Requirement of mammalian DNA polymerase-beta in base-excision repair, *Nature 379*, 183-186.
63. Asagoshi, K., Liu, Y., Masaoka, A., Lan, L., Prasad, R., Horton, J. K., Brown, A. R., Wang, X. H., Bdour, H. M., Sobol, R. W., Taylor, J. S., Yasui, A., and Wilson, S. H. (2010) DNA polymerase beta-dependent long patch base excision repair in living cells, *DNA Repair (Amst) 9*, 109-119.
64. Asagoshi, K., Tano, K., Chastain, P. D., 2nd, Adachi, N., Sonoda, E., Kikuchi, K., Koyama, H., Nagata, K., Kaufman, D. G., Takeda, S., Wilson, S. H., Watanabe, M., Swenberg, J. A., and

Nakamura, J. (2010) FEN1 functions in long patch base excision repair under conditions of oxidative stress in vertebrate cells, *Mol Cancer Res 8*, 204-215.
65. Klungland, A., and Lindahl, T. (1997) Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1), *Embo J 16*, 3341-3348.

Figure 1



B.

Protein	Cb	St
POLβ	1	1
FEN1	6	1.5
XRCC1	3.5	1.6
LIG1	2	0.4
LIG3	2	1.5
APE1	15	5
PCNA	2	1

Numbers correspond to protein molar ratios relative to POLβ

Table 1

	nt number	Sequence 5'-3'*
Target		
D1	100	5'CGAGTCATCTAGCATCCGTAUAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG CAGCAGCAGCAGCAGCAGTACGTAGACTTACTCATTGC 3'
M1	100	5'CGAGTCATCTAGCATCCGTAUTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG CTGCTGCTGCTGCTGCTGTACGTAGACTTACTCATTGC 3'
M2	100	5'CGAGTCATCTAGCATCCGTACAUCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG CAGCAGCAGCAGCAGCAGTACGTAGACTTACTCATTGC 3'
M3	100	5'CGAGTCATCTAGCATCCGTACAGCAGCAGCAGCAGCAGCAGCAGCAUCAGCAGCAGCAGCAG CAGCAGCAGCAGCAGCAGTACGTAGACTTACTCATTGC 3'
M4	100	5'CGAGTCATCTAGCATCCGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG CAGCAGCAGCAGCAGCAUTACGTAGACTTACTCATTGC 3'
M5	100	5'CGAGTCATCTAGCATCCGTACTUCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG CTGCTGCTGCTGCTGCTGTACGTAGACTTACTCATTGC 3'
M6	100	5'CGAGTCATCTAGCATCCGTACTGCTGCTGCTGCTGCTGCTGCTGCTUCTGCTGCTGCTGCTG CTGCTGCTGCTGCTGCTGTACGTAGACTTACTCATTGC 3'
M7	100	5'CGAGTCATCTAGCATCCGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG CTGCTGCTGCTGCTGCTUTACGTAGACTTACTCATTGC 3'
U14	33	5'CGTGACCTAATGCUCCTTTAAGTCTGAAGGCTG 3'
U16	33	5'CGTGACCTAATGCGCUTTTAAGTCTGAAGGCTG 3'
Template		
D1c/M2c	100	5'GCAATGAGTAAGTCTACGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT GCTGCTGCTGCTGCTGCTGTACGGATGCTAGATGACTCG 3'
M1c/M5c	100	5'GCAATGAGTAAGTCTACGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGTACGGATGCTAGATGACTCG 3'
G16	33	5'CAGCCTTCAGACTTAAAGGCGCATTAGGTCACG 3'

*The modified base is in green; U, deoxyuridine. CAG and CTG repeats are shown in blue and red, respectively.

Figure 2

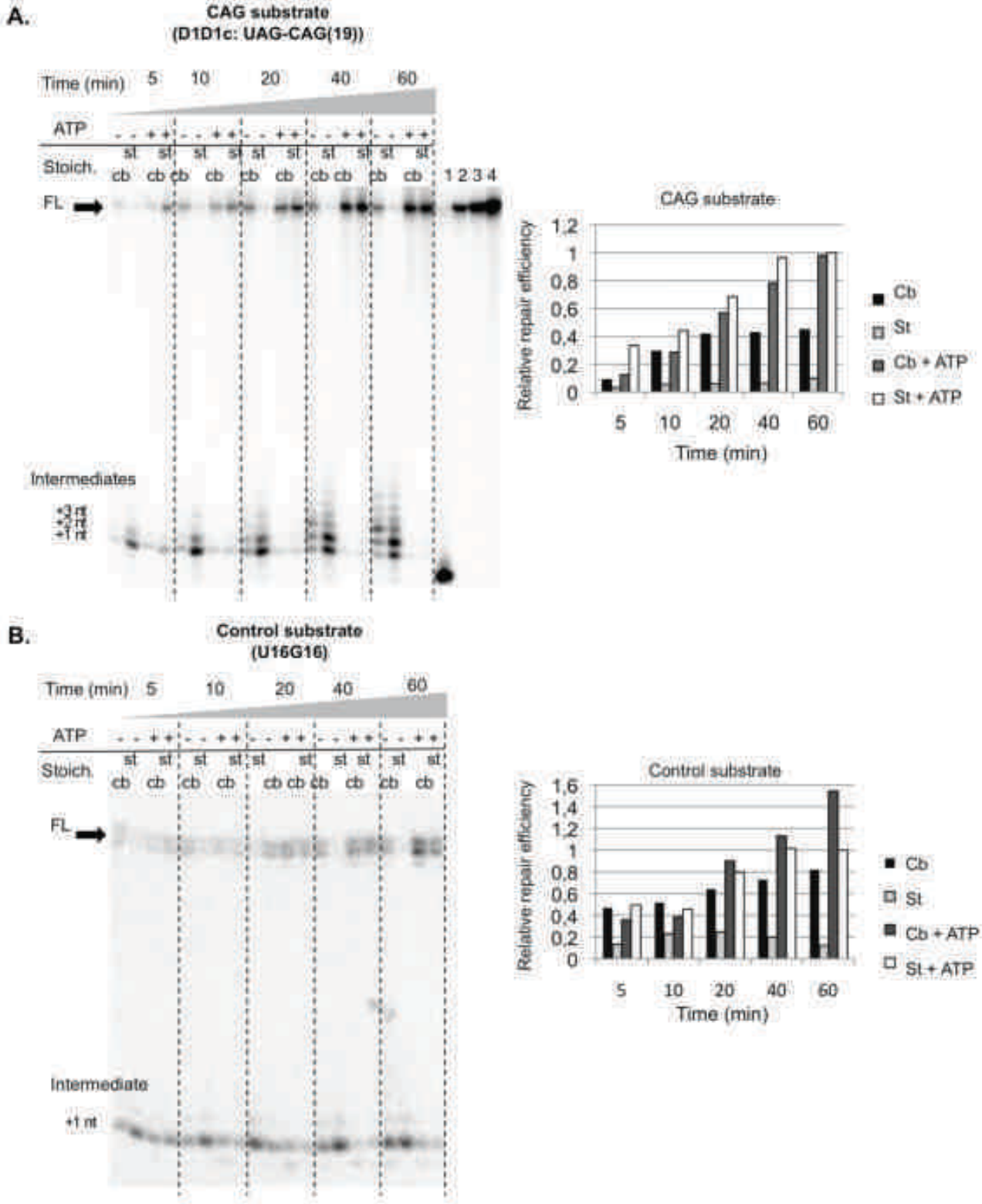

A.
CAG substrate
(D1D1c: UAG-CAG(19))
Time (min) 5 10 20 40 60
ATP
Stoich.
FL
1 2 3 4
Intermediates
+3 nt
+2 nt
+1 nt
CAG substrate
Relative repair efficiency
1,2
1
0,8
0,6
0,4
0,2
0
5 10 20 40 60
Time (min)
Cb
St
Cb + ATP
St + ATP
B.
Control substrate
(U16G16)
Time (min) 5 10 20 40 60
ATP
Stoich.
FL
Intermediate
+1 nt
Control substrate
Relative repair efficiency
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0
5 10 20 40 60
Time (min)
Cb
St
Cb + ATP
St + ATP

Figure 3

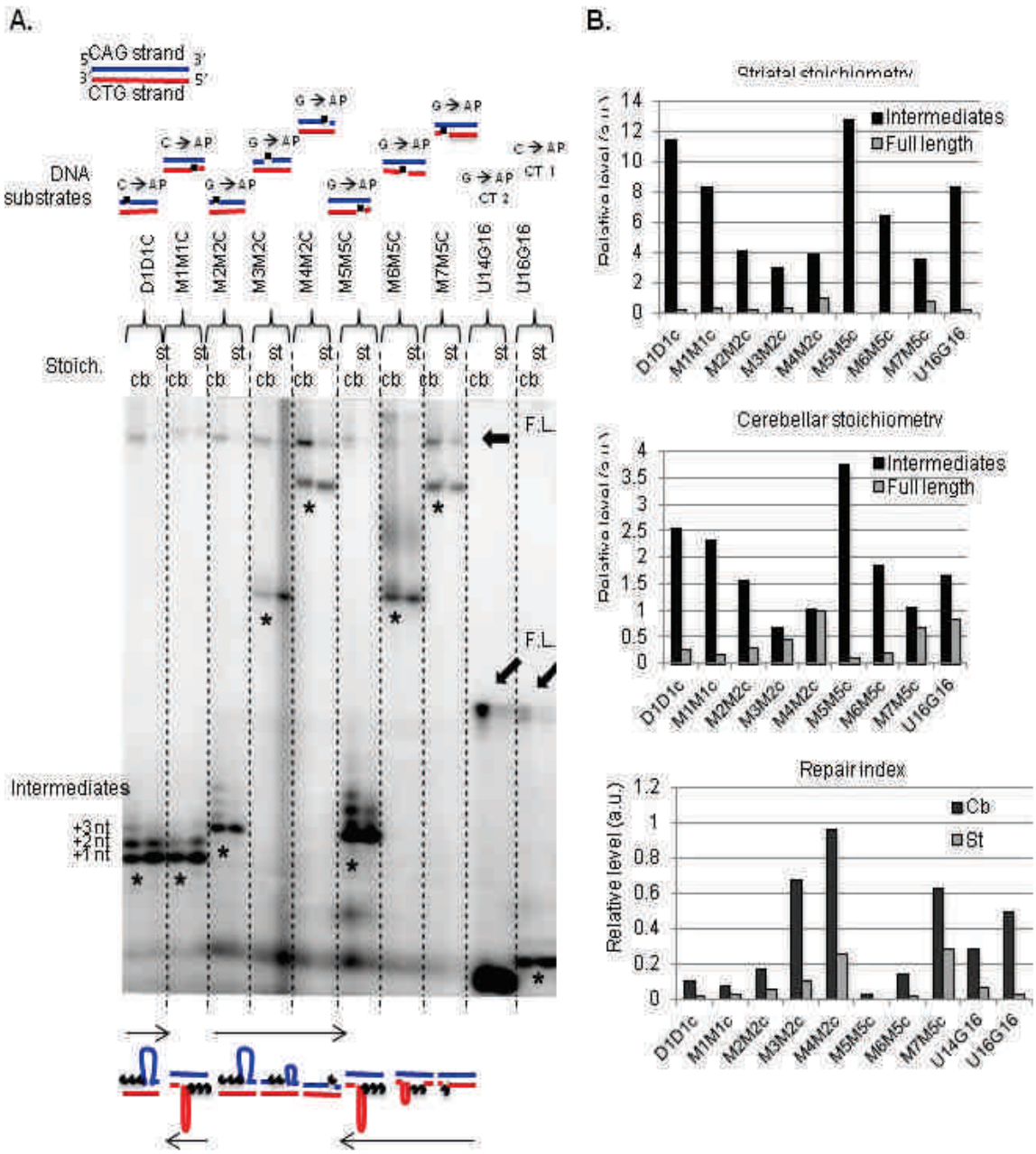
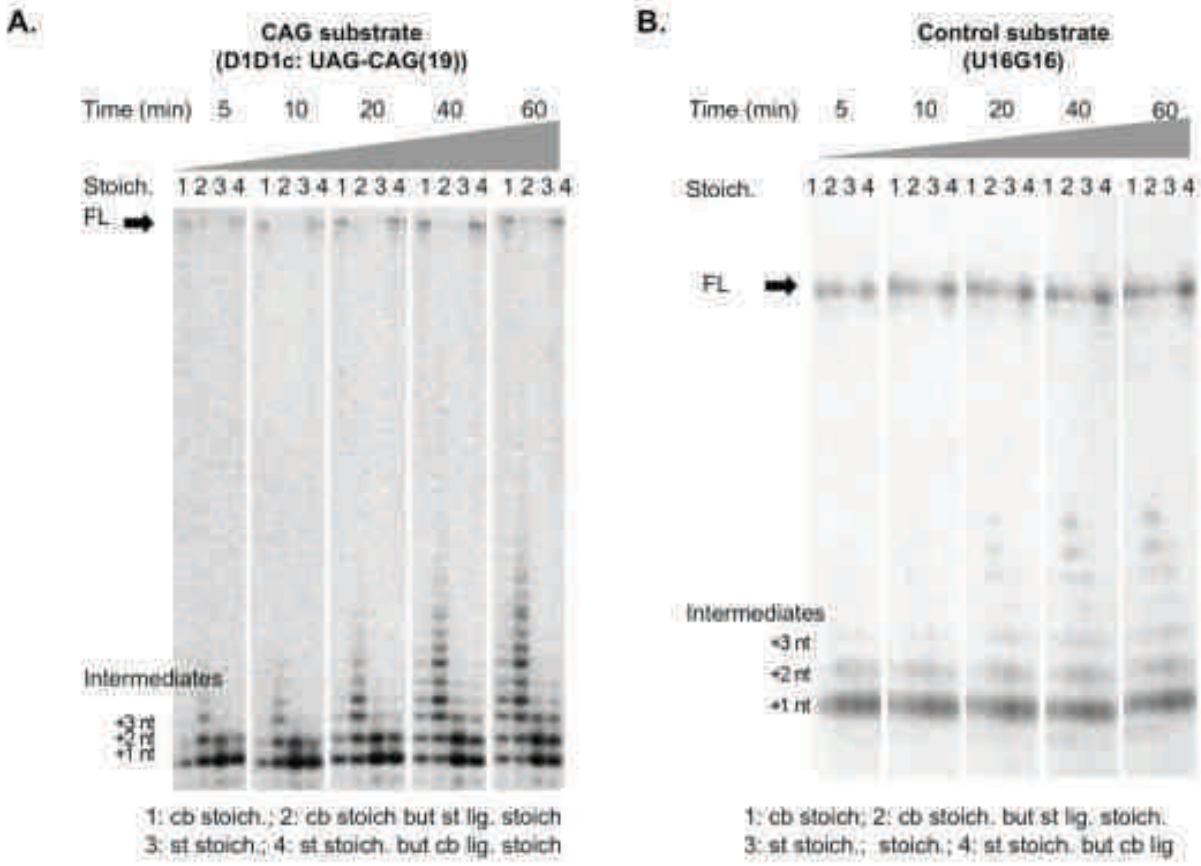
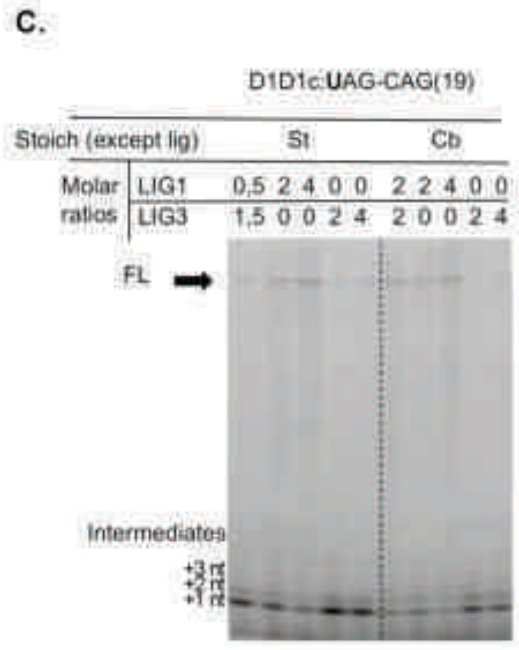
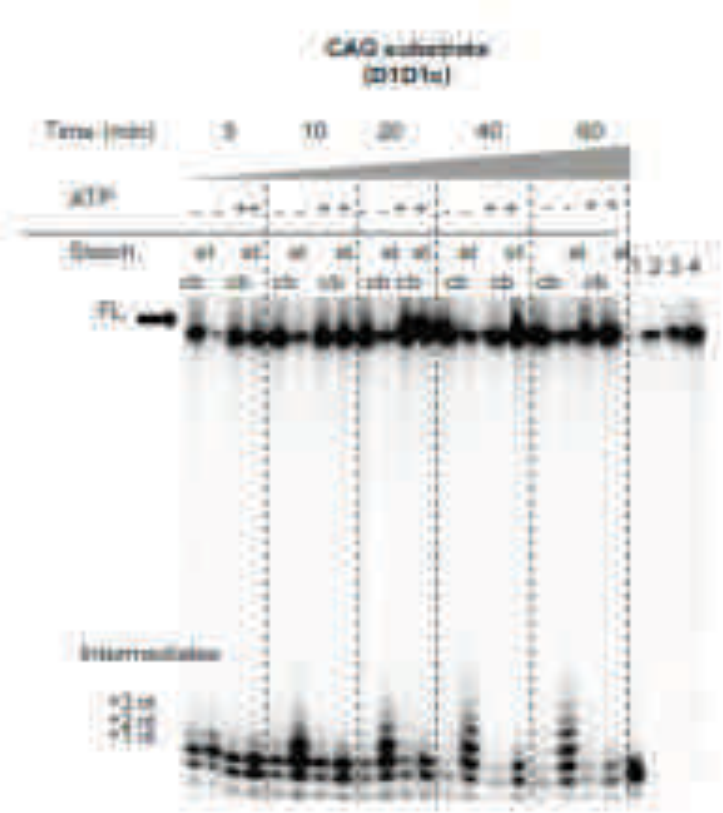

A.
CAG strand
CTG strand
DNA substrates
C→AP
G→AP
CT 1
CT 2
D1D1C
M1M1C
M2M2C
M3M2C
M4M2C
M5M5C
M6M5C
M7M5C
U14G16
U16G16
Stoich.
St
Cb
F.L.
Intermediates
+3nt
+2nt
+1nt
B.
Striatal stoichiometry
Intermediates
Full length
Relative level (a.u.)
Cerebellar stoichiometry
Repair index
Cb
St

Figure 4

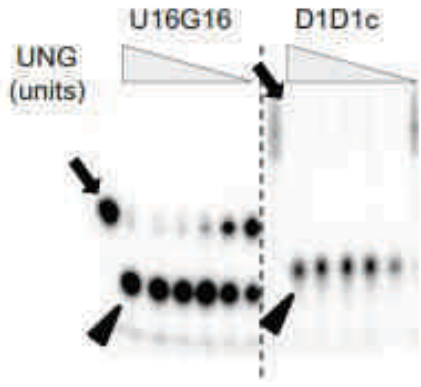

A.
CAG substrate
(D1D1c: UAG-CAG(19))
Time (min) 5 10 20 40 60
Stoich. 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
FL
Intermediates
+3 nt
+2 nt
+1 nt
1: cb stoich.; 2: cb stoich but st lig. stoich
3: st stoich.; 4: st stoich. but cb lig. stoich
B.
Control substrate
(U16G16)
Time (min) 5 10 20 40 60
Stoich. 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
FL
Intermediates
+3 nt
+2 nt
+1 nt
1: cb stoich; 2: cb stoich. but st lig. stoich.
3: st stoich.; stoich.; 4: st stoich. but cb lig


C.
D1D1c:UAG-CAG(19)
Stoich (except lig) St Cb
Molar ratios LIG1 0,5 2 4 0 0 2 2 4 0 0
LIG3 1,5 0 0 2 4 2 0 0 2 4
FL
Intermediates
+3 nt
+2 nt
+1 nt

Supplementary Figure 1



Supplementary Figure 2



III) Discussion

1. Update on BER in the tissue-specific instability of TNR

Our work allowed gaining insights into the potential role of BER in triplet repeat instability. We first showed that lesions do accumulate at CAG/CTG repeats, suggesting a higher propensity of those sequences to recruit repair proteins. The findings by Kovtun showed that deficiency of the DNA glycosylase *Ogg1* in HD R6/1 mice reduced somatic CAG instability (Kovtun et al., 2007). The accumulation of oxidant lesions in the ageing brain that they observed led them to support the "toxic oxidation cycle" model. However, our data show that DNA damage levels at CAG/CTG repeats are independent of mouse age, suggesting that age-dependent increase of DNA damage is not responsible for age-dependent increase of somatic CAG instability (Goula et al., 2009). As CAG/CTG repeats form hairpins *in vitro* (Owen et al., 2005), we examined the outcome of DNA lesions located at CAG/CTG hairpins, and observed that their accessibility to BER proteins is decreased, which might result in lesion accumulation at CAG/CTG repeats (Goula et al., 2009). Assessment of main BER activities from protein extracts prepared from the striatum and cerebellum of R6/1 mice revealed globally higher activities in the cerebellum compared to the striatum. In addition, we found that the stoichiometry of BER proteins was different in the striatum and in the cerebellum of R6/1 mice (Goula et al., 2009). While the levels of OGG1 protein and activity were similar in the two tissues, suggesting similar initiation of repair, the subsequent repair proteins and/or activities were different in the two tissues. In particular, the levels of FEN1, a LP-BER enzyme, was ≈ five-fold higher in the cerebellum compared to the striatum, resulting in higher FEN1/Polβ in the cerebellum and more efficient processing of the lesion (Goula et al., 2009). The importance of FEN:Polβ stoichiometry in modulating CAG repeat length is also supported by the results by Liu (Liu et al., 2009). We investigated further the hypothesis that processing of a lesion by LP-BER would be more efficient under the cerebellar BER protein stoichiometry as compared to the striatal BER protein stoichiometry (Goula et al. submitted), thereby increasing the probability of CAG/CTG instability. Using a reconstitution repair assay and various CAG/CTG substrates containing an AP lesion at located in different positions within the substrate, we show that the AP lesion is preferentially processed via LP-BER, particularly when the substrate was prone to

hairpin formation during BER (i.e. when the lesion was located 5' to the repeat tract or on the CTG strand). Moreover, repair was more efficient when using the cerebellar BER protein stoichiometry, as compared to the striatal BER protein stoichiometry, suggesting faster processing of the lesion in the cerebellum as opposed to stalling of the repair reaction in the striatum. Those results suggest that the tissue-specific stoichiometry of BER proteins might predispose or limit the propensity of trinucleotide repeats to instability.

Together, we propose that CAG/CTG repeats form secondary structures that lead to accumulation of DNA lesions. We suggest that some lesions might become more accessible, for instance during transcription, allowing their processing by BER. Our results support that processing of a lesion by BER at CAG/CTG repeats is less efficient when using the striatal BER protein stoichiometry as compared to the cerebellar BER protein stoichiometry. We suggest that inefficient processing of oxidative lesions in the striatum contributes to tissue-selective CAG instability in HD.

2. What is the nature of oxidative lesions at CAG repeats?

An increase of the amount of some oxidant lesions, such as 8-oxodG and 5-OH-uracil, in ageing mouse brain has been reported (Kovtun et al., 2007). The finding that in R6/1 mice deficiency of the DNA glycosylase OGG1, which mainly targets 8-oxodG lesions, led to limitation of repeat instability, supported the "toxic oxidation cycle" model. However, deficiency of other DNA glycosylases in R6/1 mice such as AAG and NTH1 did not alter repeat instability, suggesting that TNR instability is not affected to the same extent by the different oxidative lesions, and that 8-OxodG play an important role in repeat instability in HD. Whether additional DNA glycosylases targeting other specific lesions, including Neil1 and Neil2, contribute to TNR instability remains to be investigated. Alternatively, functional redundancy between DNA glycosylases may also mask the effect of deficiency of a specific glycosylase and might explain why inactivation of AAG or NTH1 in R6/1 had no effect on instability. Thus, it could be of interest to inactivate several DNA glycosylases simultaneously in R6/1 mice. In addition, it

would also be interesting to screen different tissues for several DNA glycosylase activities relative to tissue-specific repeat instability levels, in order to correlate instability levels with specific glycosylase requirement. Another open question is the level of oxidant lesions present at CAG/CTG repeats. Kovtun et al., (*Nature* 2007), did not specifically examine the content of oxidant lesions at CAG/CTG repeats but measured globally the level of lesions in total ageing brains. However, when we compared the amount of lesions at the CAG/CTG repeats in the cerebellum versus the striatum, we found that the levels of lesions at repeats were increased, but were not tissue- or age-dependent, suggesting that the tissue-specificity of repeat instability is not dependent on the number of lesions at CAG/CTG repeats. Finally, it would be interesting to investigate whether the involvement of lesions and DNA glycosylases in repeat instability in the HD mice is shared by other CAG/CTG polyQ or DM1 models, or by other trinucleotide repeat sequences such as those causing FRDA or FXS.

3. What mechanism(s) lead(s) to accumulation of oxidative lesions at CAG repeats?

Accumulation of lesions was age- and tissue-independent, though significantly higher at CAG repeats compared to control sequence in R6/1 tissues (Goula et al., 2009). CAG/CTG repeats were shown to form stable secondary structures *in vitro*, in a repeat length-dependent manner (Figueroa et al., 2011; Gacy et al., 1995; Pearson and Sinden, 1996; Pearson et al., 1998b) Repeat hairpins can be efficiently removed or, alternatively, contribute to the accumulation of DNA damages (Hou et al., 2009; Zhang et al., 2011). We showed that lesions located at loops of hairpins are refractory to repair by BER initiating enzymes, including OGG1 and APE1 (Goula et al., 2009). Additional studies show that hairpins can promote damage and delay repair (Jarem et al., 2009; Jarem et al., 2011), likely due to a higher exposure of DNA bases to oxidized or mutagenic agents or to intrinsic characteristics of hairpin structures. Interestingly, triplet repeat pre-mutation sizes were reported to adopt a non-B conformation, which is thermodynamically stable enough to participate to the instability process, by a potential accumulation of DNA damage at those structures (Avila Figueroa and Delaney, 2010;

Figueroa et al., 2011; Jarem and Delaney, 2011; Volker et al., 2002). Altogether those results suggest that lesions accumulate at stable CAG/CTG secondary structures, and ageing may not contribute substantially to this process.

Transcription requires an accessible chromatin structure and could therefore play a role in TNR instability, by promoting DNA damage accumulation. Interestingly, knock-down of *Csb* in human cells led to stabilization of CAG/CTG instability (Lin and Wilson, 2007). In a similar manner, R6/1 mice deficient for *Csb* showed reduction of intergenerational instability and decrease in somatic instability in the brain (Kovtun et al., 2011). Those results suggest an implication of TC-NER in repeat instability. Indeed since CSB is not only a TC-NER factor but is also participating in processing of oxidant lesions, it might suggest that BER could contribute to transcription-dependent CAG instability (Thorslund et al., 2005). However, the fact that double deficiency of *Csb* and *Ogg1* in R6/1 mice led to exacerbation of instability rather suggests that those two factors do not lead to repeat instability by acting on the same pathway.

4. Does BER, and more particularly LP-BER directly contribute to CAG/CTG instability?

How could BER modulate CAG/CTG instability in disease? It has been suggested that the contribution of BER to CAG/CTG instability may be direct. In an *in vitro* repair assay, expansion products were generated with CAG-substrates (similar to the substrates used in our experiments, see § Publication 2) using cell extracts supplemented or not with BER proteins, such as APE1, Polβ and FEN1 (Liu et al., 2009). This experimental setting revealed that the amount of expansion was dependent upon the presence of BER proteins. However, in our repair assay, where protein ratios were reconstituted to reflect the stoichiometries observed in mouse striatum or cerebellum, we did not detect expansion products, suggesting that BER on its own is not sufficient to induce CAG/CTG instability. Furthermore, the level of somatic CAG/CTG instability was not changed in DM1 or HD mice haploinsufficient for *Fen1*, or in DM1 mice expressing a mutated version of *LigI* with <5% residual ligase activity (Spiro and McMurray, 2003; Tome et al., 2011; van den Broek et al., 2006). As complete inactivation of

Fen1 and *Lig1* is embryonic lethal (Bentley et al., 1996; Kucherlapati et al., 2002; Larsen et al., 2003; Petrini et al., 1995), thus precluding the use of such null mice, the significance of the results with the mutant animals above is not yet clear, but could suggest that BER does not directly contribute to CAG/CTG instability or is not an obligatory pathway. Accordingly, somatic instability was only moderately reduced in HD mice deficient for *Ogg1*, in contrast to the severe reduction seen in HD and DM1 mice deficient for *Msh2* or *Msh3* (Dragileva et al., 2009; Foiry et al., 2006; Kovtun et al., 2007; Manley et al., 1999; Savouret et al., 2004; van den Broek et al., 2002). Whether BER contribution to CAG/CTG instability is dependent upon MMR is an intriguing possibility that would remain to be explored.

IV) REFERENCES

Acharya, S., Wilson, T., Gradia, S., Kane, M.F., Guerrette, S., Marsischky, G.T., Kolodner, R., and Fishel, R. (1996). hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. Proc Natl Acad Sci U S A *93*, 13629-13634.

Al-Mahdawi, S., Pinto, R.M., Ismail, O., Varshney, D., Lymperi, S., Sandi, C., Trabzuni, D., and Pook, M. (2008). The Friedreich ataxia GAA repeat expansion mutation induces comparable epigenetic changes in human and transgenic mouse brain and heart tissues. Hum Mol Genet *17*, 735-746.

Andrew, S.E., Goldberg, Y.P., Kremer, B., Telenius, H., Theilmann, J., Adam, S., Starr, E., Squitieri, F., Lin, B., Kalchman, M.A*., et al.* (1993). The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. Nature Genet *4*, 398-403.

Anvret, M., Ahlberg, G., Grandell, U., Hedberg, B., Johnson, K., and Edstrom, L. (1993). Larger expansions of the CTG repeat in musde compared to lymphocytes from patients with myotonic dystrophy. Hum Mol Genet *2*, 1397-1400.

Arveiler, B., Alembik, Y., Hanauer, A., Jacobs, P., Tranebjaerg, L., Mikkelsen, M., Puissant, H., Piet, L.L., and Mandel, J.L. (1988). Linkage analysis suggests at least two loci for X-linked non-specific mental retardation. Am J Med Genet *30*, 473-483.

Asagoshi, K., Liu, Y., Masaoka, A., Lan, L., Prasad, R., Horton, J.K., Brown, A.R., Wang, X.H., Bdour, H.M., Sobol, R.W*., et al.* (2010a). DNA polymerase beta-dependent long patch base excision repair in living cells. DNA Repair (Amst) *9*, 109-119.

Asagoshi, K., Tano, K., Chastain, P.D., 2nd, Adachi, N., Sonoda, E., Kikuchi, K., Koyama, H., Nagata, K., Kaufman, D.G., Takeda, S*., et al.* (2010b). FEN1 functions in long patch base excision repair under conditions of oxidative stress in vertebrate cells. Mol Cancer Res *8*, 204-215.

Ashley-Koch, A.E., Robinson, H., Glicksman, A.E., Nolin, S.L., Schwartz, C.E., Brown, W.T., Turner, G., and Sherman, S.L. (1998). Examination of factors associated with instability of the FMR1 CGG repeat. Am J Hum Genet *63*, 776-785.

Avila Figueroa, A., and Delaney, S. (2010). Mechanistic studies of hairpin to duplex conversion for trinudeotide repeat sequences. J Biol Chem *285*, 14648-14657.

Aziz, N.A., van Belzen, M.J., Coops, I.D., Belfroid, R.D., and Roos, R.A. (2011). Parent-of-origin differences of mutant HTT CAG repeat instability in Huntington's disease. Eur J Med Genet *54*, e413-418.

Bai, H., Madabushi, A., Guan, X., and Lu, A.L. (2010). Interaction between human mismatch repair recognition proteins and checkpoint sensor Rad9-Rad1-Hus1. DNA Repair (Amst) *9*, 478-487.

Balakrishnan, L., Brandt, P.D., Lindsey-Boltz, L.A., Sancar, A., and Bambara, R.A. (2009). Long patch base excision repair proceeds via coordinated stimulation of the multienzyme DNA repair complex. J Biol Chem *284*, 15158-15172.

Barnes, D.E., Tomkinson, A.E., Lehmann, A.R., Webster, A.D., and Lindahl, T. (1992). Mutations in the DNA ligase I gene of an individual with immunodeficiencies and cellular hypersensitivity to DNA-damaging agents. Cell *69*, 495-503.

Baskaran, S., Datta, S., Mandal, A., Gulati, N., Totey, S., Anand, R.R., and Brahmachari, V. (2002). Instability of CGG repeats in transgenic mice. Genomics *80*, 151-157.

Belotserkovskii, B.P., Liu, R., Tornaletti, S., Krasilnikova, M.M., Mirkin, S.M., and Hanawalt, P.C. (2010). Mechanisms and implications of transcription blockage by guanine-rich DNA sequences. Proc Natl Acad Sci U S A *107*, 12816-12821.

Benitez, J., Robledo, M., Ramos, C., Ayuso, C., Astarloa, R., Garcia Yebenes, J., and Brambati, B. (1995). Somatic stability in chorionic villi samples and other Huntington fetal tissues. Hum Genet *96*, 229-232.

Bentley, D., Selfridge, J., Millar, J.K., Samuel, K., Hole, N., Ansell, J.D., and Melton, D.W. (1996). DNA ligase I is required for fetal liver erythropoiesis but is not essential for mammalian cell viability. Nat Genet *13*, 489-491.

Bertoni, A., Giuliano, P., Galgani, M., Rotoli, D., Ulianich, L., Adornetto, A., Santillo, M.R., Porcellini, A., and Avvedimento, V.E. (2011). Early and late events induced by polyQ-expanded proteins: identification of a common pathogenic property of polYQ-expanded proteins. J Biol Chem *286*, 4727-4741.

Biacsi, R., Kumari, D., and Usdin, K. (2008). SIRT1 inhibition alleviates gene silencing in Fragile X mental retardation syndrome. PLoS Genet *4*, e1000017.

Bidichandani, S.I., Ashizawa, T., and Patel, P.I. (1998). The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure. Am J Hum Genet *62*, 111-121.

Boiteux, S., and Radicella, J.P. (1999). Base excision repair of 8-hydroxyguanine protects DNA from endogenous oxidative stress. Biochimie *81*, 59-67.

Boiteux, S., and Radicella, J.P. (2000). The human OGG1 gene: structure, functions, and its implication in the process of carcinogenesis. Arch Biochem Biophys *377*, 1-8.

Bonini, N.M. (2002). Chaperoning brain degeneration. Proc Natl Acad Sci U S A *99 Suppl 4*, 16407-16411.

Bowater, R.P., Jaworski, A., Larson, J.E., Parniewski, P., and Wells, R.D. (1997). Transcription increases the deletion frequency of long CTG.CAG triplet repeats from plasmids in Escherichia coli. Nucleic Acids Res *25*, 2861-2868.

Brais, B., Bouchard, J.P., Xie, Y.G., Rochefort, D.L., Chretien, N., Tome, F.M., Lafreniere, R.G., Rommens, J.M., Uyama, E., Nohira, O*., et al.* (1998). Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy [published erratum appears in Nat Genet 1998 Aug;19(4):404]. Nat Genet *18*, 164-167.

Braithwaite, E.K., Kedar, P.S., Stumpo, D.J., Bertocci, B., Freedman, J.H., Samson, L.D., and Wilson, S.H. (2010). DNA polymerases beta and lambda mediate overlapping and independent roles in base excision repair in mouse embryonic fibroblasts. PLoS One *5*, e12229.

Cabelof, D.C., Guo, Z., Raffoul, J.J., Sobol, R.W., Wilson, S.H., Richardson, A., and Heydari, A.R. (2003). Base excision repair deficiency caused by polymerase beta haploinsufficiency: accelerated DNA damage and increased mutational response to carcinogens. Cancer Res *63*, 5799-5807.

Caldecott, K.W., Aoufouchi, S., Johnson, P., and Shall, S. (1996). XRCC1 polypeptide interacts with DNA polymerase beta and possibly poly (ADP-ribose) polymerase, and DNA ligase III is a novel molecular 'nick-sensor' in vitro. Nucleic Acids Res *24*, 4387-4394.

Callahan, J.L., Andrews, K.J., Zakian, V.A., and Freudenreich, C.H. (2003). Mutations in yeast replication proteins that increase CAG/CTG expansions also increase repeat fragility. Mol Cell Biol *23*, 7849-7860.

Campuzano, V., Montermini, L., Molto, M.D., Pianese, L., Cossee, M., Cavalcanti, F., Monros, E., Rodius, F., Duclos, F., Monticelli, A.*, et al.* (1996). Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion [see comments]. Science *271*, 1423-1427.

Cappelli, E., Taylor, R., Cevasco, M., Abbondandolo, A., Caldecott, K., and Frosina, G. (1997). Involvement of XRCC1 and DNA ligase III gene products in DNA base excision repair. J Biol Chem *272*, 23970-23975.

Chen, X., Pascal, J., Vijayakumar, S., Wilson, G.M., Ellenberger, T., and Tomkinson, A.E. (2006). Human DNA ligases I, III, and IV-purification and new specific assays for these enzymes. Methods Enzymol *409*, 39-52.

Cho, D.H., Thienes, C.P., Mahoney, S.E., Analau, E., Filippova, G.N., and Tapscott, S.J. (2005). Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. Mol Cell *20*, 483-489.

Chon, H., Vassilev, A., DePamphilis, M.L., Zhao, Y., Zhang, J., Burgers, P.M., Crouch, R.J., and Cerritelli, S.M. (2009). Contributions of the two accessory subunits, RNASEH2B and RNASEH2C, to the activity and properties of the human RNase H2 complex. Nucleic Acids Res *37*, 96-110.

Chong, S.S., McCall, A.E., Cota, J., Subramony, S.H., Orr, H.T., Hughes, M.R., and Zoghbi, H.Y. (1995a). Gametic and somatic tissue-specific heterogeneity of the expanded SCA1 CAG repeat in spinocerebellar ataxia type 1. Nature Genet *10*, 344-350.

Chong, S.S., McCall, A.E., Cota, J., Subramony, S.H., Orr, H.T., Hughes, M.R., and Zoghbi, H.Y. (1995b). Gametic and somatic tissue-specific heterogeneity of the expanded SCA1 CAG repeat in spinocerebellar ataxia type 1. Nat Genet *10*, 344-350.

Chung, D.W., Rudnicki, D.D., Yu, L., and Margolis, R.L. (2011). A natural antisense transcript at the Huntington's disease repeat locus regulates HTT expression. Hum Mol Genet *20*, 3467-3477.

Chung, M.Y., Ranum, L.P., Duvick, L.A., Servadio, A., Zoghbi, H.Y., and Orr, H.T. (1993). Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. Nat Genet *5*, 254-258.

Clark, R.M., De Biase, I., Malykhina, A.P., Al-Mahdawi, S., Pook, M., and Bidichandani, S.I. (2007). The GAA triplet-repeat is unstable in the context of the human FXN locus and displays age-dependent expansions in cerebellum and DRG in a transgenic mouse model. Hum Genet *120*, 633-640.

Cleary, J.D., Nichol, K., Wang, Y.H., and Pearson, C.E. (2002). Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. Nat Genet *31*, 37-46.

Cleary, J.D., and Pearson, C.E. (2005). Replication fork dynamics and dynamic mutations: the fork-shift model of repeat instability. Trends Genet *21*, 272-280.

Cleary, J.D., Tome, S., Lopez Castel, A., Panigrahi, G.B., Foiry, L., Hagerman, K.A., Sroka, H., Chitayat, D., Gourdon, G., and Pearson, C.E. (2010). Tissue- and age-specific DNA replication patterns at the CTG/CAG-expanded human myotonic dystrophy type 1 locus. Nat Struct Mol Biol *17*, 1079-1087.

Coffee, B., Zhang, F., Ceman, S., Warren, S.T., and Reines, D. (2002). Histone modifications depict an aberrantly heterochromatinized FMR1 gene in fragile x syndrome. Am J Hum Genet *71*, 923-932.

Coffee, B., Zhang, F., Warren, S.T., and Reines, D. (1999). Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells. Nat Genet *22*, 98-101.

Cohen, S., Masyn, K., Adams, J., Hessl, D., Rivera, S., Tassone, F., Brunberg, J., DeCarli, C., Zhang, L., Cogswell, J.*, et al.* (2006). Molecular and imaging correlates of the fragile X-associated tremor/ataxia syndrome. Neurology *67*, 1426-1431.

Corrette-Bennett, S.E., Mohlman, N.L., Rosado, Z., Miret, J.J., Hess, P.M., Parker, B.O., and Lahue, R.S. (2001). Efficient repair of large DNA loops in Saccharomyces cerevisiae. Nucleic Acids Res *29*, 4134-4143.

Cossee, M., Schmitt, M., Campuzano, V., Reutenauer, L., Moutou, C., Mandel, J.L., and Koenig, M. (1997). Evolution of the Friedreich's ataxia trinucleotide repeat expansion: founder effect and premutations. Proc Natl Acad Sci U S A *94*, 7452-7457.

Cowan, K.J., Diamond, M.I., and Welch, W.J. (2003). Polyglutamine protein aggregation and toxicity are linked to the cellular stress response. Hum Mol Genet *12*, 1377-1391.

Das, A., Wiederhold, L., Leppard, J.B., Kedar, P., Prasad, R., Wang, H., Boldogh, I., Karimi-Busheri, F., Weinfeld, M., Tomkinson, A.E.*, et al.* (2006). NEIL2-initiated, APE-independent repair of oxidized bases in DNA: Evidence for a repair complex in human cells. DNA Repair (Amst) *5*, 1439-1448.

Datta, S., Alam, M.P., Majumdar, S.S., Mehta, A.K., Maiti, S., Wadhwa, N., and Brahmachari, V. (2011). Nucleosomal occupancy and CGG repeat expansion: a comparative analysis of triplet repeat region from mouse and human fragile X mental retardation gene 1. Chromosome Res *19*, 445-455.

Daughters, R.S., Tuttle, D.L., Gao, W., Ikeda, Y., Moseley, M.L., Ebner, T.J., Swanson, M.S., and Ranum, L.P. (2009). RNA gain-of-function in spinocerebellar ataxia type 8. PLoS Genet *5*, e1000600.

David, G., Durr, A., Stevanin, G., Cancel, G., Abbas, N., Benomar, A., Belal, S., Lebre, A.S., Abada-Bendib, M., Grid, D.*, et al.* (1998). Molecular and clinical correlations in autosomal dominant cerebellar ataxia with progressive macular dystrophy (SCA7). Hum Mol Genet *7*, 165-170.

Davranche, A., Aviolat, H., Zeder-Lutz, G., Busso, D., Altschuh, D., Trottier, Y., and Klein, F.A. (2011). Huntingtin affinity for partners is not changed by polyglutamine length: aggregation itself triggers aberrant interactions. Hum Mol Genet *20*, 2795-2806.

De Biase, I., Chutake, Y.K., Rindler, P.M., and Bidichandani, S.I. (2009). Epigenetic silencing in Friedreich ataxia is associated with depletion of CTCF (CCCTC-binding factor) and antisense transcription. PLoS One *4*, e7914.

De Biase, I., Rasmussen, A., Endres, D., Al-Mahdawi, S., Monticelli, A., Cocozza, S., Pook, M., and Bidichandani, S.I. (2007a). Progressive GAA expansions in dorsal root ganglia of Friedreich's ataxia patients. Ann Neurol *61*, 55-60.

De Biase, I., Rasmussen, A., Monticelli, A., Al-Mahdawi, S., Pook, M., Cocozza, S., and Bidichandani, S.I. (2007b). Somatic instability of the expanded GAA triplet-repeat sequence in Friedreich ataxia progresses throughout life. Genomics *90*, 1-5.

De Michele, G., Cavalcanti, F., Criscuolo, C., Pianese, L., Monticelli, A., Filla, A., and Cocozza, S. (1998). Parental gender, age at birth and expansion length influence GAA repeat intergenerational instability in the X25 gene: pedigree studies and analysis of sperm from patients with Friedreich's ataxia. Hum Mol Genet *7*, 1901-1906.

De Temmerman, N., Sermon, K., Seneca, S., De Rycke, M., Hilven, P., Lissens, W., Van Steirteghem, A., and Liebaers, I. (2004). Intergenerational instability of the expanded CTG repeat in the DMPK gene: studies in human gametes and preimplantation embryos. Am J Hum Genet *75*, 325-329.

Delatycki, M.B., Paris, D., Gardner, R.J., Forshaw, K., Nicholson, G.A., Nassif, N., Williamson, R., and Forrest, S.M. (1998). Sperm DNA analysis in a Friedreich ataxia premutation carrier suggests both meiotic and mitotic expansion in the FRDA gene. J Med Genet *35*, 713-716.

Della-Maria, J., Zhou, Y., Tsai, M.S., Kuhnlein, J., Carney, J.P., Paull, T.T., and Tomkinson, A.E. (2011). Human Mre11/Human Rad50/Nbs1 and DNA Ligase III{alpha}/XRCC1 Protein Complexes Act Together in an Alternative Nonhomologous End Joining Pathway. J Biol Chem *286*, 33845-33853.

Devys, D., Biancalana, V., Rousseau, F., Boue, J., Mandel, J.L., and Oberle, I. (1992). Analysis of full fragile X mutations in fetal tissues and monozygotic twins indicate that abnormal methylation and somatic heterogeneity are established early in development. Am J Med Genet *43*, 208-216.

Dianov, G., Bischoff, C., Piotrowski, J., and Bohr, V.A. (1998). Repair pathways for processing of 8-oxoguanine in DNA by mammalian cell extracts. J Biol Chem *273*, 33811-33816.

Dianov, G., Price, A., and Lindahl, T. (1992). Generation of single-nucleotide repair patches following excision of uracil residues from DNA. Mol Cell Biol *12*, 1605-1612.

Dianov, G.L., Prasad, R., Wilson, S.H., and Bohr, V.A. (1999). Role of DNA polymerase beta in the excision step of long patch mammalian base excision repair. J Biol Chem *274*, 13741-13743.

Dianov, G.L., Thybo, T., Dianova, II, Lipinski, L.J., and Bohr, V.A. (2000). Single nucleotide patch base excision repair is the major pathway for removal of thymine glycol from DNA in human cell extracts. J Biol Chem *275*, 11809-11813.

Dianova, II, Sleeth, K.M., Allinson, S.L., Parsons, J.L., Breslin, C., Caldecott, K.W., and Dianov, G.L. (2004). XRCC1-DNA polymerase beta interaction is required for efficient base excision repair. Nucleic Acids Res *32*, 2550-2555.

Dion, V., Lin, Y., Hubert, L., Jr., Waterland, R.A., and Wilson, J.H. (2008). Dnmt1 deficiency promotes CAG repeat expansion in the mouse germline. Hum Mol Genet *17*, 1306-1317.

Dixon, M.J., and Lahue, R.S. (2002). Examining the potential role of DNA polymerases eta and zeta in triplet repeat instability in yeast. DNA Repair (Amst) *1*, 763-770.

Dogliotti, E., Fortini, P., Pascucci, B., and Parlanti, E. (2001). The mechanism of switching among multiple BER pathways. Prog Nudeic Acid Res Mol Biol *68*, 3-27.

Dragileva, E., Hendricks, A., Teed, A., Gillis, T., Lopez, E.T., Friedberg, E.C., Kucherlapati, R., Edelmann, W., Lunetta, K.L., MacDonald, M.E.*, et al.* (2009). Intergenerational and striatal CAG repeat instability in Huntington's disease knock-in mice involve different DNA repair genes. Neurobiol Dis *33*, 37-47.

Duyao, M., Ambrose, C., Myers, R., Novelletto, A., Persichetti, F., Frontali, M., Folstein, S., Ross, C., Franz, M., Abbott, M.*, et al.* (1993). Trinucleotide repeat length instability and age of onset in Huntington's disease. Nat Genet *4*, 387-392.

Eichler, E.E., Holden, J.J., Popovich, B.W., Reiss, A.L., Snow, K., Thibodeau, S.N., Richards, C.S., Ward, P.A., and Nelson, D.L. (1994). Length of uninterrupted CGG repeats determines instability in the FMR1 gene. Nat Genet *8*, 88-94.

Engelward, B.P., Weeda, G., Wyatt, M.D., Broekhof, J.L., de Wit, J., Donker, I., Allan, J.M., Gold, B., Hoeijmakers, J.H., and Samson, L.D. (1997). Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase. Proc Natl Acad Sci U S A *94*, 13087-13092.

Erzberger, J.P., Barsky, D., Scharer, O.D., Colvin, M.E., and Wilson, D.M., 3rd (1998). Elements in abasic site recognition by the major human and Escherichia coli apurinic/apyrimidinic endonucleases. Nucleic Acids Res *26*, 2771-2778.

Fan, J., and Wilson, D.M., 3rd (2005). Protein-protein interactions and posttranslational modifications in mammalian base excision repair. Free Radic Biol Med *38*, 1121-1138.

Figueroa, A.A., Cattie, D., and Delaney, S. (2011). Structure of even/odd trinucleotide repeat sequences modulates persistence of non-B conformations and conversion to duplex. Biochemistry *50*, 4441-4450.

Filippova, G.N., Fagerlie, S., Klenova, E.M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P.E., Collins, S.J., and Lobanenkov, V.V. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. Mol Cell Biol *16*, 2802-2813.

Filippova, G.N., Thienes, C.P., Penn, B.H., Cho, D.H., Hu, Y.J., Moore, J.M., Klesert, T.R., Lobanenkov, V.V., and Tapscott, S.J. (2001). CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. Nat Genet *28*, 335-343.

Foiry, L., Dong, L., Savouret, C., Hubert, L., te Riele, H., Junien, C., and Gourdon, G. (2006). Msh3 is a limiting factor in the formation of intergenerational CTG expansions in DM1 transgenic mice. Hum Genet *119*, 520-526.

Fortini, P., and Dogliotti, E. (2007). Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways. DNA Repair (Amst) *6*, 398-409.

Fortune, M.T., Vassilopoulos, C., Coolbaugh, M.I., Siciliano, M.J., and Monckton, D.G. (2000). Dramatic, expansion-biased, age-dependent, tissue-specific somatic mosaicism in a transgenic mouse model of triplet repeat instability. Hum Mol Genet *9*, 439-445.

Fousteri, M., and Mullenders, L.H. (2008). Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects. Cell Res *18*, 73-84.

Freudenreich, C.H., Kantrow, S.M., and Zakian, V.A. (1998). Expansion and length-dependent fragility of CTG repeats in yeast. Science *279*, 853-856.

Fry, M., and Loeb, L.A. (1994). The fragile X syndrome d(CGG)n nucleotide repeats form a stable tetrahelical structure. Proc Natl Acad Sci U S A *91*, 4950-4954.

Fu, Y.H., Kuhl, D.P., Pizzuti, A., Pieretti, M., Sutcliffe, J.S., Richards, S., Verkerk, A.J., Holden, J.J., Fenwick, R.G., Jr., Warren, S.T., *et al.* (1991). Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell *67*, 1047-1058.

Gacy, A.M., Goellner, G., Juranic, N., Macura, S., and Mcmurray, C.T. (1995). Trinucleotide repeats that expand in human disease form hairpin structures in vitro. Cell *81*, 533-540.

Gacy, A.M., Goellner, G.M., Spiro, C., Chen, X., Gupta, G., Bradbury, E.M., Dyer, R.B., Mikesell, M.J., Yao, J.Z., Johnson, A.J., *et al.* (1998). GAA instability in Friedreich's Ataxia shares a common, DNA-directed and intraallelic mechanism with other trinucleotide diseases. Mol Cell *1*, 583-593.

Gary, R., Kim, K., Cornelius, H.L., Park, M.S., and Matsumoto, Y. (1999). Proliferating cell nuclear antigen facilitates excision in long-patch base excision repair. J Biol Chem *274*, 4354-4363.

Gatchel, J.R., and Zoghbi, H.Y. (2005). Diseases of unstable repeat expansion: mechanisms and common principles. Nat Rev Genet *6*, 743-755.

Gauthier, L.R., Charrin, B.C., Borrell-Pages, M., Dompierre, J.P., Rangone, H., Cordelieres, F.P., De Mey, J., MacDonald, M.E., Lessmann, V., Humbert, S., *et al.* (2004). Huntingtin controls neurotrophic support and survival of neurons by enhancing BDNF vesicular transport along microtubules. Cell *118*, 127-138.

Gembka, A., Toueille, M., Smirnova, E., Poltz, R., Ferrari, E., Villani, G., and Hubscher, U. (2007). The checkpoint clamp, Rad9-Rad1-Hus1 complex, preferentially stimulates the activity of apurinic/apyrimidinic endonuclease 1 and DNA polymerase beta in long patch base excision repair. Nucleic Acids Res *35*, 2596-2608.

Genschel, J., Littman, S.J., Drummond, J.T., and Modrich, P. (1998). Isolation of MutSbeta from human cells and comparison of the mismatch repair specificities of MutSbeta and MutSalpha. J Biol Chem *273*, 19895-19901.

Gil, J.M., and Rego, A.C. (2008). Mechanisms of neurodegeneration in Huntington's disease. Eur J Neurosci *27*, 2803-2820.

Godde, J.S., Kass, S.U., Hirst, M.C., and Wolffe, A.P. (1996). Nucleosome assembly on methylated CGG triplet repeats in the fragile X mental retardation gene 1 promoter. J Biol Chem *271*, 24325-24328.

Godde, J.S., and Wolffe, A.P. (1996). Nucleosome assembly on CTG triplet repeats. J Biol Chem *271*, 15222-15229.

Gomes-Pereira, M., Fortune, M.T., Ingram, L., McAbney, J.P., and Monckton, D.G. (2004). Pms2 is a genetic enhancer of trinucleotide CAG.CTG repeat somatic mosaicism: implications for the mechanism of triplet repeat expansion. Hum Mol Genet *13*, 1815-1825.

Gonitel, R., Moffitt, H., Sathasivam, K., Woodman, B., Detloff, P.J., Faull, R.L., and Bates, G.P. (2008). DNA instability in postmitotic neurons. Proc Natl Acad Sci U S A *105*, 3467-3472.

Goodman, F.R., Bacchelli, C., Brady, A.F., Brueton, L.A., Fryns, J.P., Mortlock, D.P., Innis, J.W., Holmes, L.B., Donnenfeld, A.E., Feingold, M., *et al.* (2000). Novel HOXA13 mutations and the phenotypic spectrum of hand-foot-genital syndrome. Am J Hum Genet *67*, 197-202.

Goula, A.V., Berquist, B.R., Wilson, D.M., 3rd, Wheeler, V.C., Trottier, Y., and Merienne, K. (2009). Stoichiometry of base excision repair proteins correlates with increased somatic CAG instability in striatum over cerebellum In Huntington's disease transgenic mice. PLoS Genet *5*, e1000749.

Gouw, L.G., Castaneda, M.A., McKenna, C.K., Digre, K.B., Pulst, S.M., Perlman, S., Lee, M.S., Gomez, C., Fischbeck, K., Gagnon, D*., et al.* (1998). Analysis of the dynamic mutation in the SCA7 gene shows marked parental effects on CAG repeat transmission. Hum Mol Genet *7*, 525-532.

Grabczyk, E., and Usdin, K. (2000a). Alleviating transcript insufficiency caused by Friedreich's ataxia triplet repeats. Nucleic Acids Res *28*, 4930-4937.

Grabczyk, E., and Usdin, K. (2000b). The GAA*TTC triplet repeat expanded in Friedreich's ataxia impedes transcription elongation by T7 RNA polymerase in a length and supercoil dependent manner. Nucleic Acids Res *28*, 2815-2822.

Greene, E., Mahishi, L., Entezam, A., Kumari, D., and Usdin, K. (2007). Repeat-induced epigenetic changes in intron 1 of the frataxin gene and its consequences in Friedreich ataxia. Nucleic Acids Res *35*, 3383-3390.

Groh, W.J., Groh, M.R., Shen, C., Monckton, D.G., Bodkin, C.L., and Pascuzzi, R.M. (2011). Survival and CTG repeat expansion in adults with myotonic dystrophy type 1. Muscle Nerve *43*, 648-651.

Groh, W.J., Lowe, M.R., and Zipes, D.P. (2002). Severity of cardiac conduction involvement and arrhythmias in myotonic dystrophy type 1 correlates with age and CTG repeat length. J Cardiovasc Electrophysiol *13*, 444-448.

Guiraud-Dogan, C., Huguet, A., Gomes-Pereira, M., Brisson, E., Bassez, G., Junien, C., and Gourdon, G. (2007). DM1 CTG expansions affect insulin receptor isoforms expression in various tissues of transgenic mice. Biochim Biophys Acta *1772*, 1183-1191.

Gunawardena, S., and Goldstein, L.S. (2005). Polyglutamine diseases and transport problems: deadly traffic jams on neuronal highways. Arch Neurol *62*, 46-51.

Gunawardena, S., Her, L.S., Brusch, R.G., Laymon, R.A., Niesman, I.R., Gordesky-Gold, B., Sintasath, L., Bonini, N.M., and Goldstein, L.S. (2003). Disruption of axonal transport by loss of huntingtin or expression of pathogenic polyQ proteins in Drosophila. Neuron *40*, 25-40.

Guyenet SJ, L.S.A. (2005). Encyclopedia of Molecular Cell Biology and Molecular Medecine, Vol 15 (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim).

Hagerman, P.J. (2008). The fragile X prevalence paradox. J Med Genet *45*, 498-499.

Han, I., You, Y., Kordower, J.H., Brady, S.T., and Morfini, G.A. (2010). Differential vulnerability of neurons in Huntington's disease: the role of cell type-specific features. J Neurochem *113*, 1073-1091.

Hanawalt, P.C., and Spivak, G. (2008). Transcription-coupled DNA repair: two decades of progress and surprises. Nat Rev Mol Cell Biol *9*, 958-970.

Hands, S.L., and Wyttenbach, A. (2011). Neurotoxic protein oligomerisation associated with polyglutamine diseases. Acta Neuropathol *120*, 419-437.

Harper, P.S. (1992). The epidemiology of Huntington's disease. HumGenet *89*, 365-376.

Harrison, C., Ketchen, A.M., Redhead, N.J., O'Sullivan, M.J., and Melton, D.W. (2002). Replication failure, genome instability, and increased cancer susceptibility in mice with a point mutation in the DNA ligase I gene. Cancer Res *62*, 4065-4074.

Hartenstine, M.J., Goodman, M.F., and Petruska, J. (2002). Weak strand displacement activity enables human DNA polymerase beta to expand CAG/CTG triplet repeats at strand breaks. J Biol Chem *277*, 41379-41389.

Hashida, H., Goto, J., Suzuki, T., Jeong, S., Masuda, N., Ooie, T., Tachiiri, Y., Tsuchiya, H., and Kanazawa, I. (2001). Single cell analysis of CAG repeat in brains of dentatorubral-pallidoluysian atrophy (DRPLA). J Neurol Sci *190*, 87-93.

Hazra, T.K., Izumi, T., Boldogh, I., Imhoff, B., Kow, Y.W., Jaruga, P., Dizdaroglu, M., and Mitra, S. (2002a). Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. Proc Natl Acad Sci U S A *99*, 3523-3528.

Hazra, T.K., Kow, Y.W., Hatahet, Z., Imhoff, B., Boldogh, I., Mokkapati, S.K., Mitra, S., and Izumi, T. (2002b). Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions. J Biol Chem *277*, 30417-30420.

Hebert, M.L., Spitz, L.A., and Wells, R.D. (2004). DNA double-strand breaks induce deletion of CTG.CAG repeats in an orientation-dependent manner in Escherichia coli. J Mol Biol *336*, 655-672.

Hebert, M.L., and Wells, R.D. (2005). Roles of double-strand breaks, nicks, and gaps in stimulating deletions of CTG.CAG repeats by intramolecular DNA repair. J Mol Biol *353*, 961-979.

Heidenfelder, B.L., Makhov, A.M., and Topal, M.D. (2003). Hairpin formation in Friedreich's ataxia triplet repeat expansion. J Biol Chem *278*, 2425-2431.

Heilig, R., Oberle, I., Arveiler, B., Hanauer, A., Vidaud, M., and Mandel, J.L. (1988). Improved DNA markers for efficient analysis of fragile X families. Am J Med Genet *30*, 543-550.

Henricksen, L.A., Tom, S., Liu, Y., and Bambara, R.A. (2000). Inhibition of flap endonuclease 1 by flap secondary structure and relevance to repeat sequence expansion. J Biol Chem *275*, 16420-16427.

Herman, D., Jenssen, K., Burnett, R., Soragni, E., Perlman, S.L., and Gottesfeld, J.M. (2006). Histone deacetylase inhibitors reverse gene silencing in Friedreich's ataxia. Nat Chem Biol *2*, 551-558.

Hoem, G., Raske, C.R., Garcia-Arocena, D., Tassone, F., Sanchez, E., Ludwig, A.L., Iwahashi, C.K., Kumar, M., Yang, J.E., and Hagerman, P.J. (2011). CGG-repeat length threshold for FMR1 RNA pathogenesis in a cellular model for FXTAS. Hum Mol Genet *20*, 2161-2170.

Hornstra, I.K., Nelson, D.L., Warren, S.T., and Yang, T.P. (1993). High resolution methylation analysis of the FMR1 gene trinucleotide repeat region in fragile X syndrome. Hum Mol Genet *2*, 1659-1665.

Horton, J.K., Prasad, R., Hou, E., and Wilson, S.H. (2000). Protection against methylation-induced cytotoxicity by DNA polymerase beta-dependent long patch base excision repair. J Biol Chem *275*, 2211-2218.

Hou, C., Chan, N.L., Gu, L., and Li, G.M. (2009). Incision-dependent and error-free repair of (CAG)(n)/(CTG)(n) hairpins in human cell extracts. Nat Struct Mol Biol *16*, 869-875.

Hou, C., Zhang, T., Tian, L., Huang, J., Gu, L., and Li, G.M. (2011). The Role of XPG in Processing (CAG)n/(CTG)n DNA Hairpins. Cell Biosci *1*, 11.

Huang, J.C., Hsu, D.S., Kazantsev, A., and Sancar, A. (1994). Substrate spectrum of human excinuclease: repair of abasic sites, methylated bases, mismatches, and bulky adducts. Proc Natl Acad Sci U S A *91*, 12213-12217.

Hubert, L., Jr., Lin, Y., Dion, V., and Wilson, J.H. (2011). Xpa deficiency reduces CAG trinucleotide repeat instability in neuronal tissues in a mouse model of SCA1. Hum Mol Genet *20*, 4822-4830.

Hubscher, U., Kuenzle, C.C., and Spadari, S. (1977). Variation of DNA polymerases-alpha, -beta. and -gamma during perinatal tissue growth and differentiation. Nucleic Acids Res *4*, 2917-2929.

Hughes, A., and Rando, O.J. (2009). Chromatin 'programming' by sequence--is there more to the nucleosome code than %GC? J Biol *8*, 96.

Ide, H., and Kotera, M. (2004). Human DNA glycosylases involved in the repair of oxidatively damaged DNA. Biol Pharm Bull *27*, 480-485.

Ikeda, Y., Daughters, R.S., and Ranum, L.P. (2008). Bidirectional expression of the SCA8 expansion mutation: one mutation, two genes. Cerebellum *7*, 150-158.

Ikeuchi, T., Koide, R., Tanaka, H., Onodera, O., Igarashi, S., Takahashi, H., Kondo, R., Ishikawa, A., Tomoda, A., Miike, T.*, et al.* (1995a). Dentatorubral-pallidoluysian atrophy: Clinical features are closely related to unstable expansions of trinucleotide (CAG) repeat. Annals of Neurology *37*, 769-775.

Ikeuchi, T., Onodera, O., Oyake, M., Koide, R., Tanaka, H., and Tsuji, S. (1995b). Dentatorubral-pallidoluysian atrophy (DRPLA): Close correlation of CAG repeat expansions with the wide spectrum of clinical presentations and prominent anticipation. SeminCell Biol *6*, 37-44.

Imam, S.Z., Karahalil, B., Hogue, B.A., Souza-Pinto, N.C., and Bohr, V.A. (2006). Mitochondrial and nuclear DNA-repair capacity of various brain regions in mouse is altered in an age-dependent manner. Neurobiol Aging *27*, 1129-1136.

Innis, J.W., Mortlock, D., Chen, Z., Ludwig, M., Williams, M.E., Williams, T.M., Doyle, C.D., Shao, Z., Glynn, M., Mikulic, D.*, et al.* (2004). Polyalanine expansion in HOXA13: three new affected families and the molecular consequences in a mouse model. Hum Mol Genet *13*, 2841-2851.

Intano, G.W., Cho, E.J., McMahan, C.A., and Walter, C.A. (2003). Age-related base excision repair activity in mouse brain and liver nuclear extracts. J Gerontol A Biol Sci Med Sci *58*, 205-211.

Ishiguro, H., Yamada, K., Sawada, H., Nishii, K., Ichino, N., Sawada, M., Kurosawa, Y., Matsushita, N., Kobayashi, K., Goto, J.*, et al.* (2001). Age-dependent and tissue-specific CAG repeat instability occurs in mouse knock-in for a mutant Huntington's disease gene. J Neurosci Res *65*, 289-297.

Ishii, S., Nishio, T., Sunohara, N., Yoshihara, T., Takemura, K., Hikiji, K., Tsujino, S., and Sakuragawa, N. (1996). Small increase in triplet repeat length of cerebellum from patients with myotonic dystrophy. Hum Genet *98*, 138-140.

Itaya, M. (1990). Isolation and characterization of a second RNase H (RNase HII) of Escherichia coli K-12 encoded by the rnhB gene. Proc Natl Acad Sci U S A *87*, 8587-8591.

Izumi, T., Hazra, T.K., Boldogh, I., Tomkinson, A.E., Park, M.S., Ikeda, S., and Mitra, S. (2000). Requirement for human AP endonuclease 1 for repair of 3'-blocking damage at DNA single-strand breaks induced by reactive oxygen species. Carcinogenesis *21*, 1329-1334.

Jankowski, C., Nasar, F., and Nag, D.K. (2000). Meiotic instability of CAG repeat tracts occurs by double-strand break repair in yeast. Proc Natl Acad Sci U S A *97*, 2134-2139.

Jansen, G., Willems, P., Coerwinkel, M., Nillesen, W., Smeets, H., Vits, L., Howeler, C., Brunner, H., and Wieringa, B. (1994). Gonosomal mosaicism in myotonic dystrophy patients: involvement of mitotic events in (CTG)n repeat variation and selection against extreme expansion in sperm. Am J Hum Genet *54*, 575-585.

Jarem, D.A., and Delaney, S. (2011). Premutation huntingtin allele adopts a non-B conformation and contains a hot spot for DNA damage. Biochem Biophys Res Commun *416*, 146-152.

Jarem, D.A., Wilson, N.R., and Delaney, S. (2009). Structure-dependent DNA damage and repair in a trinucleotide repeat sequence. Biochemistry *48*, 6655-6663.

Jarem, D.A., Wilson, N.R., Schermerhorn, K.M., and Delaney, S. (2011). Incidence and persistence of 8-oxo-7,8-dihydroguanine within a hairpin intermediate exacerbates a toxic oxidation cycle associated with trinucleotide repeat expansion. DNA Repair (Amst) *10*, 887-896.

Jedele, K.B., Wahl, D., Chahrokh-Zadeh, S., Wirtz, A., Murken, J., and Holinski-Feder, E. (1998). Spinal and bulbar muscular atrophy (SBMA): somatic stability of an expanded CAG repeat in fetal tissues. Clin Genet *54*, 148-151.

Jin, P., Zarnescu, D.C., Zhang, F., Pearson, C.E., Lucchesi, J.C., Moses, K., and Warren, S.T. (2003). RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in Drosophila. Neuron *39*, 739-747.

Jung, J., and Bonini, N. (2007). CREB-binding protein modulates repeat instability in a Drosophila model for polyQ disease. Science *315*, 1857-1859.

Kahlem, P., and Djian, P. (2000). The expanded CAG repeat associated with juvenile Huntington disease shows a common origin of most or all neurons and glia in human cerebrum. Neurosci Lett *286*, 203-207.

Kang, S., Jaworski, A., Ohshima, K., and Wells, R.D. (1995). Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in E. coli. Nat Genet *10*, 213-218.

Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J., *et al.* (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. Nature *458*, 362-366.

Karahalil, B., Hogue, B.A., de Souza-Pinto, N.C., and Bohr, V.A. (2002). Base excision repair capacity in mitochondria and nuclei: tissue-specific variations. Faseb J *16*, 1895-1902.

Kato, T., Tanaka, F., Yamamoto, M., Yosida, E., Indo, T., Watanabe, H., Yoshiwara, T., Doyu, M., and Sobue, G. (2000). Sisters homozygous for the spinocerebellar ataxia type 6 (SCA6)/CACNA1A gene associated with different clinical phenotypes. Clin Genet *58*, 69-73.

Kaytor, M.D., Burright, E.N., Duvick, L.A., Zoghbi, H.Y., and Orr, H.T. (1997). Increased trinucleotide repeat instability with advanced maternal age. Hum Mol Genet *6*, 2135-2139.

Kennedy, L., Evans, E., Chen, C.M., Craven, L., Detloff, P.J., Ennis, M., and Shelbourne, P.F. (2003). Dramatic tissue-specific mutation length increases are an early molecular event in Huntington disease pathogenesis. Hum Mol Genet *12*, 3359-3367.

Kennedy, L., and Shelbourne, P.F. (2000). Dramatic mutation instability in HD mouse striatum: does polyglutamine load contribute to cell-specific vulnerability in Huntington's disease? Hum Mol Genet *9*, 2539-2544.

Kerrest, A., Anand, R.P., Sundararajan, R., Bermejo, R., Liberi, G., Dujon, B., Freudenreich, C.H., and Richard, G.F. (2009). SRS2 and SGS1 prevent chromosomal breaks and stabilize triplet repeats by restraining recombination. Nat Struct Mol Biol *16*, 159-167.

Kiliszek, A., Kierzek, R., Krzyzosiak, W.J., and Rypniewski, W. (2011). Crystal structures of CGG RNA repeats with implications for fragile X-associated tremor ataxia syndrome. Nucleic Acids Res *39*, 7308-7315.

Kim, H.M., Narayanan, V., Mieczkowski, P.A., Petes, T.D., Krasilnikova, M.M., Mirkin, S.M., and Lobachev, K.S. (2008). Chromosome fragility at GAA tracts in yeast depends on repeat orientation and requires mismatch repair. Embo J *27*, 2896-2906.

Kirkpatrick, D.T., and Petes, T.D. (1997). Repair of DNA loops involves DNA-mismatch and nucleotide-excision repair proteins. Nature *387*, 929-931.

Klungland, A., Hoss, M., Gunz, D., Constantinou, A., Clarkson, S.G., Doetsch, P.W., Bolton, P.H., Wood, R.D., and Lindahl, T. (1999). Base excision repair of oxidative DNA damage activated by XPG protein. Mol Cell *3*, 33-42.

Klungland, A., and Lindahl, T. (1997). Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). Embo J *16*, 3341-3348.

Kovtun, I.V., Johnson, K.O., and McMurray, C.T. (2011). Cockayne syndrome B protein antagonizes OGG1 in modulating CAG repeat length in vivo. Aging (Albany NY) *3*, 509-514.

Kovtun, I.V., Liu, Y., Bjoras, M., Klungland, A., Wilson, S.H., and McMurray, C.T. (2007). OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. Nature *447*, 447-452.

Kovtun, I.V., and McMurray, C.T. (2001). Trinucleotide expansion in haploid germ cells by gap repair. Nat Genet *27*, 407-411.

Kucherlapati, M., Yang, K., Kuraguchi, M., Zhao, J., Lia, M., Heyer, J., Kane, M.F., Fan, K., Russell, R., Brown, A.M.*, et al.* (2002). Haploinsufficiency of Flap endonuclease (Fen1) leads to rapid tumor progression. Proc Natl Acad Sci U S A *99*, 9924-9929.

Kumari, D., Biacsi, R.E., and Usdin, K. (2011). Repeat expansion affects both transcription initiation and elongation in friedreich ataxia cells. J Biol Chem *286*, 4209-4215.

Ladd, P.D., Smith, L.E., Rabaia, N.A., Moore, J.M., Georges, S.A., Hansen, R.S., Hagerman, R.J., Tassone, F., Tapscott, S.J., and Filippova, G.N. (2007). An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals. Hum Mol Genet *16*, 3174-3187.

Larsen, E., Gran, C., Saether, B.E., Seeberg, E., and Klungland, A. (2003). Proliferation failure and gamma radiation sensitivity of Fen1 null mutant mice at the blastocyst stage. Mol Cell Biol *23*, 5346-5353.

Lee, B.I., and Wilson, D.M., 3rd (1999). The RAD2 domain of human exonuclease 1 exhibits 5' to 3' exonuclease and flap structure-specific endonuclease activities. J Biol Chem *274*, 37763-37769.

Leeflang, E.P., Zhang, L., Tavare, S., Hubert, R., Srinidhi, J., MacDonald, M.E., Myers, R.H., de Young, M., Wexler, N.S., Gusella, J.F.*, et al.* (1995). Single sperm analysis of the trinucleotide repeats in the Huntington's disease gene: quantification of the mutation frequency spectrum. HumMolGenet *4*, 1519-1526.

Lenzi, M.L., Smith, J., Snowden, T., Kim, M., Fishel, R., Poulos, B.K., and Cohen, P.E. (2005). Extreme heterogeneity in the molecular events leading to the establishment of chiasmata during meiosis i in human oocytes. Am J Hum Genet *76*, 112-127.

LeProust, E.M., Pearson, C.E., Sinden, R.R., and Gao, X. (2000). Unexpected formation of parallel duplex in GAA and TTC trinucleotide repeats of Friedreich's ataxia. J Mol Biol *302*, 1063-1080.

Levin, D.S., Bai, W., Yao, N., O'Donnell, M., and Tomkinson, A.E. (1997). An interaction between DNA ligase I and proliferating cell nuclear antigen: implications for Okazaki fragment synthesis and joining. Proc Natl Acad Sci U S A *94*, 12863-12868.

Levin, D.S., McKenna, A.E., Motycka, T.A., Matsumoto, Y., and Tomkinson, A.E. (2000). Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair. Curr Biol *10*, 919-922.

Li, L.B., Yu, Z., Teng, X., and Bonini, N.M. (2008). RNA toxicity is a component of ataxin-3 degeneration in Drosophila. Nature *453*, 1107-1111.

Li, X.J., Li, S.H., Sharp, A.H., Nucifora, F.C., Jr., Schilling, G., Lanahan, A., Worley, P., Snyder, S.H., and Ross, C.A. (1995). A huntingtin-associated protein enriched in brain with implications for pathology. Nature *378*, 398-402.

Lia, A.S., Seznec, H., Hofmann-Radvanyi, H., Radvanyi, F., Duros, C., Saquet, C., Blanche, M., Junien, C., and Gourdon, G. (1998). Somatic instability of the CTG repeat in mice transgenic for the myotonic dystrophy region is age dependent but not correlated to the relative intertissue transcription levels and proliferative capacities. Hum Mol Genet *7*, 1285-1291.

Libby, R.T., Hagerman, K.A., Pineda, V.V., Lau, R., Cho, D.H., Baccam, S.L., Axford, M.M., Cleary, J.D., Moore, J.M., Sopher, B.L.*, et al.* (2008). CTCF cis-regulates trinucleotide repeat instability in an epigenetic manner: a novel basis for mutational hot spot determination. PLoS Genet *4*, e1000257.

Lieber, M.R. (2008). The mechanism of human nonhomologous DNA end joining. J Biol Chem *283*, 1-5.

Lin, Y., Dent, S.Y., Wilson, J.H., Wells, R.D., and Napierala, M. (2010a). R loops stimulate genetic instability of CTG.CAG repeats. Proc Natl Acad Sci U S A *107*, 692-697.

Lin, Y., Dion, V., and Wilson, J.H. (2006). Transcription promotes contraction of CAG repeat tracts in human cells. Nat Struct Mol Biol *13*, 179-180.

Lin, Y., Leng, M., Wan, M., and Wilson, J.H. (2010b). Convergent transcription through a long CAG tract destabilizes repeats and induces apoptosis. Mol Cell Biol *30*, 4435-4451.

Lin, Y., and Wilson, J.H. (2007). Transcription-induced CAG repeat contraction in human cells is mediated in part by transcription-coupled nudeotide excision repair. Mol Cell Biol *27*, 6209-6217.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature *362*, 709-715.

Lindahl, T. (2000). Suppression of spontaneous mutagenesis in human cells by DNA base excision-repair. Mutat Res *462*, 129-135.

Littman, S.J., Fang, W.H., and Modrich, P. (1999). Repair of large insertion/deletion heterologies in human nudear extracts is directed by a 5' single-strand break and is independent of the mismatch repair system. J Biol Chem *274*, 7474-7481.

Liu, G., Chen, X., Bissler, J.J., Sinden, R.R., and Leffak, M. (2010a). Replication-dependent instability at (CTG) x (CAG) repeat hairpins in human cells. Nat Chem Biol *6*, 652-659.

Liu, Y., Beard, W.A., Shock, D.D., Prasad, R., Hou, E.W., and Wilson, S.H. (2005). DNA polymerase beta and flap endonuclease 1 enzymatic specificities sustain DNA synthesis for long patch base excision repair. J Biol Chem *280*, 3665-3674.

Liu, Y., Prasad, R., Beard, W.A., Hou, E.W., Horton, J.K., McMurray, C.T., and Wilson, S.H. (2009). Coordination between polymerase beta and FEN1 can modulate CAG repeat expansion. J Biol Chem *284*, 28352-28366.

Liu, Y., Prasad, R., Beard, W.A., Kedar, P.S., Hou, E.W., Shock, D.D., and Wilson, S.H. (2007). Coordination of steps in single-nucleotide base excision repair mediated by apurinic/apyrimidinic endonuclease 1 and DNA polymerase beta. J Biol Chem *282*, 13532-13541.

Liu, Y., Prasad, R., and Wilson, S.H. (2010b). HMGB1: roles in base excision repair and related function. Biochim Biophys Acta *1799*, 119-130.

Liu, Y., Zhang, H., Veeraraghavan, J., Bambara, R.A., and Freudenreich, C.H. (2004). Saccharomyces cerevisiae flap endonuclease 1 uses flap equilibration to maintain triplet repeat stability. Mol Cell Biol *24*, 4049-4064.

Loesch, D.Z., Godler, D.E., Evans, A., Bui, Q.M., Gehling, F., Kotschet, K.E., Trost, N., Storey, E., Stimpson, P., Kinsella, G.*, et al.* (2011). Evidence for the toxicity of bidirectional transcripts and mitochondrial dysfunction in blood associated with small CGG expansions in the FMR1 gene in patients with parkinsonism. Genet Med *13*, 392-399.

Lopes-Cendes, I., Maciel, P., Kish, S., Gaspar, C., Robitaille, Y., Clark, H.B., Koeppen, A.H., Nance, M., Schut, L., Silveira, I.*, et al.* (1996). Somatic mosaicism in the central nervous system in spinocerebellar ataxia type 1 and Machado-Joseph disease. Ann Neurol *40*, 199-206.

Lopez Castel, A., Cleary, J.D., and Pearson, C.E. (2010). Repeat instability as the basis for human diseases and as a potential target for therapy. Nat Rev Mol Cell Biol *11*, 165-170.

Lopez Castel, A., Nakamori, M., Tome, S., Chitayat, D., Gourdon, G., Thornton, C.A., and Pearson, C.E. (2011). Expanded CTG repeat demarcates a boundary for abnormal CpG methylation in myotonic dystrophy patient tissues. Hum Mol Genet *20*, 1-15.

Lopez Castel, A., Tomkinson, A.E., and Pearson, C.E. (2009). CTG/CAG repeat instability is modulated by the levels of human DNA ligase i and its interaction with proliferating cell nuclear antigen: a distinction between replication and slipped-DNA repair. J Biol Chem *284*, 26631-26645.

Lorincz, M.C., Dickerson, D.R., Schmitt, M., and Groudine, M. (2004). Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. Nat Struct Mol Biol *11*, 1068-1075.

Malter, H.E., Iber, J.C., Willemsen, R., de Graaff, E., Tarleton, J.C., Leisti, J., Warren, S.T., and Oostra, B.A. (1997). Characterization of the full fragile X syndrome mutation in fetal gametes. Nat Genet *15*, 165-169.

Mangiarini, L., Sathasivam, K., Mahal, A., Mott, R., Seller, M., and Bates, G.P. (1997). Instability of highly expanded CAG repeats in mice transgenic for the Huntington's disease mutation [see comments]. Nat Genet *15*, 197-200.

Mangiarini, L., Sathasivam, K., Seller, M., Cozens, B., Harper, A., Hetherington, C., Lawton, M., Trottier, Y., Lehrach, H., Davies, S.W.*, et al.* (1996a). Exon 1 of the *HD* gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. Cell *87*, 493-506.

Mangiarini, L., Sathasivam, K., Seller, M., Cozens, B., Harper, A., Hetherington, C., Lawton, M., Trottier, Y., Lehrach, H., Davies, S.W.*, et al.* (1996b). Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. Cell *87*, 493-506.

Manley, K., Shirley, T.L., Flaherty, L., and Messer, A. (1999). Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. Nat Genet *23*, 471-473.

Mao, Z., Bozzella, M., Seluanov, A., and Gorbunova, V. (2008). Comparison of nonhomologous end joining and homologous recombination in human cells. DNA Repair (Amst) *7*, 1765-1771.

Marcadier, J.L., and Pearson, C.E. (2003). Fidelity of primate cell repair of a double-strand break within a (CTG).(CAG) tract. Effect of slipped DNA structures. J Biol Chem *278*, 33848-33856.

Martorell, L., Johnson, K., Boucher, C.A., and Baiget, M. (1997). Somatic instability of the myotonic dystrophy (CTG)n repeat during human fetal development. Hum Mol Genet *6*, 877-880.

Martorell, L., Martinez, J.M., Carey, N., Johnson, K., and Baiget, M. (1995). Comparison of CTG repeat length expansion and clinical progression of myotonic dystrophy over a five year period. J Med Genet *32*, 593-596.

Martorell, L., Monckton, D.G., Gamez, J., and Baiget, M. (2000). Complex patterns of male germline instability and somatic mosaicism in myotonic dystrophy type 1. Eur J Hum Genet *8*, 423-430.

Martorell, L., Monckton, D.G., Gamez, J., Johnson, K.J., Gich, I., Lopez de Munain, A., and Baiget, M. (1998). Progression of somatic CTG repeat length heterogeneity in the blood cells of myotonic dystrophy patients. Hum Mol Genet *7*, 307-312.

Matsumoto, Y., and Kim, K. (1995). Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair. Science *269*, 699-702.

Matsuyama, Z., Kawakami, H., Maruyama, H., Izumi, Y., Komure, O., Udaka, F., Kameyama, M., Nishio, T., Kuroda, Y., Nishimura, M.*, et al.* (1997). Molecular features of the CAG repeats of spinocerebellar ataxia 6 (SCA6). Hum Mol Genet *6*, 1283-1287.

McCulloch, S.D., Gu, L., and Li, G.M. (2003). Nick-dependent and -independent processing of large DNA loops in human cells. J Biol Chem *278*, 50803-50809.

McMurray, C.T. (2010). Mechanisms of trinucleotide repeat instability during human development. Nat Rev Genet *11*, 786-799.

Merienne, K., Helmlinger, D., Perkin, G.R., Devys, D., and Trottier, Y. (2003). Polyglutamine expansion induces a protein-damaging stress connecting heat shock protein 70 to the JNK pathway. J Biol Chem *278*, 16957-16967.

Mirkin, S.M. (2007). Expandable DNA repeats and human disease. Nature *447*, 932-940.

Mitas, M., Yu, A., Dill, J., Kamp, T.J., Chambers, E.J., and Haworth, I.S. (1995). Hairpin properties of single-stranded DNA containing a GC-rich triplet repeat: (CTG)15. Nucleic Acids Res *23*, 1050-1059.

Mittelman, D., Moye, C., Morton, J., Sykoudis, K., Lin, Y., Carroll, D., and Wilson, J.H. (2009). Zinc-finger directed double-strand breaks within CAG repeat tracts promote repeat instability in human cells. Proc Natl Acad Sci U S A *106*, 9607-9612.

Mochmann, L.H., and Wells, R.D. (2004). Transcription influences the types of deletion and expansion products in an orientation-dependent manner from GAC*GTC repeats. Nucleic Acids Res *32*, 4469-4479.

Mol, C.D., Arvai, A.S., Slupphaug, G., Kavli, B., Alseth, I., Krokan, H.E., and Tainer, J.A. (1995). Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. Cell *80*, 869-878.

Monckton, D.G., Cayuela, M.L., Gould, F.K., Brock, G.J., Silva, R., and Ashizawa, T. (1999). Very large (CAG)(n) DNA repeat expansions in the sperm of two spinocerebellar ataxia type 7 males. Hum Mol Genet *8*, 2473-2478.

Monckton, D.G., Wong, L.J., Ashizawa, T., and Caskey, C.T. (1995). Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses. Hum Mol Genet *4*, 1-8.

Monros, E., Molto, M.D., Martinez, F., Canizares, J., Blanca, J., Vilchez, J.J., Prieto, F., de Frutos, R., and Palau, F. (1997). Phenotype correlation and intergenerational dynamics of the Friedreich ataxia GAA trinucleotide repeat. Am J Hum Genet *61*, 101-110.

Mortlock, D.P., Post, L.C., and Innis, J.W. (1996). The molecular basis of hypodactyly (Hd): a deletion in Hoxa 13 leads to arrest of digital arch formation. Nat Genet *13*, 284-289.

Moseley, M.L., Schut, L.J., Bird, T.D., Koob, M.D., Day, J.W., and Ranum, L.P. (2000). SCA8 CTG repeat: en masse contractions in sperm and intergenerational sequence changes may play a role in reduced penetrance. Hum Mol Genet *9*, 2125-2130.

Moseley, M.L., Zu, T., Ikeda, Y., Gao, W., Mosemiller, A.K., Daughters, R.S., Chen, G., Weatherspoon, M.R., Clark, H.B., Ebner, T.J., *et al.* (2006). Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. Nat Genet *38*, 758-769.

Moutou, C., Vincent, M.C., Biancalana, V., and Mandel, J.L. (1997). Transition from premutation to full mutation in fragile X syndrome is likely to be prezygotic. Hum Mol Genet *6*, 971-979.

Muller, J.G., Duarte, V., Hickerson, R.P., and Burrows, C.J. (1998). Gel electrophoretic detection of 7,8-dihydro-8-oxoguanine and 7, 8-dihydro-8-oxoadenine via oxidation by Ir (IV). Nucleic Acids Res *26*, 2247-2249.

Mulvihill, D.J., Nichol Edamura, K., Hagerman, K.A., Pearson, C.E., and Wang, Y.H. (2005). Effect of CAT or AGG interruptions and CpG methylation on nucleosome assembly upon trinucleotide repeats on spinocerebellar ataxia, type 1 and fragile X syndrome. J Biol Chem *280*, 4498-4503.

Muragaki, Y., Mundlos, S., Upton, J., and Olsen, B.R. (1996). Altered growth and branching patterns in synpolydactyly caused by mutations in HOXD13 [see comments]. Science *272*, 548-551.

Nakamori, M., Pearson, C.E., and Thornton, C.A. (2011). Bidirectional transcription stimulates expansion and contraction of expanded (CTG)*(CAG) repeats. Hum Mol Genet *20*, 580-588.

Napierala, M., and Krzyzosiak, W.J. (1997). CUG repeats present in myotonin kinase RNA form metastable "slippery" hairpins. J Biol Chem *272*, 31079-31085.

Narciso, L., Fortini, P., Pajalunga, D., Franchitto, A., Liu, P., Degan, P., Frechet, M., Demple, B., Crescenzi, M., and Dogliotti, E. (2007). Terminally differentiated muscle cells are defective in base excision DNA repair and hypersensitive to oxygen injury. Proc Natl Acad Sci U S A *104*, 17010-17015.

Nemes, J.P., Benzow, K.A., Moseley, M.L., Ranum, L.P., and Koob, M.D. (2000). The SCA8 transcript is an antisense RNA to a brain-specific transcript encoding a novel actin-binding protein (KLHL1). Hum Mol Genet *9*, 1543-1551.

Nguyen, L.H., Barsky, D., Erzberger, J.P., and Wilson, D.M., 3rd (2000). Mapping the protein-DNA interface and the metal-binding site of the major human apurinic/apyrimidinic endonuclease. J Mol Biol *298*, 447-459.

Nolin, S.L., Houck, G.E., Jr., Gargano, A.D., Blumstein, H., Dobkin, C.S., and Brown, W.T. (1999). FMR1 CGG-repeat instability in single sperm and lymphocytes of fragile-X premutation males. Am J Hum Genet *65*, 680-688.

Nucifora, F.C., Jr., Sasaki, M., Peters, M.F., Huang, H., Cooper, J.K., Yamada, M., Takahashi, H., Tsuji, S., Troncoso, J., Dawson, V.L., *et al.* (2001). Interference by huntingtin and atrophin-1 with cbp-mediated transcription leading to cellular toxicity. Science *291*, 2423-2428.

O'Brien, P.J., and Ellenberger, T. (2003). Human alkyladenine DNA glycosylase uses acid-base catalysis for selective excision of damaged purines. Biochemistry *42*, 12418-12429.

O'Rourke, J.R., and Swanson, M.S. (2009). Mechanisms of RNA-mediated disease. J Biol Chem *284*, 7419-7423.

Oberle, I., Rousseau, F., Heitz, D., Kretz, C., Devys, D., Hanauer, A., Boue, J., Bertheas, M.F., and Mandel, J.L. (1991). Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. Science *252*, 1097-1102.

Oei, S.L., and Ziegler, M. (2000). ATP for the DNA ligation step in base excision repair is generated from poly(ADP-ribose). J Biol Chem *275*, 23234-23239.

Osterod, M., Larsen, E., Le Page, F., Hengstler, J.G., Van Der Horst, G.T., Boiteux, S., Klungland, A., and Epe, B. (2002). A global DNA repair mechanism involving the Cockayne syndrome B (CSB) gene product can prevent the in vivo accumulation of endogenous oxidative DNA base damage. Oncogene *21*, 8232-8239.

Oussatcheva, E.A., Hashem, V.I., Zou, Y., Sinden, R.R., and Potaman, V.N. (2001). Involvement of the nucleotide excision repair protein UvrA in instability of CAG*CTG repeat sequences in Escherichia coli. J Biol Chem *276*, 30878-30884.

Owen, B.A., Yang, Z., Lai, M., Gajec, M., Badger, J.D., 2nd, Hayes, J.J., Edelmann, W., Kucherlapati, R., Wilson, T.M., and McMurray, C.T. (2005). (CAG)(n)-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. Nat Struct Mol Biol *12*, 663-670.

Panigrahi, G.B., Lau, R., Montgomery, S.E., Leonard, M.R., and Pearson, C.E. (2005). Slipped (CTG)*(CAG) repeats can be correctly repaired, escape repair or undergo error-prone repair. Nat Struct Mol Biol *12*, 654-662.

Panigrahi, G.B., Slean, M.M., Simard, J.P., Gileadi, O., and Pearson, C.E. (2010). Isolated short CTG/CAG DNA slip-outs are repaired efficiently by hMutSbeta, but clustered slip-outs are poorly repaired. Proc Natl Acad Sci U S A *107*, 12593-12598.

Parniewski, P., Bacolla, A., Jaworski, A., and Wells, R.D. (1999). Nucleotide excision repair affects the stability of long transcribed (CTG*CAG) tracts in an orientation-dependent manner in Escherichia coli. Nucleic Acids Res *27*, 616-623.

Parsons, J.L., Dianova, II, Allinson, S.L., and Dianov, G.L. (2005). DNA polymerase beta promotes recruitment of DNA ligase III alpha-XRCC1 to sites of base excision repair. Biochemistry *44*, 10613-10619.

Parsons, J.L., Dianova, II, and Dianov, G.L. (2004). APE1 is the major 3'-phosphoglycolate activity in human cell extracts. Nucleic Acids Res *32*, 3531-3536.

Pascucci, B., Maga, G., Hubscher, U., Bjoras, M., Seeberg, E., Hickson, I.D., Villani, G., Giordano, C., Cellai, L., and Dogliotti, E. (2002). Reconstitution of the base excision repair pathway for 7,8-dihydro-8-oxoguanine with purified human proteins. Nucleic Acids Res *30*, 2124-2130.

Pascucci, B., Stucki, M., Jonsson, Z.O., Dogliotti, E., and Hubscher, U. (1999). Long patch base excision repair with purified human proteins. DNA ligase I as patch size mediator for DNA polymerases delta and epsilon. J Biol Chem *274*, 33696-33702.

Pearson, C.E., Eichler, E.E., Lorenzetti, D., Kramer, S.F., Zoghbi, H.Y., Nelson, D.L., and Sinden, R.R. (1998a). Interruptions in the triplet repeats of SCA1 and FRAXA reduce the propensity and complexity of slipped strand DNA (S-DNA) formation. Biochemistry *37*, 2701-2708.

Pearson, C.E., Ewel, A., Acharya, S., Fishel, R.A., and Sinden, R.R. (1997). Human MSH2 binds to trinucleotide repeat DNA structures associated with neurodegenerative diseases. Hum Mol Genet *6*, 1117-1123.

Pearson, C.E., Nichol Edamura, K., and Cleary, J.D. (2005). Repeat instability: mechanisms of dynamic mutations. Nat Rev Genet *6*, 729-742.

Pearson, C.E., and Sinden, R.R. (1996). Alternative structures in duplex DNA formed within the trinucleotide repeats of the myotonic dystrophy and fragile X loci. Biochemistry *35*, 5041-5053.

Pearson, C.E., Tam, M., Wang, Y.H., Montgomery, S.E., Dar, A.C., Cleary, J.D., and Nichol, K. (2002). Slipped-strand DNAs formed by long (CAG)*(CTG) repeats: slipped-out repeats and slip-out junctions. Nucleic Acids Res *30*, 4534-4547.

Pearson, C.E., Wang, Y.H., Griffith, J.D., and Sinden, R.R. (1998b). Structural analysis of slipped-strand DNA (S-DNA) formed in (CTG)n. (CAG)n repeats from the myotonic dystrophy locus. Nucleic Acids Res *26*, 816-823.

Petermann, E., Keil, C., and Oei, S.L. (2006). Roles of DNA ligase III and XRCC1 in regulating the switch between short patch and long patch BER. DNA Repair (Amst) *5*, 544-555.

Petermann, E., Ziegler, M., and Oei, S.L. (2003). ATP-dependent selection between single nucleotide and long patch base excision repair. DNA Repair (Amst) *2*, 1101-1114.

Petrini, J.H., Xiao, Y., and Weaver, D.T. (1995). DNA ligase I mediates essential functions in mammalian cells. Mol Cell Biol *15*, 4303-4308.

Petruska, J., Arnheim, N., and Goodman, M.F. (1996). Stability of intrastrand hairpin structures formed by the CAG/CTG class of DNA triplet repeats associated with neurological diseases. Nucleic Acids Res *24*, 1992-1998.

Pirozzi, F., Tabolacci, E., and Neri, G. (2011). The FRAXopathies: definition, overview, and update. Am J Med Genet A *155A*, 1803-1816.

Podlutsky, A.J., Dianova, II, Podust, V.N., Bohr, V.A., and Dianov, G.L. (2001). Human DNA polymerase beta initiates DNA synthesis during long-patch repair of reduced AP sites in DNA. Embo J *20*, 1477-1482.

Pollard, L.M., Bourn, R.L., and Bidichandani, S.I. (2008). Repair of DNA double-strand breaks within the (GAA*TTC)n sequence results in frequent deletion of the triplet-repeat sequence. Nucleic Acids Res *36*, 489-500.

Prasad, R., Dianov, G.L., Bohr, V.A., and Wilson, S.H. (2000). FEN1 stimulation of DNA polymerase beta mediates an excision step in mammalian long patch base excision repair. J Biol Chem *275*, 4460-4466.

Prasad, R., Liu, Y., Deterding, L.J., Poltoratsky, V.P., Kedar, P.S., Horton, J.K., Kanno, S., Asagoshi, K., Hou, E.W., Khodyreva, S.N.*, et al.* (2007). HMGB1 is a cofactor in mammalian base excision repair. Mol Cell *27*, 829-841.

Prasad, R., Longley, M.J., Sharief, F.S., Hou, E.W., Copeland, W.C., and Wilson, S.H. (2009). Human DNA polymerase theta possesses 5'-dRP lyase activity and functions in single-nucleotide base excision repair in vitro. Nucleic Acids Res *37*, 1868-1877.

Rai, M., Soragni, E., Jenssen, K., Burnett, R., Herman, D., Coppola, G., Geschwind, D.H., Gottesfeld, J.M., and Pandolfo, M. (2008). HDAC inhibitors correct frataxin deficiency in a Friedreich ataxia mouse model. PLoS One *3*, e1958.

Raji, N.S., Krishna, T.H., and Rao, K.S. (2002). DNA-polymerase alpha, beta, delta and epsilon activities in isolated neuronal and astroglial cell fractions from developing and aging rat cerebral cortex. Int J Dev Neurosci *20*, 491-496.

Ranen, N.G., Stine, O.C., Abbott, M.H., Sherr, M., Codori, A.M., Franz, M.L., Chao, N.I., Chung, A.S., Pleasant, N., Callahan, C.*, et al.* (1995). Anticipation and instability of IT-15 (CAG)n repeats in parent-offspring pairs with Huntington disease. Am J Hum Genet *57*, 593-602.

Ranum, L.P., and Cooper, T.A. (2006). RNA-Mediated Neuromuscular Disorders. Annu Rev Neurosci.

Rao, K.S., Annapurna, V.V., and Raji, N.S. (2001). DNA polymerase-beta may be the main player for defective DNA repair in aging rat neurons. Ann N Y Acad Sci *928*, 113-120.

Reddy, K., Tam, M., Bowater, R.P., Barber, M., Tomlinson, M., Nichol Edamura, K., Wang, Y.H., and Pearson, C.E. (2010). Determinants of R-loop formation at convergent bidirectionally transcribed trinucleotide repeats. Nucleic Acids Res.

Reyniers, E., Vits, L., De Boulle, K., Van Roy, B., Van Velzen, D., de Graaff, E., Verkerk, A.J., Jorens, H.Z., Darby, J.K., Oostra, B.*, et al.* (1993). The full mutation in the FMR-1 gene of male fragile X patients is absent in their sperm. Nat Genet *4*, 143-146.

Richard, G.F., Dujon, B., and Haber, J.E. (1999). Double-strand break repair can lead to high frequencies of deletions within short CAG/CTG trinucleotide repeats. Mol Gen Genet *261*, 871-882.

Richard, G.F., Goellner, G.M., McMurray, C.T., and Haber, J.E. (2000). Recombination-induced CAG trinucleotide repeat expansions in yeast involve the MRE11-RAD50-XRS2 complex. EMBO J *19*, 2381-2390.

Riess, O., Laccone, F.A., Gispert, S., Schols, L., Zuhlke, C., Vieira-Saecker, A.M., Herlt, S., Wessel, K., Epplen, J.T., Weber, B.H.*, et al.* (1997). SCA2 trinucleotide expansion in German SCA patients. Neurogenetics *1*, 59-64.

Robertson, A.B., Klungland, A., Rognes, T., and Leiros, I. (2009). DNA repair in mammalian cells: Base excision repair: the long and short of it. Cell Mol Life Sci *66*, 981-993.

Roos, R.A. (2010). Huntington's disease: a clinical review. Orphanet J Rare Dis *5*, 40.

Rossi, M.L., Pike, J.E., Wang, W., Burgers, P.M., Campbell, J.L., and Bambara, R.A. (2008). Pif1 helicase directs eukaryotic Okazaki fragments toward the two-nuclease cleavage pathway for primer removal. J Biol Chem *283*, 27483-27493.

Rousseau, F., Heitz, D., Biancalana, V., Blumenfeld, S., Kretz, C., Boue, J., Tommerup, N., Van Der Hagen, C., DeLozier-Blanchet, C., Croquette, M.F.*, et al.* (1991). Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation. N Engl J Med *325*, 1673-1681.

Sakamoto, N., Chastain, P.D., Parniewski, P., Ohshima, K., Pandolfo, M., Griffith, J.D., and Wells, R.D. (1999). Sticky DNA: self-association properties of long GAA.TTC repeats in R.R.Y triplex structures from Friedreich's ataxia. Mol Cell *3*, 465-475.

Sakamoto, N., Larson, J.E., Iyer, R.R., Montermini, L., Pandolfo, M., and Wells, R.D. (2001a). GGA*TCC-interrupted triplets in long GAA*TTC repeats inhibit the formation of triplex and sticky DNA structures, alleviate transcription inhibition, and reduce genetic instabilities. J Biol Chem *276*, 27178-27187.

Sakamoto, N., Ohshima, K., Montermini, L., Pandolfo, M., and Wells, R.D. (2001b). Sticky DNA, a self-associated complex formed at long GAA*TTC repeats in intron 1 of the frataxin gene, inhibits transcription. J Biol Chem *276*, 27171-27177.

Salinas-Rios, V., Belotserkovskii, B.P., and Hanawalt, P.C. (2011). DNA slip-outs cause RNA polymerase II arrest in vitro: potential implications for genetic instability. Nucleic Acids Res *39*, 7444-7454.

Sancar, A. (1996). DNA excision repair. Annu Rev Biochem *65*, 43-81.

Sandi, C., Pinto, R.M., Al-Mahdawi, S., Ezzatizadeh, V., Barnes, G., Jones, S., Rusche, J.R., Gottesfeld, J.M., and Pook, M.A. (2011). Prolonged treatment with pimelic o-aminobenzamide HDAC inhibitors ameliorates the disease phenotype of a Friedreich ataxia mouse model. Neurobiol Dis *42*, 496-505.

Sato, T., Oyake, M., Nakamura, K., Nakao, K., Fukusima, Y., Onodera, O., Igarashi, S., Takano, H., Kikugawa, K., Ishida, Y.*, et al.* (1999). Transgenic mice harboring a full-length human mutant DRPLA gene exhibit age-dependent intergenerational and somatic instabilities of CAG repeats comparable with those in DRPLA patients. Hum Mol Genet *8*, 99-106.

Savouret, C., Garcia-Cordier, C., Megret, J., te Riele, H., Junien, C., and Gourdon, G. (2004). MSH2-dependent germinal CTG repeat expansions are produced continuously in spermatogonia from DM1 transgenic mice. Mol Cell Biol *24*, 629-637.

Savva, R., McAuley-Hecht, K., Brown, T., and Pearl, L. (1995). The structural basis of specific base-excision repair by uracil-DNA glycosylase. Nature *373*, 487-493.

Schmidt, K.H., Abbott, C.M., and Leach, D.R. (2000). Two opposing effects of mismatch repair on CTG repeat instability in Escherichia coli. Mol Microbiol *35*, 463-471.

Schumacher, S., Pinet, I., and Bichara, M. (2001). Modulation of transcription reveals a new mechanism of triplet repeat instability in Escherichia coli. J Mol Biol *307*, 39-49.

Schweitzer, J.K., and Livingston, D.M. (1999). The effect of DNA replication mutations on CAG tract stability in yeast. Genetics *152*, 953-963.

Seeberg, E., Eide, L., and Bjoras, M. (1995). The base excision repair pathway. Trends Biochem Sci *20*, 391-397.

Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thastrom, A., Field, Y., Moore, I.K., Wang, J.P., and Widom, J. (2006). A genomic code for nucleosome positioning. Nature *442*, 772-778.

Sellier, C., Rau, F., Liu, Y., Tassone, F., Hukema, R.K., Gattoni, R., Schneider, A., Richard, S., Willemsen, R., Elliott, D.J.*, et al.* (2010). Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. Embo J *29*, 1248-1261.

Seriola, A., Spits, C., Simard, J.P., Hilven, P., Haentjens, P., Pearson, C.E., and Sermon, K. (2011). Huntington's and myotonic dystrophy hESCs: down-regulated trinucleotide repeat instability and mismatch repair machinery expression upon differentiation. Hum Mol Genet *20*, 176-185.

Seznec, H., Lia-Baldini, A.S., Duros, C., Fouquet, C., Lacroix, C., Hofmann-Radvanyi, H., Junien, C., and Gourdon, G. (2000). Transgenic mice carrying large human genomic sequences with expanded CTG repeat mimic closely the DM CTG repeat intergenerational and somatic instability. Hum Mol Genet *9*, 1185-1194.

Shelbourne, P.F., Keller-McGandy, C., Bi, W.L., Yoon, S.R., Dubeau, L., Veitch, N.J., Vonsattel, J.P., Wexler, N.S., Arnheim, N., and Augood, S.J. (2007). Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain. Hum Mol Genet *16*, 1133-1142.

Sherman, S.L., Jacobs, P.A., Morton, N.E., Froster-Iskenius, U., Howard-Peebles, P.N., Nielsen, K.B., Partington, M.W., Sutherland, G.R., Turner, G., and Watson, M. (1985). Further segregation analysis of the fragile X syndrome with special reference to transmitting males. Hum Genet *69*, 289-299.

Sherman, S.L., Morton, N.E., Jacobs, P.A., and Turner, G. (1984). The marker (X) syndrome: a cytogenetic and genetic analysis. Ann Hum Genet *48*, 21-37.

Shieh, S.Y., and Bonini, N.M. (2011). Genes and pathways affected by CAG-repeat RNA-based toxicity in Drosophila. Hum Mol Genet *20*, 4810-4821.

Shuck, S.C., Short, E.A., and Turchi, J.J. (2008). Eukaryotic nucleotide excision repair: from understanding mechanisms to influencing biology. Cell Res *18*, 64-72.

Sleeth, K.M., Robson, R.L., and Dianov, G.L. (2004). Exchangeability of mammalian DNA ligases between base excision repair pathways. Biochemistry *43*, 12924-12930.

Slupphaug, G., Kavli, B., and Krokan, H.E. (2003). The interacting pathways for prevention and repair of oxidative DNA damage. Mutat Res *531*, 231-251.

Smirnova, E., Toueille, M., Markkanen, E., and Hubscher, U. (2005). The human checkpoint sensor and alternative DNA clamp Rad9-Rad1-Hus1 modulates the activity of DNA ligase I, a component of the long-patch base excision repair machinery. Biochem J *389*, 13-17.

Snell, R.G., MacMillan, J.C., Cheadle, J.P., Fenton, I., Lazarou, L.P., Davies, P., MacDonald, M.E., Gusella, J.F., Harper, P.S., and Shaw, D.J. (1993). Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. Nature Genet *4*, 393-397.

Sobczak, K., de Mezer, M., Michlewski, G., Krol, J., and Krzyzosiak, W.J. (2003). RNA structure of trinucleotide repeats associated with human neurological diseases. Nucleic Acids Res *31*, 5469-5482.

Sobol, R.W., Horton, J.K., Kuhn, R., Gu, H., Singhal, R.K., Prasad, R., Rajewsky, K., and Wilson, S.H. (1996). Requirement of mammalian DNA polymerase-beta in base-excision repair. Nature *379*, 183-186.

Sofola, O.A., Jin, P., Qin, Y., Duan, R., Liu, H., de Haro, M., Nelson, D.L., and Botas, J. (2007). RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CGG premutation repeat-induced neurodegeneration in a Drosophila model of FXTAS. Neuron *55*, 565-571.

Solvsten, C., and Nielsen, A.L. (2011). FMR1 CGG repeat lengths mediate different regulation of reporter gene expression in comparative transient and locus specific integration assays. Gene *486*, 15-22.

Sopher, B.L., Ladd, P.D., Pineda, V.V., Libby, R.T., Sunkin, S.M., Hurley, J.B., Thienes, C.P., Gaasterland, T., Filippova, G.N., and La Spada, A.R. (2011). CTCF regulates ataxin-7 expression through promotion of a convergently transcribed, antisense noncoding RNA. Neuron *70*, 1071-1084.

Soragni, E., Herman, D., Dent, S.Y., Gottesfeld, J.M., Wells, R.D., and Napierala, M. (2008). Long intronic GAA*TTC repeats induce epigenetic changes and reporter gene silencing in a molecular model of Friedreich ataxia. Nucleic Acids Res *36*, 6056-6065.

Spiro, C., and McMurray, C.T. (2003). Nuclease-deficient FEN-1 blocks Rad51/BRCA1-mediated repair and causes trinudeotide repeat instability. Mol Cell Biol *23*, 6063-6074.

Spiro, C., Pelletier, R., Rolfsmeier, M.L., Dixon, M.J., Lahue, R.S., Gupta, G., Park, M.S., Chen, X., Mariappan, S.V., and McMurray, C.T. (1999). Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats. Mol Cell *4*, 1079-1085.

Stucki, M., Pascucci, B., Parlanti, E., Fortini, P., Wilson, S.H., Hubscher, U., and Dogliotti, E. (1998). Mammalian base excision repair by DNA polymerases delta and epsilon. Oncogene *17*, 835-843.

Subramanian, J., Vijayakumar, S., Tomkinson, A.E., and Arnheim, N. (2005). Genetic instability induced by overexpression of DNA ligase I in budding yeast. Genetics *171*, 427-441.

Sukhanova, M., Khodyreva, S., and Lavrik, O. (2010). Poly(ADP-ribose) polymerase 1 regulates activity of DNA polymerase beta in long patch base excision repair. Mutat Res *685*, 80-89.

Sukhanova, M.V., Khodyreva, S.N., Lebedeva, N.A., Prasad, R., Wilson, S.H., and Lavrik, O.I. (2005). Human base excision repair enzymes apurinic/apyrimidinic endonuclease1 (APE1), DNA polymerase beta and poly(ADP-ribose) polymerase 1: interplay between strand-displacement DNA synthesis and proofreading exonuclease activity. Nucleic Acids Res *33*, 1222-1229.

Sundararajan, R., Gellon, L., Zunder, R.M., and Freudenreich, C.H. (2010). Double-strand break repair pathways protect against CAG/CTG repeat expansions, contractions and repeat-mediated chromosomal fragility in Saccharomyces cerevisiae. Genetics *184*, 65-77.

Sunesen, M., Stevnsner, T., Brosh, R.M., Jr., Dianov, G.L., and Bohr, V.A. (2002). Global genome repair of 8-oxoG in hamster cells requires a functional CSB gene product. Oncogene *21*, 3571-3578.

Suominen, T., Bachinski, L.L., Auvinen, S., Hackman, P., Baggerly, K.A., Angelini, C., Peltonen, L., Krahe, R., and Udd, B. (2011). Population frequency of myotonic dystrophy: higher than expected frequency of myotonic dystrophy type 2 (DM2) mutation in Finland. Eur J Hum Genet *19*, 776-782.

Sutcliffe, J.S., Nelson, D.L., Zhang, F., Pieretti, M., Caskey, C.T., Saxe, D., and Warren, S.T. (1992). DNA methylation represses FMR-1 transcription in fragile X syndrome. Hum Mol Genet *1*, 397-400.

Swami, M., Hendricks, A.E., Gillis, T., Massood, T., Mysore, J., Myers, R.H., and Wheeler, V.C. (2009). Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. Hum Mol Genet *18*, 3039-3047.

Szczesny, B., Tann, A.W., and Mitra, S. (2010). Age- and tissue-specific changes in mitochondrial and nuclear DNA base excision repair activity in mice: Susceptibility of skeletal muscles to oxidative injury. Mech Ageing Dev *131*, 330-337.

Takano, H., Onodera, O., Takahashi, H., Igarashi, S., Yamada, M., Oyake, M., Ikeuchi, T., Koide, R., Tanaka, H., Iwabuchi, K.*, et al.* (1996). Somatic mosaicism of expanded CAG repeats in brains of patients with dentatorubral-pallidoluysian atrophy: cellular population-dependent dynamics of mitotic instability. Am J Hum Genet *58*, 1212-1222.

Takiyama, Y., Sakoe, K., Amaike, M., Soutome, M., Ogawa, T., Nakano, I., and Nishizawa, M. (1999). Single sperm analysis of the CAG repeats in the gene for dentatorubral-pallidoluysian atrophy (DRPLA): the instability of the CAG repeats in the DRPLA gene is prominent among the CAG repeat diseases. Hum Mol Genet *8*, 453-457.

Takiyama, Y., Sakoe, K., Namekawa, M., Soutome, M., Esumi, E., Ogawa, T., Ishikawa, K., Mizusawa, H., Nakano, I., and Nishizawa, M. (1998). A Japanese family with spinocerebellar ataxia type 6 which includes three individuals homozygous for an expanded CAG repeat in the SCA6/CACNL1A4 gene. J Neurol Sci *158*, 141-147.

Takiyama, Y., Sakoe, K., Soutome, M., Namekawa, M., Ogawa, T., Nakano, I., Igarashi, S., Oyake, M., Tanaka, H., Tsuji, S.*, et al.* (1997). Single sperm analysis of the CAG repeats in the gene for Machado-Joseph disease (MJD1): evidence for non-Mendelian transmission of the MJD1 gene and for the effect of the intragenic CGG/GGG polymorphism on the intergenerational instability. Hum Mol Genet *6*, 1063-1068.

Tanaka, F., Reeves, M.F., Ito, Y., Matsumoto, M., Li, M., Miwa, S., Inukai, A., Yamamoto, M., Doyu, M., Yoshida, M.*, et al.* (1999). 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 *65*, 966-973.

Tanaka, F., Sobue, G., Doyu, M., Ito, Y., Yamamoto, M., Shimada, N., Yamamoto, K., Riku, S., Hshizume, Y., and Mitsuma, T. (1996). Differential pattern in tissue-specific somatic mosaicism of expanded CAG trinucleotide repeats in dentatorubral-pallidoluysian atrophy, Machado-Joseph disease, and X-linked recessive spinal and bulbar muscular atrophy. J Neurol Sci *135*, 43-50.

Tebbs, R.S., Flannery, M.L., Meneses, J.J., Hartmann, A., Tucker, J.D., Thompson, L.H., Cleaver, J.E., and Pedersen, R.A. (1999). Requirement for the Xrcc1 DNA base excision repair gene during early mouse development. Dev Biol *208*, 513-529.

Telenius, H., Almqvist, E., Kremer, B., Spence, N., Squitieri, F., Nichol, K., Grandell, U., Starr, E., Benjamin, C., Castaldo, I.*, et al.* (1995). Somatic mosaicism in sperm is associated with intergenerational (CAG)n changes in Huntington disease (vol 4, pg 189, 1995). HumMolGenet *4*, 974.

Telenius, H., Kremer, B., Goldberg, Y.P., Theilmann, J., Andrew, S.E., Zeisler, J., Adam, S., Greenberg, C., Ives, E.J., Clarke, L.A.*, et al.* (1994). Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm. Nat Genet *6*, 409-414.

Thompson, L.H., and Schild, D. (2001). Homologous recombinational repair of DNA ensures mammalian chromosome stability. Mutat Res *477*, 131-153.

Thornton, C.A., Johnson, K., and Moxley, R.T., 3rd (1994). Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. Ann Neurol *35*, 104-107.

Thorslund, T., von Kobbe, C., Harrigan, J.A., Indig, F.E., Christiansen, M., Stevnsner, T., and Bohr, V.A. (2005). Cooperation of the Cockayne syndrome group B protein and poly(ADP-ribose) polymerase 1 in the response to oxidative stress. Mol Cell Biol *25*, 7625-7636.

Tian, L., Hou, C., Tian, K., Holcomb, N.C., Gu, L., and Li, G.M. (2009). Mismatch recognition protein MutSbeta does not hijack (CAG)n hairpin repair in vitro. J Biol Chem *284*, 20452-20456.

Todd, P.K., Oh, S.Y., Krans, A., Pandey, U.B., Di Prospero, N.A., Min, K.T., Taylor, J.P., and Paulson, H.L. (2010). Histone deacetylases suppress CGG repeat-induced neurodegeneration via transcriptional silencing in models of fragile X tremor ataxia syndrome. PLoS Genet *6*, e1001240.

Tome, S., Holt, I., Edelmann, W., Morris, G.E., Munnich, A., Pearson, C.E., and Gourdon, G. (2009). MSH2 ATPase domain mutation affects CTG*CAG repeat instability in transgenic mice. PLoS Genet *5*, e1000482.

Tome, S., Panigrahi, G.B., Lopez Castel, A., Foiry, L., Melton, D.W., Gourdon, G., and Pearson, C.E. (2011). Maternal germline-specific effect of DNA ligase I on CTG/CAG instability. Hum Mol Genet *20*, 2131-2143.

Trapp, C., Reite, K., Klungland, A., and Epe, B. (2007). Deficiency of the Cockayne syndrome B (CSB) gene aggravates the genomic instability caused by endogenous oxidative DNA base damage in mice. Oncogene *26*, 4044-4048.

Tuo, J., Chen, C., Zeng, X., Christiansen, M., and Bohr, V.A. (2002). Functional crosstalk between hOgg1 and the helicase domain of Cockayne syndrome group B protein. DNA Repair (Amst) *1*, 913-927.

Tuo, J., Muftuoglu, M., Chen, C., Jaruga, P., Selzer, R.R., Brosh, R.M., Jr., Rodriguez, H., Dizdaroglu, M., and Bohr, V.A. (2001). The Cockayne Syndrome group B gene product is involved in general genome base excision repair of 8-hydroxyguanine in DNA. J Biol Chem *276*, 45772-45779.

Umar, A., Boyer, J.C., and Kunkel, T.A. (1994). DNA loop repair by human cell extracts. Science *266*, 814-816.

Umar, A., Buermeyer, A.B., Simon, J.A., Thomas, D.C., Clark, A.B., Liskay, R.M., and Kunkel, T.A. (1996). Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. Cell *87*, 65-73.

Usdin, K., and Woodford, K.J. (1995). CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro. Nucleic Acids Res *23*, 4202-4209.

Vallur, A.C., and Maizels, N. (2010). Complementary roles for exonuclease 1 and Flap endonuclease 1 in maintenance of triplet repeats. J Biol Chem *285*, 28514-28519.

Valouev, A., Johnson, S.M., Boyd, S.D., Smith, C.L., Fire, A.Z., and Sidow, A. (2011). Determinants of nucleosome organization in primary human cells. Nature *474*, 516-520.

van den Broek, W.J., Nelen, M.R., van der Heijden, G.W., Wansink, D.G., and Wieringa, B. (2006). Fen1 does not control somatic hypermutability of the (CTG)(n)*(CAG)(n) repeat in a knock-in mouse model for DM1. FEBS Lett *580*, 5208-5214.

van den Broek, W.J., Nelen, M.R., Wansink, D.G., Coerwinkel, M.M., te Riele, H., Groenen, P.J., and Wieringa, B. (2002). Somatic expansion behaviour of the (CTG)n repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. Hum Mol Genet *11*, 191-198.

Venema, J., Mullenders, L.H., Natarajan, A.T., van Zeeland, A.A., and Mayne, L.V. (1990). The genetic defect in Cockayne syndrome is associated with a defect in repair of UV-induced DNA damage in transcriptionally active DNA. Proc Natl Acad Sci U S A *87*, 4707-4711.

Volker, J., Makube, N., Plum, G.E., Klump, H.H., and Breslauer, K.J. (2002). Conformational energetics of stable and metastable states formed by DNA triplet repeat oligonucleotides: implications for triplet expansion diseases. Proc Natl Acad Sci U S A *99*, 14700-14705.

Wagner, J.R., Hu, C.C., and Ames, B.N. (1992). Endogenous oxidative damage of deoxycytidine in DNA. Proc Natl Acad Sci U S A *89*, 3380-3384.

Wang, W., Lindsey-Boltz, L.A., Sancar, A., and Bambara, R.A. (2006). Mechanism of stimulation of human DNA ligase I by the Rad9-rad1-Hus1 checkpoint complex. J Biol Chem *281*, 20865-20872.

Wang, Y.H., Amirhaeri, S., Kang, S., Wells, R.D., and Griffith, J.D. (1994). Preferential nucleosome assembly at DNA triplet repeats from the myotonic dystrophy gene. Science *265*, 669-671.

Wang, Y.H., and Griffith, J. (1996). Methylation of expanded CCG triplet repeat DNA from fragile X syndrome patients enhances nucleosome exclusion. J Biol Chem *271*, 22937-22940.

Warrick, J.M., Chan, H.Y.E., Gray-Board, G.L., Chai, Y., Paulson, H.L, and Bonini, N.M. (1999). Suppression of polyglutamine-mediated neurodegeneration in Drosophila by the molecular chaperone Hsp70. Nature Genet *23*, 425-428.

Webster, A.D., Barnes, D.E., Arlett, C.F., Lehmann, A.R., and Lindahl, T. (1992). Growth retardation and immunodeficiency in a patient with mutations in the DNA ligase I gene. Lancet *339*, 1508-1509.

Wei, W., and Englander, E.W. (2008). DNA polymerase beta-catalyzed-PCNA independent long patch base excision repair synthesis: a mechanism for repair of oxidatively damaged DNA ends in post-mitotic brain. J Neurochem *107*, 734-744.

Weterings, E., and Chen, D.J. (2008). The endless tale of non-homologous end-joining. Cell Res *18*, 114-124.

Wheeler, T.M., and Thornton, C.A. (2007). Myotonic dystrophy: RNA-mediated musde disease. Curr Opin Neurol *20*, 572-576.

Wheeler, V.C., Lebel, L.A., Vrbanac, V., Teed, A., te Riele, H., and MacDonald, M.E. (2003). Mismatch repair gene Msh2 modifies the timing of early disease in Hdh(Q111) striatum. Hum Mol Genet *12*, 273-281.

Wheeler, V.C., Persichetti, F., McNeil, S.M., Mysore, J.S., Mysore, S.S., MacDonald, M.E., Myers, R.H., Gusella, J.F., and Wexler, N.S. (2007). Factors associated with HD CAG repeat instability in Huntington disease. J Med Genet *44*, 695-701.

Wiederhold, L., Leppard, J.B., Kedar, P., Karimi-Busheri, F., Rasouli-Nia, A., Weinfeld, M., Tomkinson, A.E., Izumi, T., Prasad, R., Wilson, S.H., *et al.* (2004). AP endonuclease-independent DNA base excision repair in human cells. Mol Cell *15*, 209-220.

Wilburn, B., Rudnicki, D.D., Zhao, J., Weitz, T.M., Cheng, Y., Gu, X., Greiner, E., Park, C.S., Wang, N., Sopher, B.L., *et al.* (2011). An antisense CAG repeat transcript at JPH3 locus mediates expanded polyglutamine protein toxicity in Huntington's disease-like 2 mice. Neuron *70*, 427-440.

Wilson, D.M., 3rd, and Barsky, D. (2001). The major human abasic endonuclease: formation, consequences and repair of abasic lesions in DNA. Mutat Res *485*, 283-307.

Wilson, D.M., 3rd, and McNeill, D.R. (2007). Base excision repair and the central nervous system. Neuroscience *145*, 1187-1200.

Wilson, S.H., Beard, W.A., Shock, D.D., Batra, V.K., Cavanaugh, N.A., Prasad, R., Hou, E.W., Liu, Y., Asagoshi, K., Horton, J.K., *et al.* (2010). Base excision repair and design of small molecule inhibitors of human DNA polymerase beta. Cell Mol Life Sci *67*, 3633-3647.

Wong, L.J., Ashizawa, T., Monckton, D.G., Caskey, C.T., and Richards, C.S. (1995). Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent. Am J Hum Genet *56*, 114-122.

Wu, X., Li, J., Li, X., Hsieh, C.L., Burgers, P.M., and Lieber, M.R. (1996). Processing of branched DNA intermediates by a complex of human FEN-1 and PCNA. Nucleic Acids Res *24*, 2036-2043.

Xanthoudakis, S., Smeyne, R.J., Wallace, J.D., and Curran, T. (1996). The redox/DNA repair protein, Ref-1, is essential for early embryonic development in mice. Proc Natl Acad Sci U S A *93*, 8919-8923.

Yang, J., and Freudenreich, C.H. (2007). Haploinsufficiency of yeast FEN1 causes instability of expanded CAG/CTG tracts in a length-dependent manner. Gene *393*, 110-115.

Yang, Z., Lau, R., Marcadier, J.L., Chitayat, D., and Pearson, C.E. (2003). Replication inhibitors modulate instability of an expanded trinucleotide repeat at the myotonic dystrophy type 1 disease locus in human cells. Am J Hum Genet *73*, 1092-1105.

Yoon, S.R., Dubeau, L., de Young, M., Wexler, N.S., and Arnheim, N. (2003). Huntington disease expansion mutations in humans can occur before meiosis is completed. Proc Natl Acad Sci U S A *100*, 8834-8838.

Yoshimura, M., Kohzaki, M., Nakamura, J., Asagoshi, K., Sonoda, E., Hou, E., Prasad, R., Wilson, S.H., Tano, K., Yasui, A., *et al.* (2006). Vertebrate POLQ and POLbeta cooperate in base excision repair of oxidative DNA damage. Mol Cell *24*, 115-125.

Zala, D., Colin, E., Rangone, H., Liot, G., Humbert, S., and Saudou, F. (2008). Phosphorylation of mutant huntingtin at S421 restores anterograde and retrograde transport in neurons. Hum Mol Genet *17*, 3837-3846.

Zatz, M., Passos-Bueno, M.R., Cerqueira, A., Marie, S.K., Vainzof, M., and Pavanello, R.C. (1995). Analysis of the CTG repeat in skeletal musde of young and adult myotonic dystrophy patients: when does the expansion occur? Hum Mol Genet *4*, 401-406.

Zhang, T., Huang, J., Gu, L., and Li, G.M. (2011). In vitro repair of DNA hairpins containing various numbers of CAG/CTG trinucleotide repeats. DNA Repair (Amst).

Zheng, M., Huang, X., Smith, G.K., Yang, X., and Gao, X. (1996). Genetically unstable CXG repeats are structurally dynamic and have a high propensity for folding. An NMR and UV spectroscopic study. J Mol Biol *264*, 323-336.

Zu, T., Gibbens, B., Doty, N.S., Gomes-Pereira, M., Huguet, A., Stone, M.D., Margolis, J., Peterson, M., Markowski, T.W., Ingram, M.A.*, et al.* (2010). Non-ATG-initiated translation directed by microsatellite expansions. Proc Natl Acad Sci U S A *108*, 260-265.

Zuhlke, C., Riess, O., Bockel, B., Lange, H., and Thies, U. (1993). Mitotic stability and meiotic variability of the (CAG)n repeat in the Huntington disease gene. Hum Mol Genet *2*, 2063-2067.

Zumwalt, M., Ludwig, A., Hagerman, P.J., and Dieckmann, T. (2007). Secondary structure and dynamics of the r(CGG) repeat in the mRNA of the fragile X mental retardation 1 (FMR1) gene. RNA Biol *4*, 93-100.

Agathi-Vasiliki GOULA

Implication des lésions oxydantes et du mécanisme de réparation par excision de base dans la sélectivité tissulaire de l'instabilité somatique des répétitions CAG dans la maladie de Huntington

Résumé

La maladie de Huntington (MH) est la plus courante parmi les maladies neurodégénératives à répétitions trinucléotidiques. La MH est une maladie progressive et fatale, causée par l'expansion anormale des répétitions CAG du gène de la Huntingtine. La longueur de l'expansion est instable et proportionnelle à la gravité de la maladie. L'instabilité affecte différemment les tissus, dont le cerveau, où le striatum ayant une forte instabilité dégénère, alors que le cervelet ayant une instabilité limitée est épargné par la maladie. Nous avons étudié le rôle des lésions oxydatives et du mécanisme de réparation par excision de base (BER) dans la sélectivité tissulaire de l'instabilité dans le striatum et le cervelet des souris R6/1. Nous avons observé un niveau de lésions oxydantes similaire dans les deux tissus, suggérant qu'il ne corrèle pas avec l'instabilité tissu-spécifique. Les lésions situées à des structures en épingle à cheveux étaient réfractaires à réparer. De plus, les niveaux et les activités des principales protéines BER étaient globalement diminués dans le striatum, en particulier FEN1 et LIG1 deux protéines de LP-BER. En outre, l'efficacité de réparation dépendait de la stœchiométrie de BER, la position de la lésion et la séquence d'ADN. En particulier, des cocktails protéiques aux stœchiométries BER du striatum réparaient moins efficacement par rapport à ceux du cervelet et les substrats contenant des CAG/CTG étaient préférentiellement réparés par LP-BER, la position de la lésion modulant son utilisation. Nos résultats suggèrent une faible coopération entre les activités BER, associée à la spécificité tissulaire de l'instabilité somatique CAG/CTG dans la MH.

Keywords : Maladie de Huntington, instabilité somatique, sélectivité tissulaire, lésions oxydantes, BER.

Summary

Huntington's disease (HD) is the most common among the neurodegenerative disorders associated to trinucleotide repeats expansion. HD is a progressive and fatal disease caused by the abnormal expansion of CAG repeats in the Huntingtin gene. The expansion length is unstable and proportional to the disease severity. The instability affects differently several tissues, among which the brain, where the striatum that shows a high instability degenerates, whereas the cerebellum that shows limited instability is spared from the disease. We addressed the role of oxidative lesions and Base Excision Repair (BER) in the tissue-selectivity of the somatic instability in striatum and cerebellum of R6/1 mouse model. Interestingly, we observed a similar level of oxidative lesions at both tissues, thus not correlating with the tissue-specific instability. We showed that lesions located at hairpin structures were refractory to repair. Additionally, levels and activities of main BER proteins were globally decreased in striatum relative to cerebellum, especially Fen1 and Lig1 two LP-BER proteins. Moreover we found that repair outcome is dependent upon BER stoichiometries, lesion location and sequence. In particular we showed that reconstituted striatal BER stoichiometries result in less efficient repair relative to the cerebellar ones and that CAG/CTG-containing substrates are preferentially repaired through LP-BER, the lesion position within the repeats modulating the requirement for LP-BER. Our results suggest a poor cooperation between BER activities that could underlie tissue-specificity of somatic CAG/CTG instability in HD.

Keywords : Huntington's disease, somatic instability, tissue-specificity, oxidative lesions, BER.