[Frontiers in Bioscience 13, 4467-4484, May 1, 2008]

Tandem repeats in human disorders: mechanisms and evolution

Pratibha Siwach[1], Subramaniam Ganesh[1]

[1]*Department of Biological Sciences and Bioengineering, Indian Institute of Technology, Kanpur, India*

TABLE OF CONTENTS

1. ABSTRACT

One of the most compelling reasons for the study of repetitive DNA sequence in the human genome has been the instability of simple repeat sequences associating with a growing and an interesting group of disorders affecting the neurological, neuromuscular or developmental processes. As a result, the molecular processes that underlie this unique form of mutation and the pathological pathways that lead to the disorders are being uncovered rapidly and are being intensively investigated. Genes with expanded repeats exhibit either loss-of-function or gain-of-function effect at the protein and/or RNA level. In this review, we aim to provide an overview of the recent advances in molecular pathology of disorders associated with heritable changes in the length of the repeat sequences, and examine how dynamism in these repeats is regulated.

2. INTRODUCTION

The three billion base pairs of our genome consist of a combination of four nucleotide bases in a random but functionally conserved arrangement. In spite of this randomness, many nucleotide sequences occur repeatedly to constitute a tandemly repeated signature motif. Recent analysis of the human genome sequence advocates that roughly 95% of genome is made up of tandem repeat sequences (TRS) (1-2). The TRS consists of consecutive and head-to-tail copies of sequence and posses varied levels of compositional complexity ranging from a few bases to more then hundred or so. For example, surveys based on database analysis showed that A/T repeats are maximal among mononucleotides, CA/GT as dinucleotide repeats, and that the GC rich repeats are overrepresented in tri-, and tetra-nucleotides (2-3). TRS are spread more or less evenly

throughout the entire human genome, as they are found to be located in the coding, non-coding and regulatory regions of the genes (3-4).

Numerous evidences demonstrate that TRS are significantly important in genomic organization because of their effect on various cellular activities like chromatin organization, regulation of gene expression, replication, recombination, mismatch repair systems, and monitoring of cell cycle (5). TRS are dynamic in nature as they are quite prone to undergo repeat length variations (6). Although the mechanisms that generate and maintain TRS in genomes are poorly understood, available evidences suggest that, due to the reiterated nucleotide sequences, processes coupled with replication and recombination could lead to their expansion and/or contraction (7-8). Perhaps the most compelling reason for studying TRS is their involvement in plethora of human diseases. Though length variation (expansion or contraction) in TRS largely remain neutral because of their location in nongenic regions, there are sufficient examples of their involvement in disturbing gene functions and regulations (4). Particularly, TRS lying with in or close to gene sequences are more prone to perturb normal gene function by altering their activity. Indeed, TRS are known to be associated both with heritable (germ-line mutations) and non-heritable (somatic mutations) disorders in humans (4). More often TRS manifest "dynamism" in heritable disorder wherein the probability of repeat expansion is directly proportional to number of repeating units within the TRS (7, 9-11). In these disorders, therefore, symptoms appear at an early age when the expanded TRS is passed on to the next generation (7, 10-11). In this review, we focus on progress made in understanding the significance of TRS in human disorders. We discuss the molecular mechanisms of genome instability associated with TRS with the hope of providing insights into the mechanisms and the evolution of this novel group of disorders.

3. TRS EXPANSION AND THEIR BIOLOGY IN RELATION TO DISORDER

3.1. Functional significance of TRS in non-coding regions of genes

There is growing number of evidences that suggest critical roles for TRS in gene control and regulation. For example, removal of CA repeats normally present in the 3'-untranslated region (UTR) of BCL2 gene led to stabilization of BCL2 mRNA abundance (4), and removal of CAG repeat stretch in the 5'-UTR of the calmodulin-1 gene reduced its expression by 45% (12). Repeat tracts present in the 3'-UTR perhaps influence various aspect of mRNA metabolism by serving as binding sites for RNA binding proteins (13). Curiously, a majority of transcripts having GATA repeats in the 3'-UTR were found to encode membrane proteins, indicating that such repeats could serve as signals for translational regulation of genes involved in membrane functions (14). Intronic TRS are also known to regulate gene expression. For example, a tetranucleotide repeat located in the first intron of the gene encoding tyrosine hydroxylase in human is shown to function as a transcriptional enhancer, suggesting that TRS in introns might have a direct influence on gene expression (15-16). Similar to TRS present in the intergenic regions, the TRS lying in the genic region also show a higher mutation rate as compared to regions that do not harbor repetitive sequences (17). Considering the essential role that such repeat elements play in gene function, elongation/shortening events in TRS are expected to affect the normal gene function which eventually could lead to a disorder (Figure 1). At least 10 genetic disorders are known in humans in which the primary cause of the defect is the expansion of TRS located in the non-coding region of corresponding genes (Table 1).

3.1.1. Disease associated TRS in the 5'-flanking regions of genes

Three neurological disorders are associated with expansion of TRS located in the 5'-flanking region that lead to the silencing of corresponding gene functions (Table 1). These include progressive myoclonic epilepsy of Unverricht and Lundborg (EPM1) associated with mutations in CSTB gene, the fragile X mental retardation syndrome associated with mutations in FMR1 gene (a common form known as FMR), or the FMR2 gene (a milder subsyndrome known as FRAXE) (Table 1). The CSTB gene encodes cystatin B protein, a cysteine protease inhibitor, which is thought to play a role in protecting against the proteases leaking from lysosomes (18). Expansion of a dodecamer repeat in the 5'-flanking region of CSTB gene is the common mutation mechanism in EPM1 (19). This expansion is associated with a marked reduction of CSTB mRNA and subsequent loss of cystatin B activity, leading to the onset of EPM1 (20). Among the most probable explanation for the reduced transcription of CSTB gene include hypermethylation of the CpG elements, altered spacing of promoter elements, altered chromatin structure, and recruitment of transcriptional repressors on the dodecameric repeat (21). Similar observations were made for the genes FMR1 and FMR2, except that the trinucleotide repeat sequences in these genes fall within the transcribed region (5'-UTR) (Table 1). Expression of both FMR1 and FMR2 genes was shown to be downregulated by repeat expansion and methylation (22-24). FMR1 gene encodes an RNA binding protein that associates with polyribosome and regulates the process of translation (25-26). Thus, translational dysregulation of mRNAs normally associated with FMR1 protein was thought be the primary cause of FMR syndrome (27). FMR2 encodes a DNA binding nuclear protein, and could perhaps function as a transcription factor (28). Loss-of-function of corresponding genes in disorders EPM1 and FMR is the likely to be the primary cause for the resulting phenotype as mutations other than repeat expansion are also reported for CSTB and FMR1 genes (29-31). TRS in FMR1 and FMR2 genes represent rare folate-sensitive fragile sites (known as FRAXA and FRAXE, respectively) that are especially prone to forming gaps on metaphase chromosomes when cells are cultured *in vitro* under conditions that block DNA replication (32). While chromosomal breaks in FRAXA or FRAXE loci have not yet been reported in FMR patients, another fragile site, FRA11B located at 11q23.3 and harboring CCG-repeats, was shown to associate with

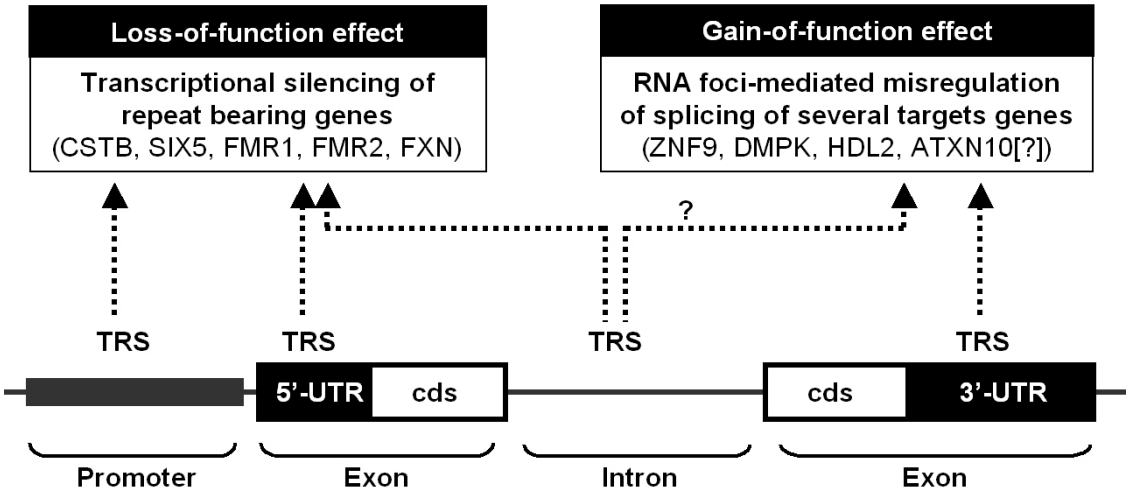

Figure 1. Mechanism of pathogenesis associated with expansion of non-coding TRS. *Loss-of-function effect:* In EPM1, SIX5, FMR and FRAXE, then unstable TRS element is localized either the promoter or the 5'-UTR (Table 1), and expansion of this TRS leads to reduced transcription of the respective gene. The molecular mechanism of gene-silencing could be hypermethylation of the CpG elements, altered spacing of promoter elements, altered chromatin structure, and/or recruitment of transcriptional repressors. Thus, loss-of-function of corresponding genes is the likely to be the primary cause for the resulting phenotype. *Gain-of-function effect*: In DM1, DM2, and HDL2, the unstable TRS fall in the transcribed part of the gene (Table 1). Transcripts bearing the expanded TRS form distinct RNA foci in the nucleus, and affect the RNA splicing processes of several downstream genes. Thus, the TRS in these disorders could have a 'trans-dominant' effect on a select group of genes and this may explain their dominant mode of inheritance (Table 1). The intronic TRS of ATX10 gene is believed to have the same effect although yet to be supported by experimental evidences and therefore identified with a question mark in the figure. *TRS, location of tandem repeat sequence, cds, coding sequence.*

hromosomal deletions in Jacobsen syndrome, a rare congenital disorder (33). Interestingly, a proportion of such patients inherited expanded CGG repeats at FRA11B site that falls within the 5'UTR of the proto-oncogene CBL2 (34). This observation raises an interesting possibility that the CGG repeat present in the fragile site could mediate chromosomal breakage in Jacobsen syndrome (33).

Similar to the dodecamer repeat in CSTB gene, expansion of CAG repeats in the promoter region of PPP2R2B is shown to be associated with one of the several form hereditary ataxia, named SCA12 (35). PPP2R2B gene encodes a brain-specific regulatory subunit of the protein phosphatase PP2A. While no study has elucidated the effect of CAG repeat on PPP2R2B gene activity, it is believed that the expansion of the repeat might alter the expression level of this gene, which could in turn affect activity of the enzyme PP2A (36). Unlike EPM1 and FMR, the phenotypes linked with SCA12 are autosomal dominant in nature, therefore an increased expression of PPP2R2B gene is likely to be the etiology behind SCA12 (36). Expansion of CTG repeat in the promoter region of SIX5, leading to its reduced expression, is known to be associated with a subset of symptoms of muscular dystrophy type 1(Table 1) (37, see section 3.1.3 for details)

3.1.2. Disease associated TRS in the intronic regions of genes

Instability in TRS located in the intronic regions of genes is linked with three distinct forms of hereditary ataxias and two form of myotonic dystrophies (Table 1). The clinical phenotypes of ataxias are characterized by predominantly cerebellar symptoms such as incoordination and unsteady gait (38). A majority of cases with one of the major forms of ataxias, known as Friedreich ataxia, are due to expansion of a GAA trinucleotide repeat in intron 1 of the FXN gene (39-40). This nuclear gene encoded mitochondrial protein, named frataxin, was shown to regulate mitochondrial iron transport and respiration (41). The expanded alleles of FXN gene produce lesser levels of mature FXN mRNA than the alleles in the normal range (42). The reduction in the FXN messenger level could be due to the repeat mediated stalling of RNA polymerase (43) or due to the repeat mediated epigenetic changes in the chromatin flanking the FXN gene (44). In either case, expansion of GAA repeats would result in frataxin deficiency, leading to impaired iron metabolism, oxidative damage, and progressive iron accumulation (41). The second form of ataxia (SCA10) is associated with an expansion of a pentanucleotide (ATTCT) repeat in intron 9 of the ATXN10 gene (45). Unlike FXN, however, the expansion of ATTCT repeat did not alter the normal level of ATX10 transcripts in the affected (46). However, in mice, homozygous null mutants for ATX10 gene resulted in embryonic lethality and heterozygotes were asymptomatic (46). SCA10 is being a dominant disorder in humans, a simple loss of function or haploinsufficiency of ATXN10 is unlikely to be the cause for the disease phenotype (46). Thus the precise mechanism by which the intronic repeats exert a dominant role in the etiology of SCA10 is yet to be resolved.

Table 1. Molecular and genetic features of disease-associated unstable repeats falling in the non-coding regions of genes

Diseases	Gene implicated	Chromosomal location of the gene	Affected protein	Location of the repeat in the gene	Repeat unit	Number of repeats: wild-type (disease-associated)	Mode of inheritance[1]	Parental gender-bias[1]	Effect of mutation[1]	OMIM#
Progressive myoclonus epilepsy	CSTB	21q22.3	Cystatin B	Promoter	CCCCGCC CCGCG	<12 (>17)	AR	NR	LOF	601145
Spino-cerebellar ataxia type -12	PPP2R2B	5q31-5q32	Phosphatase 2 regulatory subunit B	Promoter	CAG	<28 (>66)	AD	NR	LOF	604325
Ocular cataract phenotype associated with DM1	SIX5	19q13.32	DM locus-associated homeodomain protein	Promoter	CTG	<35 (>50)	AD	M	LOF	600963
Fragile X syndrome	FMR1	Xq27.3	FMRP	5'-UTR	CGG	<50 (>200)	XD	M	LOF	309550
Fragile XE mental retardation	FMR2	Xq28	FMR2 protein	5'-UTR	CCG	<35 (>300)	XR	M	LOF	309548
Jacobson's syndrome	CBL2	11q23.3	Adapter for RPTK	5'-UTR	CCG	<11 (>100)	NR	P	Chromo-somal deletion	165360
Spinocerebellar ataxia type 10	ATXN10	22q13.3	Ataxin-10	Intron 9	ATTCT	<1000 (>1000)	AD	P	NR	603516
Myotonic dystrophy 2	ZNF9	3q21.3	RNA-binding protein	Intron 1	CCTG	<75 (>1110)	AD	NR	GOF	116955
Friedrich's ataxia	FXN	9q13–21.1	Frataxin	Intron 1	GAA	<80 (>100)	AR	M	LOF	606829
Myotonic dystrophy 1	DMPK	19q13	MD protein kinase	3'-UTR	CTG	<35 (>50)	AD	M	GOF	605377
Spino-cerebellar ataxia type-8	ATXN8OS	13q21	Non-coding RNA	Transcribed region	CTG	<34 (>127)	AD	M	GOF	603680
Huntington disease like-2	JPH3	16q23-24	Junctophilin-3	3'-UTR	CTG	<28 (>50)	AD	NR	NR	605268

[1]Abbreviations used: AR, autosomal dominant; AR, autosomal resessive; XD, X-linked dominant; XR, X-linked recessive; NR, not reported; M, maternal; P, paternal; LOF, loss-of-function; GOF, gain-of-function

Myotonic dystrophy (DM) is a multisystem disorder and also the common form of muscular dystrophy affecting the adults. One form of DM is caused by CCTG repeat expansion in intron 1 of the ZNF9 gene and the resulting phenotype is referred to as DM2 (47). The ZNF9 gene encodes an RNA-binding protein with zinc finger domains and was shown to play critical roles in translational regulation (48-49). Similar to the ATTCT repeat of ATXN10 gene, the expansion of CCTG repeat does not appear have a direct consequence on the levels or processing of ZNF9 mRNA or the encoded protein (50). Studies have shown that the nuclear RNA bearing the expanded repeats could have a 'trans-dominant' effect on a select group of genes (see section 3.1.3) and this mechanism may explain the dominant mode of inheritance for DM2.

3.1.3. Disease associated TRS in the 3'-untranslated regions of genes

A second and common form of DM, referred to as DM1, results due to an expansion of a CTG repeat located in the 3'-UTR of DMPK gene encoding a serine-threonine kinase involved in signal pathways of myogenesis (51-52) (Table 1). Three mechanisms have been suggested to explain the dominant mode of DM1 inheritance. These include: (i) Haploinsufficiency of the DMPK protein because its mRNA and protein levels were reduced in muscle biopsies and cell lines derived from patients (53). (ii) Altered expression of neighboring genes, possibly through repeat-induced changes in chromatin structure (54); For example, the CTG repeat of DMPK gene also falls on the promoter region of SIX5 gene which is located immediately downstream of DMPK (55). It was indeed shown that the expansion of CTG in DMPK results in reduction in cellular levels of SIX5 transcripts (37). Curiously, SIX5 knockout mice developed ocular cataracts, similar to DM1 patients, raising a possibility that DM1 is a contiguous gene syndrome involving, among others, deficiency of DMPK and SIX5 proteins (56). (iii) Pathogenic effects of the CUG expansion in mRNA that accumulates as ribonuclear foci and disrupts mRNA splicing of cTNT, IR, ClC-1, Tau and MTMR1 genes (57-62). Curiously, expansion of intronic CCTG in the ZNF6 gene (involved in DM2) also leads to the formation of RNA foci and abnormal splicing of above mentioned genes (63, 61). Thus, the similarity of mechanism of mutation between DM2 and DM1 is striking, and suggest that a dominant effect of expanded TRS bearing transcripts on the RNA processing of downstream genes could underlie the genesis of some subphenotypes that are common among the two forms of DM (64).

Second example in this category is the hereditary disorder known as Huntington's disease-like 2 (HDL2). Similar to HD, HDL2 is an adult onset autosomal dominant and progressive neurological disorder clinically characterized by abnormal movements, dementia, and psychiatric syndromes. HDL2 is caused by a CTG expansion mutation located within a variably spliced exon of Junctophilin-3 gene (JPH3; 65). Recent studies have identified the presence of a JPH3 splice variant wherein the CTG repeat localize to the 3'-UTR of that transcripts. Transcripts bearing expanded repeats form RNA foci, very similar the foci found in DM1 and DM2, and were detected in neurons of HDL2 patients (66). This study suggests that, similar to DM pathology, RNA toxicity could underlie the pathology in HDL2.

3.1.4 Disease associated TRS in a non-coding RNA

In one form of dominantly inherited ataxias, referred to as SCA8, an expansion of CTG repeats occurs in a region that is transcribed into what is believed to be a processed non-coding RNA (67) (Table 1). Curiously, sequence that transcribes the 5' end of this untranslated

Table 2. Molecular and genetic features of disease-associated unstable repeats falling in the coding regions of genes

Diseases	Gene implicated	Chromosomal location of the gene	Affected protein	Repeat unit: codon (amino acid)	Number of repeats: wild-type (disease-associated)	Mode of inheritance[1]	Parental gender –bias[1]	Effect of mutation[1]	OMIM#
Spinal-bulbar muscular atrophy	AR	Xq11–12	Androgen receptor	CAG (Q)	<30 (>40)	XR	M	GOF	313700
Huntington Disease	IT15	4p16.3	Huntingtin	CAG (Q)	<35 (>40)	AD	P	GOF	143100
Dentatorubral pallidoluysian atrophy	ATN1	12p13.31	Atrophin-1	CAG (Q)	<23 (>49)	AD	P	GOF	607462
Spinocerebellar ataxia type 1	ATXN1	6p23	Ataxin-1	CAG (Q)	<39 (>41)	AD	P	GOF	601556
Spinocerebellar ataxia type 2	ATXN2	12q24.1	Ataxin-2	CAG (Q)	<23 (>36)	AD	P	GOF	601517
Spinocerebellar ataxia type 3	ATXN3	14q32.1	Ataxin-3	CAG (Q)	<36 (>68)	AD	P	GOF	607047
Spinocerebellar ataxia type 7	ATXN7	3p12–13	Ataxin-7	CAG (Q)	<17 (>38)	AD	P	GOF	607640
Spinocerebellar ataxia type 6	CACNAlA	19p13.2-p13.1	Voltage-dependent calcium channels	CAG (Q)	<16 (>21)	AD	P	GOF	601011
Spinocerebellar ataxia type 17	TBP	6q27	TATA box binding Protein	CAG (Q)	<42 (>47)	AD	P	NR	600075
Infantile onset spinocerebellar ataxia-8	ATXN8	13q21	Undetermined	CAG (Q)	<34 (>74)	AR	NR	NR	603680
Blepharophimosis-ptosis-epicanthus inversus syndrome	FOXL2	3q23	Forkhead box L2 transcription factor	GCN (A)	<14 (>22)	AD	P	LOF	605597
Holoprosencephaly	ZIC2	13q32	Zinc finger protein of the cerebellum 2	GCG (A)	<15 (>25)	De novo	NR	LOF	603073
Central hypoventilation syndrome	PHOX2B	4p12	Paired-like homeobox 2b transcription factor	GCN (A)	<20 (>25)	De novo	NR	LOF	603851
Mental retardation, Epilepsy, Partington syndrome	ARX	Xp22.1-p21.3	Aristaless-related homeobox	GCN (A)	<12 (>20)	XR	NR	Partial LOF	300382
Mental retardation with growth hormone deficiency	SOX3	Xq27.1	SRY-related HMG-box transcription factor family	GCC (A)	<15 (>26)	XR	NR	LOF	313430
Cleidocranial dysplasia	RUNX2	6p21	Runt-related transcription factor 2	GCG (A)	<17 (>27)	AD	NR	LOF	600211
Synpolydactyly	HOXD13	2q31.1	Homeobox D13 transcription factor	GCN (A)	<15 (>22)	AD	NR	LOF	142989
Oculopharyngeal muscular dystrophy	PABPN1	14q11.2-q13	Poly Adenylate-binding protein	GCG (A)	<6 (>9)	AD	NR	GOF	602279
Hand–foot–genital syndrome	HOXA13	7p15-p14	Homeobox A13 transcription factor	GCC (A)	<12 (>18)	AD	NR	LOF	142959
Pseudoachondro-plasia and multiple epiphyseal dysplasia	COMP	19p13.1	Cartilage oligomeric matrix protein	GCN (D)	5 (<5>)	AD	NR	LOF	600310
Huntington disease like-2	JPH3	16q23-24	Junctophilin-3	CTG (A or L)	<28 (>50)	AD	NR	NR	605268

[1]Abbreviations used: AR, autosomal dominant; AR, autosomal resessive; XR, X-linked recessive; NR, not reported; De novo, De novo mutation in a germ cell; M, maternal; P, paternal; LOF, loss-of-function; GOF, gain-of-function #Online Mendelian Inheritance in Man (OMIM) ID (More details: http://www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM)

transcript (known as ATXN8OS) shares the first exon of another gene (KLHL1) that is transcribed in the opposite orientation (68). The CTG/CAG repeats, however, do not fall in the coding region of KLHL1 gene (68). Thus, ATXN8OS transcript appear to be an endogenous RNA that regulates the expression of sense transcript encoding a Kelch-like protein, and the expanded repeat might affect its regulatory function (68). Targeted disruption of KLHL1 gene in mouse resulted in ataxic phenotype, both in homozygous and heterozygous conditions, suggesting that loss of KLHL1 activity might play significant part in the SCA8 etiology (69). A recent study also suggests that the SCA8 locus might encode a protein with polyglutamine from a previously unidentified antiparallel transcript of ATXN8OS spanning the CAG repeats and transgenic expression of it results in neurological phenotype in mice (70) (Table 2). Thus, SCA8 pathogenesis might involve dysfunction at both the protein and RNA levels.

3.2. Functional significance of TRS in the coding regions of genes

While repeats falling in the transcribed and translated parts of the genome although makeup very insignificant proportion of total repeat content of human genome, they are of utmost functional importance. Thus any dynamism in these repeats is very likely to affect the protein that they encode. Analyses of the TRS in the coding regions of genes reveal stronger selection pressure on modulating the length of the repeat. For example, TRS of three or multiples of three bases are found to be overrepresented in the coding regions as compared to mono- or di-nucleotide motifs (71). Taken together, these observations on TRS motifs indicate that the selection against frame shifts strongly influences distribution of repeat lengths (71). Amino acid repeats coded by the TRS in the translated sequences are three times more abundant in the proteome of eukaryotes than in prokaryotes (72-74). Moreover, a majority of these repeats are clustered around non-homologous proteins suggesting that the proteins containing repeats are predominantly unique to eukaryotes (72). For example amino acid repeats are predominantly found in muscle, brain and synaptic cell adhesion proteins, but are underrepresented in very basic cellular functions. Functional classification of proteins with TRS encoded amino acid repeats suggest that these repeats might facilitate protein–protein interactions and therefore required for the formation of multi-protein complexes (75-76). These observations strongly indicate that coding repeats are evolutionally favored and have an important role in protein functions (72-76).

3.2.1. TRS in the coding region: polyglutamine disorders

Polyglutamine (PolyQ) disorders are perhaps the most infamous and the most vigorously investigated

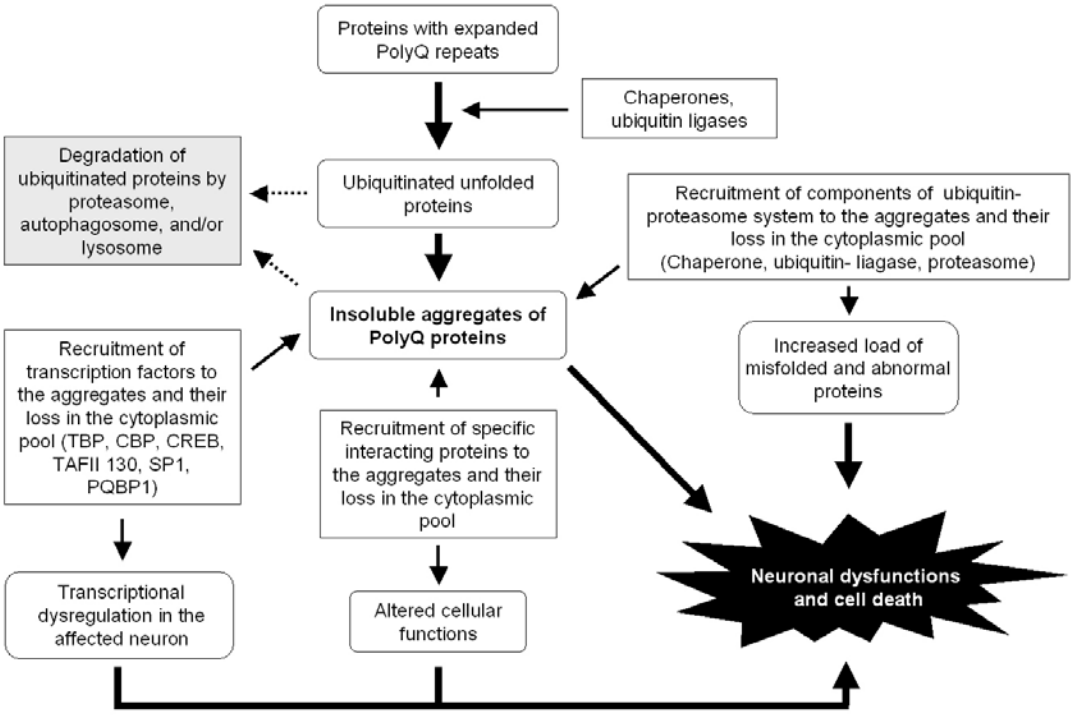

Figure 2. Mechanism of pathogenesis associated with PolyQ disorders. Protein with expanded PolyQ tract adopts abnormal conformation and accumulates as intracellular insoluble aggregates in all PolyQ disorders. These aggregates are expected to affect the cellular functions in a variety of ways. These include, recruiting molecular chaperones and components of the ubiquitin-proteasome system (UPS). This effort does remove a portion of the PolyQ proteins from the cytoplasmic pool (arrows shown with dotted lines). However, due to an imbalance between the UPS capacity and the synthesis of aggregation-prone, the aggregates continue to form, and might as well increase the cellular load f abnormal/misfolded proteins. The PolyQ aggregates are also known to affect the transcription and RNA metabolic processes by recruiting several transcriptional activators and repressors (TBP, CBP, CREB, TAFII, SP1, PQBP1) into the aggregates. Other cellular processes can also be affected by the specific interaction of proteins with the aggregates, which ultimately leading to neuronal dysfunction and death.

disorder associated with TRS instability in the coding sequence. So far 10 genetically distinct PolyQ disorders are known in humans and all of them are caused by expansion of CAG repeat in the translated region of the gene involved (Table 2). These include Huntington disease (HD), Spinal-bulbar muscular atrophy (SBMA), Dentatorubral pallidoluysian atrophy (DRPLA), and 7 types of ataxias (Table 2). Although the PolyQ proteins are involved in diverse cellular functions and expressed ubiquitously, all PolyQ disorders are generally confined to the nervous system (77). Besides symptoms that are unique to a given PolyQ disorder, all 10 of them associate with motor in-coordination, and cognitive impairment (77). While the specific symptoms could be attributed to the subset of neurons vulnerable in each disorder, the common and progressive clinical features in these disorders could result from a toxic effect of the expanded PolyQ tracts (77-80). For example, loss of huntingtin does not result in HD whereas transgenic over-expression of the expanded huntingtin results HD-like symptoms in mice (81-83). Similar observations were also made for SCA1 and SBMA mice models (81, 83, 84-85,). Moreover, ectopic expression of expanded PolyQ tracts tagged to hypoxanthine phosphoribosyltransferase protein or the green fluorescent protein cause cellular toxicity and progressive neurological phenotype in mice (86-87) Taken together, these results propose a unifying mechanism of pathogenesis for PolyQ disorders involving a dominant gain-of-function by the expanded PolyQ tract itself (77).

3.2.1.1. The role of PolyQ aggregates in neuropathology

Protein with the expanded PolyQ tract adopts abnormal conformation and accumulates as intracellular insoluble aggregates in all the PolyQ disorders (77, 88-90). These PolyQ aggregates are also known to recruit molecular chaperones and components of the ubiquitin-proteasome system (UPS) thereby raising the possibility that an overload on the UPS, due to the aggregation of misfolded proteins, is likely to promote the neurodegenerative process (77, 91-92) (Figure 2). Thus, an imbalance between the UPS capacity and the synthesis of aggregation-prone, misfolded PolyQ proteins might initiate the onset of diseases symptoms (93-95). This notion was strengthened by the findings that over-expression of chaperones and/or ubiquitin-protein ligases ameliorated the cellular toxicity of PolyQ proteins in cellular and animal models (96-101). Similarly, inhibition of proteasome increased the PolyQ-mediated cellular toxicity (102-103). Moreover, the formation of aggregates in a length and dosage-dependent manner is in agreement with this

hypothesis because longer polyglutamine repeats provoke earlier onset of the disease symptoms, and their faster progression (104). However, the functional relationship between PolyQ aggregates and their toxic effects on neurons has been challenged by several studies, and instead argued that the PolyQ aggregates are neuronal protective in nature (105-106). For example, suppression of inclusions formation in a cell-based system resulted in an increased cell-death in a cellular model (107), and neurons of a transgenic mouse that lacked E6-AP ubiquitin ligase but over-expressed expanded PolyQ protein exhibited significantly fewer aggregates but markedly increased pathology (108). Likewise, transgenic mice for SCA1 expressing an ataxin-1 variant with expanded PolyQ tract developed ataxia and neuropathology but without apparent aggregation (109). Conversely, a transgenic mouse line for huntingtin protein with expanded PolyQ failed to develop symptoms of HD despite the presence of widespread neuronal aggregates (110). While these studies did not reject the notion that PolyQ aggregates are toxic to cell, they do advocate the possibility that PolyQ proteins need not always form microscopically noticeable aggregate, and that non-visble aggregate forms (soluble or smaller aggregates) of PolyQ proteins are likely to be more toxic to the neurons (105, 111). In either case, a change in the conformation of expanded PolyQ proteins is likely to recruit other important cellular proteins through abnormal protein-protein interactions which might eventually lead to neuronal dysfunction and cell death (111-116) (Figure 2).

Clearance of toxic proteins, such as the expanded PolyQ species, is especially critical for brain tissue because the accumulating abnormal proteins cannot be diluted by the cell division as neurons are post-mitotic in nature. The ubiquitin-proteasome and autophagy-lysosome pathways are the two major routes for the clearance of unwanted and misfolded toxic proteins in eukaryotic cells (117). While clearance of short lived proteins are delegated to the ubiquitin-proteasome pathway, degradation of organelles and long-lived unwanted proteins are mediated by the autophagy-lysosome pathway. Intriguingly, the PolyQ proteins are known to be cleared by both the cellular routes because blocking either the proteasome or the autophagosome/lysosome lead to increased toxicity in cells expressing expanded PolyQ proteins (118-121). Conversely, activation of autophagy enhanced the clearance of expanded PolyQ proteins, and reduced their toxicity in cell and animal models, thereby advocating autophagy as one of the potential therapeutic targets for PolyQ disorders (122-125) (Figure 2). Recent studies suggest that autophagy could compensate for the ubiquitin-proteasome to dysfunctions and that the defects in the autophagic process might predispose to neurodegeneration (121, 126).

3.2.1.2. The role of PolyQ expansion in transcriptional dysregulation

It is of interest to note that proteins bearing PolyQ tract and associated with disorders are known to be involved in transcriptional regulation (127-128). For example, androgen receptor and TATA-binding protein (involved in SBMA and SCA17 respectively; (Table 2) are transcription factors (129-130). Atrophin-1 (DRPLA), ataxin-1 (SCA1), ataxin-3 (SCA3), and ataxin-7 (SCA7) are known to function as transcriptional regulators (131-134). In addition, huntingtin (HD) is believed function as a co-repressor for transcription factors (135). It is therefore obvious to expect that PolyQ expansion in these proteins would lead to quantitative changes in the neuronal transcriptome. Gene expression studies indeed have shown that dysregulation of neuronal genes could be an early pathogenic event leading to disorders (87, 136-139). Dysregulation involves numerous genes representing diverse cellular pathways, and therefore the mechanism by which PolyQ might cause transcriptional alteration is yet to be fully resolved. Nevertheless, transcriptional regulators such as CREB-binding protein, TBP-associated factor, SP1 transcription factor, and p53 were shown to interact with expanded PolyQ-containing proteins and to localize with their aggregates (140-142). Moreover, increased acetylation of histone H3 at the promoters of a group of genes in the retinal tissue and their transcriptional down-regulation have also been demonstrated in a SCA7 mouse model (143) (Figure 2). Intriguingly, PolyQ pathology has also been linked to micro-RNAs (miRNAs) that are known to modulate the half-life and/or the translational efficiency of mRNAs (144). For example, reduction in the cellular levels of a miRNA, known as ban, correlated with enhanced toxicity to ataxin-3, and its over expression rescued the PolyQ-mediated neurotoxocity, thereby revealing novel pathways and targets for therapeutics for the treatment of PolyQ diseases (144).

3.2.2. TRS in the coding region: polyalanine disorders

Beside the PolyQ expansion associating with several neurodegenerative disorders, recent reports demonstrate the expansion of polyalanine (PolyA) stretches as a disease mechanism in predominantly developmental and congenital malformation syndromes (134) (Table 2). Intriguingly, eight of the nine genes associated with the PolyA disorders encode transcription factors (Table 2). This is not unexpected because stretches of alanine residues are known to be conserved in transcription factors, and thought to play a role in transcriptional repression (76, 146-147). PolyA stretches, however, are relatively shorter repeats and are coded by imperfect trinucleotide repeats (GCN) as against homogenous CAG repeats encoding longer PolyQ sequence (137) (Table 2). In a majority of these disorders, mutations other than PolyA expansion also result in developmental abnormalities, thus suggesting that a different disease mechanism is involved in PolyA disorders as compared to PolyQ disorders (14). Available information suggests that PolyA expansion could lead to complete, partial loss-of-function or a dominant negative gain-of-function of the protein (Table 2). For example, in synpolydactyly limb malformation syndrome, the length of PolyA tract in the HOXD13 protein clearly correlates with the severity of anomalies in both heterozygotes and homozygotes, with homozygosity resulting in more severe phenotype (151, 152). In both human and mice, null alleles of HOXD13 gene manifest less severe phenotype of synpolydactyly than HOXD13 protein with PolyA expansion (152-154). Intriguingly, mice deficient for HoxD complex (simultaneous deletion of Hoxd13, Hoxd12 and

Hoxd11genes*)* develop anomalies similar to patients with PolyA expansion in HOXD13 protein suggesting a functional hierarchy among HoxD gene clusters (155). It has indeed been shown in mice models that the PolyA expanded allele of Hoxd13 interferes with the functions of Hoxd11 and Hoxd1, and is likely to have a dominant-negative effect over them (156). Based on the observations on the families with Hand–foot–genital syndrome and mice models, very similar mechanism was also proposed for PolyA expansion in HOXA13 protein, (149, 157-158). PolyA expansion in six other transcription factors (RUNX2, ZIC2, FOXL2, SOX-3, PHOX2B and ARX) thought to result in the loss of protein function because clinical symptoms of patients with PolyA expansions are not noticeably different from those with loss-of-function mutations such as deletions or frame-shift mutations (reviewed in 149) (Table 2). Thus, the dominant nature of the phenotype is likely to result from haploinsufficiency.

Amongst the PolyA expansion disorders, oculopharyngeal muscular dystrophy (OPMD) is unique because it is a progressive, late onset dominantly inherited syndrome and thus resembles the disease mechanism of PolyQ disorders. The OPMD is also unique because the affected gene (PABPN1) does not code for a transcription factor; PABPN1 protein binds to nascent poly-adenine tails of pre-mRNAs and regulates their formation and length (159-160). PABPN1 protein with expanded PolyA tract tend to form insoluble and ubiquitin-positive aggregates in the nucleus and are toxic to the cell (161, 162). While such aggregates recruit nucleoplasmic RNA, they do not seem to affect the length of the poly-adenine tails of pre-mRNAs, suggesting a gain-of-function effect (162). Analogous to PolyQ disorders, longer lengths of PolyA tracts directly correlate with disease severity (163).

3.2.3. TRS in the coding region: disorders with other amino acid repeats

Mutations in the gene encoding cartilage oligomeric matrix protein (COMP) are associated with two forms of autosomal dominant skeletal dysplasias (pseudoachondroplasia and multiple epiphyseal dysplasia), characterized by variable short stature, joint laxity and early-onset degenerative joint pathologies (164) (Table 2). One of the known mutations is the instability in codons that code for 5 consecutive aspartic acid residues within COMP protein, making it by far the shortest disease-causing repeat expansion mutations described (165). A unique feature of this repeat is that both expansion and shortening of the repeat cause the same disease; deletion of one repeat (4 aspartate residues) cause pseudoachondroplasia and their expansion cause multiple epiphyseal dysplasia (6 aspartate residues) and pseudoachondroplasia (7 aspartate residues) (165). These mutations are likely to impart a loss-of-function effect on COMP as null mutations also result in the same phenotype (165).

The CTG triplet repeat expansion in JPH3 gene associated with Huntington's Disease-like 2, like HD, also has the potential to code for amino acid repeats (65).The CTG repeat is contained within an alternatively spliced exon that has multiple splice sites. Alternate splicing in JPH3 gene could results in transcripts having different reading frames such that the CTG may fall outside (3'UTR) or within the coding region but in two different reading frames, to code polyalanine or polyleucine (65). No study has however confirmed the translational properties of these transcripts, and therefore their involvement in the disease in yet to be established.

4. FACTORS MODULATING TRS INSTABILITY IN THE GENOME

A majority of the TRS implicated in the disorders discussed above are polymorphic with regard to repeat length but their pathological threshold varies depending on whether the TRS lie in the coding or the non-coding region, and if present in the coding region, whether it would code for glutamine or other amino acids (Table 1 and 2) (166, 167). For example, expansions are massive (>1000 repeats) for TRS in the non-coding regions whereas in the coding regions they are relatively smaller (normally <200 repeats) (Table 1 and 2). Within the group of coding TRS, repeats coding for PolQ show larger length as against short and interrupted repeats that code for PolyA stretches (167). At least for disorder associated with CAG/CTG repeat motifs, the instability is dependent on repeat length as expansion becomes probable only when the repeat length reaches a particular threshold for respective loci (Table 1 and 2). Several studies have proposed a variety of molecular mechanisms to explain the TRS expansions associated with human disorders (reviewed in 7, 78, 168-170). These include DNA replication slippage, DNA damage repair and meiotic recombination. Although whether one or all of these processes contribute to the expansion of repeat is unknown, it is widely believed that the secondary structure that the repeat tracts might adopt could be a critical step in the initiation of expansion process (7, 166, 168, 171-173).

4.1. Replication and TRS expansion

Repeat instability can occur during replication because the extent of repeat instability is known to be dependent on the location of the replication origin (174). For example, the distance between the TRS stretches and the replication origin and the orientation of the TRS with respect to the replication origin were shown to affect the stability of the TRS in various model systems (175-178). Similarly, replication fork dynamics, repeat length, and CpG methylation of the genome closer to the TRS are also known to affect the repeat stability (179, 180). For example, CTG and CGG tracts are more prone to expansion when found in the leading strand, rather than in lagging strand of the replication fork (176, 180, 181). Interestingly, an increased degree of repeat instability was observed when TRS motifs were tested in yeast models deficient in replication proteins (182-184). Corroborating these results from yeast models, studies with human cell lines have shown that the DNA replication initiation regions fall close (<3 kb) to the CAG/CTG tracts of at least 4 different TRS loci associated with neurological disorders (185). Taken together, these studies underscore the role for origins of DNA replication in influencing the TRS length instability and favor the replication slippage model (186-187). In this model, TRS regions in the daughter strand thought to form

'loop-out' structures during DNA replication without disturbing the paring with the parent strand, eventually leading to an increase in the number of repeats in the newly formed DNA strand (172, 186).

4.2. Post-replication repair and TRS expansion

In addition to replication, repeat stability can also be influenced by defects in DNA repair mechanisms, especially in non-proliferating cells (188, 189). DNA repair process might get triggered because of the unusual DNA structure formed within the TRS and might get uncorrected by the repair mechanism (166). For example, single strands of CNG repeat tracts are inefficiently repaired during meiotic recombination in yeast model (180), and such sequence are associated with double strand breaks, leading to the repeat instability (190). An intriguing observation on the TRS instability HD comes from a moue model; the expanded HD allele becomes very unstable in striatal tissue as compared to other regions of the brain which eventually lead to increased load of PolyQ proteins and neuronal cell death (189). Molecular mechanism of such region-specific expansion is yet to be understood.

4.3. Recombination and TRS instability

Recent studies suggest that recombination mediated gene conversion play a significant role in repeat expansion (191-193). Gene conversion is a non-reciprocal transfer of genetic information wherein there is an alteration in the acceptor allele with no change in the donor allele and no change in the flanking sequence. As expansion in one allele is generally not accompanied by contraction in other allele in disorder involving dynamic TRS, gene conversion is likely to explain the process of disease-associated TRS instability. Using model systems, it has been demonstrated that gene conversion mediates several folds expansions of CTG/CAG repeats irrespective of the orientation of TRS tracts, suggesting that large expansions of TRS observed in HD and DM could be the product of such process (192). Population studies have provided evidence for the gene conversion event in the genesis of novel alleles bearing CAG tracts in SCA3 (194). Unequal recombination has been proposed as an alternate mechanism for TRS instability in PolyA disorders because PolyA tracts are coded by imperfect triplet repeats and slipping model would not explain intergenerational instability observed in disorders such as polysyndactyly (195). Analyses of mutant alleles in polysyndactyly did reveal that the expanded alleles were derived from recombination between two mispaired normal alleles (195).

4.4. TRS in the coding region and codon bias

Anomalous DNA structure that the TRS might form is thought to be one of the prerequisite for initiation of the repeat instability in the genome (196-197). Among the various repeats tested, the CTG/CAG and CGG/CCG repeat tracts were proven to have a higher potential to form secondary DNA structure (7, 198-199). In agreement with this view, a majority of the TRS that code for PolyQ domains, for example, are GC rich (1, 76, 200-201). However, this pattern did not seem to differ between the TRS encoding polymorphic and non-polymorphic PolyQ, suggesting that some *cis*-acting elements (flanking sequence) might regulate the instability of TRS (172, 202, 203). A recent study also suggests that codon bias is one of the limiting factors in the motif selection for TRS (76). This study demonstrates that rare codons, despite their GC-richness, are not favored when coding amino acid repeat tracts. In this regard, it is interesting to note that clustering of rare codons within a narrow region significantly reduce the half-life of the mRNA (204 - 206). Thus, mRNA stability could be one of the factors that work against the usage of rare codons in the TRS (76).

4.5. Parent-of-origin effect in TRS expansion

Gender bias in the transmission of unstable repeats from the parent to offspring is a common feature of dynamic mutations (Table 1 and 2). For example, the transmission of TRS through males was less stable than that through females for genes involved in DRPLA (207), HD (208), and SCA1 (209). However, it is the female sex in the case of FRDA (210) and SCA7 (211). The underlying cause for this sex-dependent transmission bias is thought to be difference in the meiotic process in the germ cell or may be determined at the time of embryonic development itself (211, 212). Alternatively, sex bias in the recombination process could be the reasons for the TRS instability. A recent study shows that the sex-specific recombination rates for ~5 mega base genomic region spanning the six genes associated with TRS expansion and disorder strongly correlate with the parental gender that positively influences the repeat instability (76). Since the recombination rate is known to be non-homogeneous in the genome (213) and that there are regions in the genome where recombination rate is particularly high in women and particularly low in men or vice versa (214), the regional differences in the recombination rate might influence the stability of TRS that they harbor (76).

4.6. Role of evolution in TRS instability

Length of TRS and its distribution within a gene sequence is nonrandom and strongly biased, and does have some selection constraints. Amino acid repeats coded by TRS are abundant in the human proteome but only a few of them are known to play any functional role (75-76, 215-217). However, a significant proportion of such genes are linked to OMIM and Morbid databases suggesting their potential association with genetic disorders (76, 218-220). Conservation of the non-coding TRS in gene orthologous is rather low; however, TRS in the coding region does show conservation at least among mammals, but not in their size or variability in the population, suggesting differential selection for TRS bearing genes (76, 201, 217, 221). It is of interest to note that, TRS encoded larger stretches of amino acid repeats are far more frequent in orphan proteins as compared to familial proteins, suggesting that the two forms of repeat coding genes are subjected to differential selection constraints (76, 201). Based on these observations, Siwach *et al* (76) proposed that the recently evolved solitary genes might acquire longer TRS stretches because of the weaker constraints placed on them. However, expansion beyond a "threshold" could become pathogenic and therefore would not be fixed in the population (76). Corroborating this view, pathogenic genes bearing longer TRS do represent orphan groups, and do not have orthologues in invertebrates (76).

5. PERSPECTIVE AND CONCLUDING REMARKS

While on one hand the genome tries to copy itself with utmost fidelity, other mechanisms like replication, recombination and the evolution of sex, on the other hand, attempts to increase variations in the genome. Variations make species more robust and adaptive to changes but considerable number of individuals in the population will have to suffer in order to increase the option of selection. Among the different types of variations known in the genome, TRS provide a wide range of genotypes (>10) at a given locus as against single nucleotide polymorphism that offer, at the most, four possible genotypes, or insertion-deletion polymorphism that offer only two genotypes. Thus instability in TRS falling within the gene sequence has greater potential to increase the phenotypic variability in the population and therefore likely to provide increased fitness under natural selection (222, 223). For example, length variations in TRS falling in the coding regions of the Alx-4 and Runx-2 genes were shown to quantitatively associate with evolution of limb and skull morphology in dogs (224). Similarly, TRS variations in the Avpr1a gene generate diversity in socio-behavioral traits in voles by modulating the expression level of the gene (225). Thus, the quantitative effects of the repeat-length variations in the TRS appear to function as "tuning knobs" for gene regulation and may have been favored by evolution as a source of phenotypic variability (222, 223).

6. ACKNOWLEDGMENTS

Work on the tandem repeats in our laboratory was supported by a research grant from the Ministry of Human Resource Development, Government of India, to SG. PS was supported by a research fellowship from the Council of Scientific and Industrial Research, Government of India.

7. REFERENCES

1. J. M. Hancock and M. Simon: Simple sequence repeats in proteins and their significance for network evolution. *Gene* 345(1), 113-118 (2005)
2. G. Toth, Z. Gaspari and J. Jurka: Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res* 10, 967-981 (2000)
3. Y. C. Li, A. B. Korol, T. Fahima, A. Beiles and E. Nevo: Microsatellites: genomic Distribution, putative functions and mutational mechanisms: a review. *Mol Ecol* 11, 2453-2465 (2002)
4. Y. C. Li, A. B. Korol, T. Fahima, A. Beiles and E. Nevo: Microsatellites with in genes: Structure, function and evolution. *Mol Biol Evol* 21(6), 991-1007 (2004)
5. D. Tautz and M. Renz: Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res* 12(10), 4127-4138 (1984)
6. D. Field and C. Wills: Long, polymorphic microsatellites in simple organisms. *Proc Biol Sci* 263(1367), 209-215 (1996)
7. C. E. Pearson and R. R. Sinden: Trinucleotide repeat DNA structures: dynamic mutations from dynamic DNA. *Curr Opin Struct Biol* 8(3), 321-330 (1998)
8. G. F. Richard, F. Paques: Mini- and microsatellite expansions: the recombination connection. *EMBO Rep* 1(2), 122-126 (2000)
9. M. A. Costa Lima and M. M. Pimentel: Dynamic mutation and human disorders: the spinocerebellar ataxias. *Int J Mol Med* 13(2), 299-302 (2004)
10. G. R. Sutherland and R. I. Richards: Simple tandem DNA repeats and human genetic disease. *Proc Natl Acad Sci U S A 92*(9), 3636-3641 (1995)
11. D. G. Monckton and C. T. Caskey: Unstable triplet repeat diseases. *Circulation* 91(2), 513-520 (1995)
12. S. L. Toutenhoofd, F. Garcia, D. A. Zacharias, R. A. Wilson and E. E. Strehler: Minimum CAG repeat in the human calmodulin-1 gene 5' untranslated region is required for full expression. *Biochim Biophys Acta* 1398(3), 315-320 (1998)
13. E. A. Grzybowska, A. Wilczynska and J. A. Siedlecki: Regulatory functions of 3'UTRs. *Biochem Biophys Res Commun* 288(2), 291-295 (2001)
14. D. E. Riley and J. N. Krieger: Short tandem repeats are associated with diverse mRNAs encoding membrane-targeted proteins. *Bioessays* 26(4), 434-444 (2004)
15. R. Meloni, V. Albanese, P. Ravassard, F. Treilhou and J. Mallet: A tetranucleotide polymorphic microsatellite, located in the first intron of the tyrosine hydroxylase gene, acts as a transcription regulatory element *in vitro*. *Hum Mol Genet* 7(3), 423-428 (1998)
16. V. Albanese, N. F. Biguet, H. Kiefer, E. Bayard, J. Mallet and R. Meloni: Quantitative effects on gene silencing by allelic variation at a tetranucleotide microsatellite. *Hum Mol Genet* 10(17), 1785-1792 (2001)
17. B. Borstnik and D. Pumpernik: Tandem repeats in protein coding regions of primate genes. *Genome Res* 12(6), 909-915 (2002)
18. S. Ceru, S. Rabzelj, N. Kopitar-Jerala, V. Turk and E. Zerovnik: Protein aggregation as a possible cause for pathology in a subset of familial Unverricht-Lundborg disease. *Med Hypotheses* 64(5), 955-959 (2005)
19. M. D. Lalioti, H. S. Scott HS, C. Buresi, C. Rossier, A. Bottani, M. A. Morris, A. Malafosse and S. E. Antonarakis: Dodecamer repeat expansion in cystatin B gene in progressive myoclonus epilepsy. *Nature* 386(6627), 847-851 (1997)
20. A. V. Delgado-Escueta, S. Ganesh and K. Yamakawa: Advances in the genetics of progressive myoclonus epilepsy. *Am J Med Genet* 106(2), 129-138 (2001)
21. M. D. Lalioti, H. S. Scott and S. E. Antonarakis: Altered spacing of promoter elements due to dodecamer repeat expansion contributes to reduced expression of the cystatin B gene in EPM1. *Hum Mol Gent* 8(9), 1791-1798 (1999)
22. B. A. Oostra and P. Chiurazzi: The fragile X gene and its function. *Clin Genet* 60, 399-408. (2001)
23. W. T. O'Donnell and S. T. Warren: A decade of molecular studies of fragile X syndrome. *Annu Rev Neurosci* 25, 315-338 (2002)
24. R. S. Hansen, S. M. Gartler, C. R. Scott, S. H. Chen and C. D. Laird: Methylation analysis of CGG sites in the CpG island of the human FMR1 gene. *Hum Mol Genet* 1(8), 571-578 (1992)
25. E. W. Khandjian, F. Corbin, S. Woerly and F. Rousseau: The fragile X mental retardation protein is associated with ribosomes. *Nat Genet* 12(1), 91-93 (1996)

26. B. Laggerbauer, D. Ostareck, E. M. Keidel, A. Ostareck-Lederer and U. Fischer U: Evidence that fragile X mental retardation protein is a negative regulator of translation. *Hum Mol Genet* 10(4), 329-338 (2001)
27. V. Brown, P. Jin, S. Ceman, J.C. Darnell, W. T. O'Donnell, S. A. Tenenbaum, X. Jin, Y. Feng, K. D. Wilkinson, J. D. Keene, R. B. Darnell and S. T. Warren: Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* 107(4), 477-487 (2001)
28. J. Gecz, B. A. Oostra, A. Hockey, P. Carbonell, G. Turner, E. A. Haan, G. R. Sutherland and J. C. Mulley: FMR2 expression in families with FRAXE mental retardation. *Hum Mol Genet* 6(3), 435-441 (1997)
29. L. A. Pennacchio, A. E. Lehesjoki, N. E. Stone, V. L. Willour, K. Virtaneva, J. Miao, E. D'Amato, L. Ramirez, M. Faham, M. Koskiniemi, J. A. Warrington, R. Norio, A. de la Chapelle, D. R. Cox and R. M. Myers: Mutations in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1) *Science* 271(5256), 1731-1734 (1996)
30. K. De Boulle, A. J. Verkerk, E. Reyniers, L. Vits, J. Hendrickx, B. Van Roy, F. Van den Bos, E. de Graaff, B. A. Oostra and P. J. Willems: A point mutation in the FMR-1 gene associated with fragile X mental retardation. *Nat Genet* 3(1), 31-35 (1993)
31. C. Verheij, E. de Graaff, C. E. Bakker, R. Willemsen, P. J. Willems, N. Meijer, H. Galjaard, A. J. Reuser, B. A. Oostra and A. T. Hoogeveen: Characterization of FMR1 proteins isolated from different tissues. *Hum Mol Genet* 4(5), 895-901 (1995)
32. G. R. Sutherland: Rare fragile sites. *Cytogenet Genome Res* 100(1-4), 77-84 (2003)
33. C. Jones, R. Mullenbach, P. Grossfeld, R. Auer, R. Favier, K. Chien, M. James, A. Tunnacliffe and F. Cotter: Co-localisation of CCG repeats and chromosome deletion breakpoints in Jacobsen syndrome: evidence for a common mechanism of chromosome breakage. *Hum Mol Genet* 9(8), 1201-1208 (2000)
34. C. Jones, L. Penny, T. Mattina, S. Yu, E. Baker, L. Voullaire, W. Y. Langdon, G. R. Sutherland, R. I. Richards and A. Tunnacliffe: Association of a chromosome deletion syndrome with a fragile site within the proto-oncogene CBL2. *Nature* 376(6536), 145-149 (1995)
35. S. E. Holmes, E. E. O'Hearn, M. G. McInnis, D. A. Gorelick-Feldman, J. J. Kleiderlein, C. Callahan, N. G. Kwak, R. G. Ingersoll-Ashworth, M. Sherr, A. J. Sumner, A. H. Sharp, U. Ananth, W. K. Seltzer, M. A. Boss, A. M. Vieria-Saecker, J. T. Epplen, O. Riess, C. A. Ross and R. L. Margolis: Expansion of a novel CAG trinucleotide repeat in the 5' region of PPP2R2B is associated with SCA12. *Nat Genet* 23(4), 391-392 (1999)
36. S. E. Holmes, E. O. Hearn, C. A. Ross and R. L. Margolis: SCA12: an unusual mutation leads to an unusual spinocerebellar ataxia. *Brain Res Bull* 56(3-4), 397-403 (2001)
37. C. A. Thornton, J. P. Wymer, Z. Simmons, C. McClain and R. T. Moxley 3rd: Expansion of the myotonic dystrophy CTG repeat reduces expression of the flanking DMAHP gene. *Nat Genet* 16(4), 407-409 (1997)
38. N. Alekseeva, A. S. Kablinger, J. Pinkston, E. C. onzalez-Toledo and A. Minagar: Hereditary ataxia and behavior. *Adv Neurol* 96, 275-283 (2005)
39. V. Campuzano, L. Montermini, M. D. Molto, L. Pianese, M. Cossee, F. Cavalcanti, E. Monros, F. Rodius, F. Duclos, A. Monticelli, F. Zara, J. Canizares, H. Koutnikova, S. I. Bidichandani, C. Gellera, A. Brice, P. Trouillas, G. De Michele, A. Filla, R. De Frutos, F. Palau, P. I. Patel, S. Di Donato, J. L. Mandel, S. Cocozza, M. Koenig and M. Pandolfo: Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science* 271(5254), 1423-1427 (1996)
40. M. B. Delatycki, M. Knight, M. Koenig, M. Cossee, R Williamson and S. M. Forrest: G130V, a common FRDA point mutation, appears to have arisen from a common founder. *Hum Genet* 105(4), 343-346 (1999)
41. P. Cavadini, H. A. O'Neill, O. Benada, and G. Isaya: Assembly and iron-binding properties of human frataxin, the protein deficient in Friedreich ataxia. *Hum Mol Genet* 11(3), 217-227 (2002)
42. L. Pianese, M. Turano, M. S. Lo Casale, I. De Biase, M. Giacchetti, A. Monticelli, C. Criscuolo, A. Filla and S. Cocozza: Real time PCR quantification of frataxin mRNA in the peripheral blood leucocytes of Friedreich ataxia patients and carriers. *J Neurol Neurosurg Psychiatry* 75(7), 1061-3 (2004)
43. E. Grabczyk and K. Usdin: 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(14), 2815-2822 (2000)
44. E.Greene, L. Mahishi, A. Entezam, D. Kumari and K. Usdin: Repeat-induced epigenetic changes in intron 1 of the frataxin gene and its consequences in Friedreich ataxia. *Nucleic Acids Res* 1-8 (2007) doi:10.1093/nar/gkm271
45. T. Matsuura, T. Yamagata, D. L. Burgess, A. Rasmussen, R. P. Grewal, K. Watase, M. Khajavi, A. E. McCall, C. F. Davis, L. Zu, M. Achari, S. M. Pulst, E. Alonso, J. L. Noebels, D. L. Nelson, H. Y. Zoghbi and T. Ashizawa: Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. *Nat Genet* 26(2), 91-94 (2000)
46. M. Wakamiya, T. Matsuura, Y. Liu, G. C.Schuster, R. Gao, W. Xu, P. S. Sarkar, X. Lin and T. Ashizawa T: The role of ataxin 10 in the pathogenesis of spinocerebellar ataxia type 10. *Neurology* 67(4), 607-613 (2006)
47. C. L. Liquori, K. Ricker, M. L. Moseley, J. F. Jacobsen, W. Kress, S. L. Naylor, J. W. Day and L. P. Ranum: Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* 293(5531), 864-867 (2001)
48. T. B. Rajavashisth, A. K. Taylor, A. Andalibi, K. L. Svenson and A. J. Lusis: Identification of a zinc finger protein that binds to the sterol regulatory element. *Science* 245(4918), 640-643 (1989)
49. V. R. Gerbasi and A. J. Link: The myotonic dystrophy type-2 protein ZNF9 is part of an ITAF complex that promotes cap-independent translation. *Mol Cell Proteomics* (2007)
50. J. M. Margolis, B. G. Schoser, M. L. Moseley, J. W. Day and L. P. Ranum: DM2 intronic expansions: evidence for CCUG accumulation without flanking sequence or effects on ZNF9 mRNA processing or protein expression.*Hum Mol Genet* 15(11), 1808-1815 (2006)
51. Y. H. Fu, A. Pizzuti, R. G. Fenwick Jr, J. King, S. Rajnarayan, P. W. Dunne, J. Dubel, G. A. Nasser, T.

Ashizawa, and P. de Jong et al.: An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* 255(5049), 1256-1258 (1992)
52. M. Carrasco, J. Canicio, M. Palacin, A. Zorzano and P Kaliman: Identification of intracellular signaling pathways that induce myotonic dystrophy protein kinase expression during myogenesis. *Endocrinology* 143(8), 3017-3025 (2002)
53. Y. H. Fu, D. L. Friedman, S. Richards, J. A. Pearlman, R. A. Gibbs, A. Pizzuti, T. Ashizawa, M. B. Perryman, G. Scarlato, R. G. Fenwick RG Jr et al.:Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. *Science* 260(5105), 235-238 (1993)
54. Y. H. Wang, S. Amirhaeri, S. Kang, R. D. Wells and J. D. Griffith: Preferential nucleosome assembly at DNA triplet repeats from the myotonic dystrophy gene. *Science* 265 (5172), 669-671 (1994)
55. C. A. Boucher, S. K. King, N. Carey, R. Krahe, C. L.Winchester, S. Rahman, T. Creavin, P. Meghji, M. E. Bailey and F. L. Chartier S. D. Brown, M.J.Siciliano, and K.J.Johnson: A novel homeodomain-encoding gene is associated with a large CpG island interrupted by the myotonic dystrophy unstable (CTG)n repeat. *Hum Mol Genet* 4(10), 1919-1925 (1995)
56. P. S. Sarkar, B. Appukuttan, J. Han, Y. Ito, C. Ai, W. Tsai, Y. Chai, J. T.Stout and S. Reddy: Heterozygous loss of Six5 in mice is sufficient to cause ocular cataracts. *Nat Genet* 25(1), 110-114 (2000)
57. K. L. Taneja, M. McCurrach, M. Schalling, D. Housman and R. H. Singer: Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J Cell Biol* 128(6), 995-1002 (1995)
58. L. T. Timchenko, N. A. Timchenko, C. T. Caskey and R. Roberts: Novel proteins with binding specificity for DNA CTG repeats and RNA CUG repeats: implications for myotonic dystrophy. *Hum Mol Genet* 5(1), 115-121 (1996)
59. A. V. Philips, L. T. Timchenko and T. A. Cooper: Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science* 280(5364), 737-741 (1998)
60. R. S. Savkur, A. V. Philips and T. A.Cooper: Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat Genet* 29(1), 40-47 (2001)
61. A. Mankodi, M. P. Takahashi, H. Jiang, C. L. Beck, W. J. Bowers, R. T. Moxley, S. C. Cannon and C. A. Thornton: Expanded CUG repeats trigger aberrant splicing of ClC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol Cell* 10(1), 35-44 (2002)
62. N. Sergeant, B. Sablonniere, S. Schraen-Maschke, A. Ghestem, C. A. Maurage, A. Wattez, P. Vermersch and A. Delacourte: Dysregulation of human brain microtubule-associated tau mRNA maturation in myotonic dystrophy type 1. *Hum Mol Genet* 10(19), 2143-2155 (2001)
63. R. S. Savkur, A. V. Philips, T. A. Cooper, J. C. Dalton, M. L. Moseley, L. P. Ranum and J. W. Day: Insulin receptor splicing alteration in myotonic dystrophy type 2. *Am J Hum Genet* 74(6), 1309-1313 (2004)
64. J. W. Day and L. P. Ranum: RNA pathogenesis of the myotonic dystrophies. *Neuromuscul Disord* 15(1), 5-16 (2005)
65. S. E. Holmes, E. O'Hearn, A. Rosenblatt, C. Callahan, H. S. Hwang, R. G. Ingersoll-Ashworth, A. Fleisher, G. Stevanin, A. Brice, N. T. Potter, C. A. Ross and R. L. Margolis: A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2. *Nat Genet* 29(4), 377-378 (2001)
66. D. D. Rudnicki, S. E. Holmes, M.W. Lin, C. A. Thornton, C. A. Ross and R. L Margolis: Huntington's disease--like 2 is associated with CUG repeat-containing RNA foci. *Ann Neurol* 61(3), 272-282 (2007)
67. M. D. Koob, M. L. Moseley, L. J. Schut, K. A. Benzow, T. D. Bird, J. W. Day and L. P. Ranum: An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8) *Nat Genet* 21(4), 379-384 (1999)
68. J. P. Nemes, K. A. Benzow, M. L. Moseley, L. P. Ranum and M. D. Koob: The SCA8 transcript is an antisense RNA to a brain-specific transcript encoding a novel actin-binding protein (KLHL1). *Hum Mol Genet* 9(10), 1543-1551 (2000)
69. Y. He, T. Zu, K. A. Benzow, H. T. Orr, H. B. Clark and M. D. Koob: Targeted deletion of a single Sca8 ataxia locus allele in mice causes abnormal gait, progressive loss of motor coordination, and Purkinje cell dendritic deficits. *J Neurosci* 26(39), 9975-9982 (2006)
70. M. L. Moseley, T. Zu, Y. Ikeda, W. Gao, A. K. Mosemiller, R. S. Daughters, G. Chen, M. R. Weatherspoon, H. B. Clark, T. J. Ebner, J. W. Day and L.P. Ranum: Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. *Nat Genet* 38(7), 758-769 (2006)
71. D. Metzgar, J. Bytof and C. Wills: Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Res* 10(1), 72-80 (2000)
72. E. M. Marcotte, M. Pellegrini, T. O. Yeates and D.Eisenberg: A census of protein repeats. *J Mol Biol* 293, 151-160 (1999)
73. M. K. Kalita, G. Ramasamy, S. Duraisamy, V. S. Chauhan and D. Gupta: Prot Repeats DB: a database of amino acid repeats in genomes. *BMC Bioinformatics* 7, 336 (2006) doi:10.1186/1471-2105-7-336
74. D. P. Depledge and A. Dalby: COPASAAR--a database for proteomic analysis of single amino acid repeats. *BMC Bioinformatics* 6, 196 (2005) doi:10.1186/1471-2105-6-196
75. N.G. Faux, S. P. Bottomley, A. M. Lesk, J. A. Irving, J. R. Morrison, M. G. de la Banda and J. C. Whisstock: Functional insights from the distribution and role of homopeptide repeat-containing proteins. *Genome Res* 15(4), 537-551 (2005)
76. P. Siwach, S. D. Pophaly and S. Ganesh: Genomic and evolutionary insights into genes encoding proteins with single amino acid repeats. *Mol Biol Evol* 23(7), 1357-1369 (2006)
77. J. R. Gatchel and H. Y. Zoghbi: Diseases of unstable repeat expansion: mechanisms and common principles. *Nat Rev Genet* 6(10), 743-755 (2005)
78. K. H. Fischbeck: Polyglutamine expansion neurodegenerative disease. *Brain Research Bulletin* 56(3-4), 161-163 (2001)
79. A. McCampbell, J. P. Taylor, A. A. Taye, J. Robitschek, M. Li, J. Walcott, D. Merry, Y. Chai, H.

Paulson, G. Sobue K. H. Fischbeck: CREB-binding protein sequestration by expanded polyglutamine. *Hum Mol Genet* 9(14), 2197-2202 (2002)
80. J. M Warrick, H. L. Paulson, G. L. Gray-Board, Q. T. Bui, K. H. Fischbeck, R. N. Pittman and N. M. Bonini: Expanded polyglutamine protein forms nuclear inclusions and causes neural degeneration in Drosophila. *Cell* 93(6), 939-949 (1998)
81. E. N. Burright, H. T. Orr and H. B. Clark: Mouse models of human CAG repeat disorders. *Brain Pathol* 7(3), 965-977 (1997)
82. L. Mangiarini, K. Sathasivam, A. Mahal, R. Mott, M. Seller and G. P. Bates: Instability of highly expanded CAG repeats in mice transgenic for the Huntington's disease mutation. *Nat Genet* 15(2), 197-200 (1997)
83. E. N. Burright, H. B. Clark, A. Servadio, T. Matilla, R. M. Feddersen, W. S. Yunis, L. A. Duvick, H. Y. Zoghbi and H. T. Orr: SCA1 transgenic mice: a model for neurodegeneration caused by an expanded CAG trinucleotide repeat. *Cell* 82(6), 937-948 (1995)
84. M. Katsuno, H. Adachi, A. Inukai and G. Sobue: Transgenic mouse models of spinal and bulbar muscular atrophy (SBMA). *Cytogenet Genome Res* 100(1-4), 243-251 (2003)
85. P. McManamny, H. S. Chy, D. I. Finkelstein, R. G. Craythorn, P. J. Crack, I. Kola, S. S. Cheema, M. K. Horne, N. G. Wreford, M. K. O'Bryan, D. M. De Kretser and J. R. Morrison: A mouse model of spinal and bulbar muscular atrophy. *Hum Mol Genet* 11(18), 2103-2111 (2002)
86. J. M. Ordway, S. Tallaksen-Greene, C. A. Gutekunst, E. M. Bernstein, J. A. Cearley, H. W. Wiener, L. S. Dure 4th, R. Lindsey, S. M. Hersch, R. S. Jope, R. L. Albin and P. J. Detloff: Ectopically expressed CAG repeats cause intranuclear inclusions and a progressive late onset neurological phenotype in the mouse. *Cell* 91(6), 753-763 (1997)
87. S. Kotliarova, N. R. Jana, N. Sakamoto, M. Kurosawa, H. Miyazaki, M. Nekooki, H. Doi, Y. Machida, H. K. Wong, T. Suzuki, C. Uchikawa, Y. Kotliarov, K. Uchida, Y. Nagao, U. Nagaoka, A. Tamaoka, K. Oyanagi, F. Oyama and N. Nukina: Decreased expression of hypothalamic neuropeptides in Huntington disease transgenic mice with expanded polyglutamine-EGFP fluorescent aggregates. *J Neurochem* 93(3), 641-653 (2005)
88. M. Yamada, M. Shimohata, T. Sato, S. Tsuji and H. Takahashi: Polyglutamine disease: recent advances in the neuropathology of dentatorubral-pallidoluysian atrophy. *Neuropathology* 26(4), 346-351 (2006)
89. M. Tanaka, I. Morishima, T. Akagi, T. Hashikawa and N. Nukina: Intra- and intermolecular beta-pleated sheet formation in glutamine-repeat inserted myoglobin as a model for polyglutamine diseases. *J Biol Chem* 276(48), 45470-45475 (2001)
90. M. F. Perutz, T. Johnson, M. Suzuki and J. T. Finch: Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *Proc Natl Acad Sci U S A* 91(12), 5355-5358 (1994)
91. A. Ciechanover and P. Brundin: The ubiquitin proteasome system in neurodegenerative diseases: sometimes the chicken, sometimes the egg. *Neuron* 40(2), 427-446 (2003)
92. H. C. Ardley and P. A. Robinson: The role of ubiquitin-protein ligases in neurodegenerative disease. *Neurodegener Dis* 1(2-3), 71-87 (2004)
93. Y. Chai, S. L. Koppenhafer, S. J. Shoesmith, M. K. Perez an H. L. Paulson: Evidence for proteasome involvement in polyglutamine disease: localization to nuclear inclusions in SCA3/MJD and suppression of polyglutamine aggregation *in vitro*. *Hum Mol Genet* 8(4), 673-682 (1999)
94. P. Rusmini, D. Sau, V. Crippa, I. Palazzolo, F. Simonini, E. Onesto, L. Martini and A. Poletti: Aggregation and proteasome: the case of elongated polyglutamine aggregation in spinal and bulbar muscular atrophy. *Neurobiol Aging* 28(7), 1099-1111 (2007)
95. N. F. Bence, R. M. Sampat and R. R. Kopito: Impairment of the ubiquitin-proteasome system by protein aggregation. *Science* 292(5521), 1552-1555 (2001)
96. Y. Kobayashi and G. Sobue: Protective effect of chaperones on polyglutamine diseases. *Brain Res Bull* 56(3-4), 165-168 (2005)
97. N. R. Jana, P. Dikshit, A. Goswami, S. Kotliarova, S. Murata, K. Tanaka and N. Nukina: Co-chaperone CHIP associates with expanded polyglutamine protein and promotes their degradation by proteasomes. *J Biol Chem* 280(12), 11635-11640 (2005)
98. J. Y. Choi, J. H. Ryu, H. S. Kim, S. G. Park, K. H. Bae, S. Kang, P. K. Myung, S. Cho, Park BC and H. Lee do: Co-chaperone CHIP promotes aggregation of ataxin-1. *Mol Cell Neurosci* 34(1), 69-79 (2007)
99. Y. C. Tsai, P. S. Fishman, N. V, Thakor and G. A. Oyler: Parkin facilitates the elimination of expanded polyglutamine proteins and leads to preservation of proteasome function. *J Biol Chem* 278(24), 22044-22055 (2003)
100. C. J. Cummings, Y. Sun, P. Opal, B. Antalffy, R. Mestril, H. T. Orr, W. H. Dillmann and H. Y. Zoghbi: Over-expression of inducible HSP70 chaperone suppresses neuropathology and improves motor function in SCA1 mice. *Hum Mol Genet* 10(14), 1511-1518 (2001)
101. H. Y. Chan, J. M. Warrick, G. L. Gray-Board, H. L. Paulson and N. M. Bonini: Mechanisms of chaperone suppression of polyglutamine disease: selectivity, synergy and modulation of protein solubility in Drosophila. *Hum Mol Genet* 9(19), 2811-2820 (2000)
102. H. Sakahira, P. Breuer, M. K. Hayer-Hartl and F. U. Hartl: Molecular chaperones as modulators of polyglutamine protein aggregation and toxicity. *Proc Natl Acad Sci U S A* 4, 16412-16418 (2002)
103. A. Michalik and C. Van Broeckhoven: Proteasome degrades soluble expanded polyglutamine completely and efficiently. *Neurobiol Dis* 16(1), 202-211 (2004)
104. D. Martindale, A. Hackam, A. Wieczorek, L. Ellerby, C. Wellington, K. McCutcheon, R. Singaraja, P. Kazemi-Esfarjani, R. Devon, S. U. Kim, D. E. Bredesen, F. Tufaro and M. R. Hayden: Length of huntingtin and its polyglutamine tract influences localization and frequency of intracellular aggregates. *Nat Genet* 18(2), 150-154 (1999)
105. R. R. Kopito: Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol* 10(12), 524-530 (2002)

106. J. P. Taylor, F. Tanaka, J. Robitschek, C. M. Sandoval, A. Taye, S. Markovic-Plese, K. H. Fischbeck: Aggresomes protect cells by enhancing the degradation of toxic polyglutamine-containing protein. *Hum Mol Genet* 12(7), 749-757 (2003)
107. F. Saudou, S. Finkbeiner, D. Devys and M. E. Greenberg: Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear inclusions. *Cell* 95(1), 55-66 (1998)
108. C. J. Cummings, E. Reinstein, Y. Sun, B. Antalffy, Y. Jiang, A. Ciechanover, H. T. Orr, A. L. Beaudet and H. Y. Zoghbi: Mutation of the E6-AP ubiquitin ligase reduces nuclear inclusion frequency while accelerating polyglutamine-induced pathology in SCA1 mice. *Neuron* 24(4), 879-892 (1999)
109. I. A. Klement, P.J. Skinner, M. D. Kaytor, H. Yi, S. M. Hersch, H. B. Clark, H. Y. Zoghbi and H. T Orr: Ataxin-1 nuclear localization and aggregation: role in polyglutamine-induced disease in SCA1 transgenic mice. *Cell* 95(1), 41-53 (1998)
110. E. J. Slow, R. K. Graham, A. P. Osmand, R. S. Devon, G. Lu, Y. Deng, J. Pearson, K. Vaid, N. Bissada, R. Wetzel, B. R. Leavitt an M. R. Hayden: Absence of behavioral abnormalities and neurodegeneration *in vivo* despite widespread neuronal huntingtin inclusions. *Proc Natl Acad Sci U S A* 102(32), 11402-11407 (2005)
111. A. Michalik and C. Van Broeckhoven: Pathogenesis of polyglutamine disorders: aggregation revisited. *Hum Mol Genet* 12 Spec No 2, R173-186 (2003)
112. C. A. Ross, J. D. Wood, G. Schilling, M. F. Peters, F. C. Jr Nucifora, J. K. Cooper, A. H. Sharp, R. L. Margolis and D. R. Borchelt: Polyglutamine pathogenesis. *Philos Trans R Soc Lond B Biol Sci* 354(1386), 1005-1011 (1999)
113. C. A. Ross: Polyglutamine pathogenesis: emergence of unifying mechanisms for Huntington's disease and related disorders. *Neuron* 35(5), 819-822 (2002)
114. G. Bates: Huntingtin aggregation and toxicity in Huntington's disease. Lancet, 361(9369), 1642-1644 (1999)
115. R. Walsh, E. Storey, D. Stefani, L. Kelly and V. Turnbull: The roles of proteolysis and nuclear localisation in the toxicity of the polyglutamine diseases. A review. *Neurotox Res* 7(1-2), 43-57 (2005)
116. P. Leznicki: Aggregation and toxicity of the proteins with polyQ repeats. *Postepy Biochem* 51(2), 215-222 (2005)
117. D.C.Rubinsztein: The roles of intracellular protein-degradation pathways in neurodegeneration. *Nature* 443(7113), 780-786 (2006)
118. B.Ravikumar,R. Duden, and D.C. Rubinsztein: Aggregate-prone proteins with polyglutamine and polyalanine expansions are degraded by autophagy. *Hum Mol Genet.* 11(9), 1107-1117 (2002)
119. K.B.Kegel, M.Kim, E.Sapp,C.McIntyre, J.G.Castaño, N.Aronin, and M.DiFiglia: Huntingtin expression stimulates endosomal-lysosomal activity, endosome tubulation, and autophagy. *J Neurosci.* 20(19), 7268-7278 (2000)
120. Z.H.Qin, Y.Wang, K.B.Kegel, A. Kazantsev, B.L.Apostol, L.M.Thompson, J.Yoder, N.Aronin, and M.DiFiglia: Autophagy regulates the processing of amino terminal huntingtin fragments. *Hum Mol Genet.* 12(24), 3231-3244 (2003)
121. A. Iwata, J.C.Christianson, M.Bucci, L.M.Ellerby, N.Nukina, L.S.Forno, and R.R. Kopito: Increased susceptibility of cytoplasmic over nuclear polyglutamine aggregates to autophagic degradation. *Proc Natl Acad Sci U S A.* 102(37), 13135-13140 (2005)
122. B. Ravikumar, Z. Berger, C.Vacher, C.J.O'Kane, and D.C. Rubinsztein: Rapamycin pre-treatment protects against apoptosis. *Hum Mol Genet*. 15(7), 1209-1216 (2006)
123. Z. Berger, B.Ravikumar, F.M.Menzies, L.G.Oroz, B.R.Underwood, M.N.Pangalos, I.Schmitt, U.Wullner, B.O.Evert, C.J.O'Kane, and D.C.Rubinsztein: Rapamycin alleviates toxicity of different aggregate-prone proteins. *Hum Mol Genet* 15(3), 433-442 (2006)
124. A. Iwata, B.E.Riley, J.A. Johnston, and R.R.Kopito: HDAC6 and microtubules are required for autophagic degradation of aggregated huntingtin. *J Biol Chem.* 280(48), 40282-40292 (2005)
125. B.Ravikumar, C.Vacher, Z.Berger, J.E.Davies, S.Luo, L.G. Oroz, F.Scaravilli, D.F.Easton, R.Duden, C.J.O'Kane, and D.C.Rubinsztein: Inhibition of mTOR induces autophagy and reduces toxicity of polyglutamine expansions in fly and mouse models of Huntington disease. *Nat Genet.* 36(6), 585-595 (2004)
126. U.B.Pandey, Z.Nie, Y.Batlevi, B.A.McCray, G.P.Ritson, N.B.Nedelsky, S.L.Schwartz, N.A.DiProspero, M.A.Knight, O.Schuldiner, R.Padmanabhan, M.Hild, D.L.Berry, D.Garza, C.C.Hubbert, T.P.Yao, E.H.Baehrecke, and J.P.Taylor: HDAC6 rescues neurodegeneration and provides an essential link between autophagy and the UPS. *Nature* 447(7146), 859-863 (2007)
127. B. E. Riley and H. T. Orr: Polyglutamine neurodegenerative diseases and regulation of transcription: assembling the puzzle. *Genes Dev* 20(16), 2183-2192 (2006)
128. D. Helmlinger, L. Tora and D. Devys: Transcriptional alterations and chromatin remodeling in polyglutamine diseases. *Trends Genet* 22(10), 562-570 (2006)
129. K. Nakamura, S.Y. Jeong, T. Uchihara, M. Anno, K. Nagashima, T. S Nagashima, S. Ikeda, S. Tsuji and I. Kanazawa: SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. *Hum Mol Genet* 10(14), 1441-1448 (2001)
130. A. R. La Spada, E. M. Wilson, D. B. Lubahn, A. E. Harding and K. H. Fischbeck: Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 352(6330), 77-79 (1991)
131. D. Helmlinger, S. Hardy, S. Sasorith, F. Klein, F. Robert, C. Weber, L. Miguet, N. Potier, A. Van-Dorsselaer, J. M. Wurtz, J. L. Mandel, L. Tora and D. Devys: Ataxin-7 is a subunit of GCN5 histone acetyltransferase-containing complexes. *Hum Mol Genet* 13(12), 1257-1265 (2004)
132. C. C. Tsai, H. Y. Kao, A. Mitzutani, E. Banayo, H. Rajan, M. McKeown and R. M.Evans: Ataxin 1, a SCA1 neurodegenerative disorder protein, is functionally linked to the silencing mediator of retinoid and thyroid hormone receptors. *Proc Natl Acad Sci U S A* 101(12), 4047-4052 (2004)
133. S. Zhang, L. Xu, J. Lee and T. Xu: Drosophila atrophin homolog functions as a transcriptional corepressor in multiple developmental processes. *Cell* 108(1), 45-56 (2002)

134. F. Li, T. Macfarlan, R. N. Pittman and D. Chakravarti: Ataxin-3 is a histone-binding protein with two independent transcriptional corepressor activities. *J Biol Chem* 277(47), 45004-145012 (2002)
135. C. Zuccato, M. Tartari, A. Crotti, D. Goffredo, M. Valenza, L. Conti, T. Cataudella, B. R. Leavitt, M. R. Hayden, T. Timmusk, D. Rigamonti and E. Cattaneo: Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled neuronal genes. *Nat Genet* 35(1), 76-83 (2003)
136. J. H. Cha, C. M. Kosinski, J. A. Kerner, S. A. Alsdorf, L. Mangiarini, S. W. Davies, J. B. Penney, G. P. Bates and a. B.Young: Altered brain neurotransmitter receptors in transgenic mice expressing a portion of an abnormal human huntington disease gene. *Proc Natl Acad Sci U S A* 95(11), 6480-6485 (1998)
137. R. Luthi-Carter, A. Strand, N. L. Peters, S. M. Solano, Z. R. Hollingsworth, A. S. Menon, A. S. Frey, B. S. Spektor, E. B. Penney, G. Schilling, C. A. Ross, D. R. Borchelt, S. J. Tapscott, A. B. Young, J. H. Cha JH and J. M Olson: Decreased expression of striatal signaling genes in a mouse model of Huntington's disease. *Hum Mol Genet* 9(9), 1259-1271 (2000)
138. R. Luthi-Carter, A. D. Strand, S. A. Hanson, C. Kooperberg, G. Schilling, A. R. La Spada, D. E. Merry, A. B. Young, C. A. Ross, D. R. Borchelt and J. M. Olson: Polyglutamine and transcription: gene expression changes shared by DRPLA and Huntington's disease mouse models reveal context-independent effects. *Hum Mol Genet* 11(17), 1927-1937 (2002)
139. A. Hodges, A. D. Strand, A. K. Aragaki, A. Kuhn, T. Sengstag, G. Hughes, L. A. Elliston, C. Hartog, D. R. Goldstein, D. Thu, Z. R. Hollingsworth, F. Collin, B. Synek, P.A. Holmans, A. B.Young, N. S. Wexler, M. Delorenzi, C. Kooperberg, S. J. Augood, R. L. Faull, J. M. Olson, L. Jones and R.Luthi-Carter: Regional and cellular gene expression changes in human Huntington's disease brain. *Hum Mol Genet* 15(6), 965-977 (2006)
140. J. S. Steffan, A. Kazantsev, O. Spasic-Boskovic, M. Greenwald, Y. Z. Zhu, H. Gohler, E. E. Wanker, G. P. Bates, D. E. Housman and L. M. Thompson: The Huntington's disease protein interacts with p53 and CREB-binding protein and represses transcription. *Proc Natl Acad Sci U S A* 97(12), 6763-6768 (2000)
141. T. Shimohata, T. Nakajima, M. Yamada, C. Uchida, O. Onodera, S. Naruse, T. Kimura, R. Koide, K. Nozaki, Y. Sano, H. Ishiguro, K. Sakoe, T. Ooshima, A. Sato, T. Ikeuchi, M. Oyake, T. Sato, Y. Aoyagi, I. Hozumi, T. Nagatsu, Y. Takiyama, M. Nishizawa, J. Goto, I. Kanazawa, I. Davidson, N. Tanese, H. Takahashi and S.Tsuji: Expanded polyglutamine stretches interact with TAFII130, interfering with CREB-dependent transcription. *Nat Genet* 26(1), 29-36 (2000)
142. A. W. Dunah, H. Jeong, A. Griffin, Y. M. Kim, D. G. Standaert, S. M. Hersch, M. M. Mouradian, A. B. Young, N. Tanese and D. Krainc: Sp1 and TAFII130 transcriptional activity disrupted in early Huntington's disease. *Science* 296(5576), 2238-2243 (2002)
143. D. Helmlinger, S. Hardy, G. Abou-Sleymane, A. Eberlin, A. B. Bowman, A. Gansmuller, S. Picaud, H. Y. Zoghbi, Y. Trottier, L. Tora and D. Devys: Glutamine-expanded ataxin-7 alters TFTC/STAGA recruitment and chromatin structure leading to photoreceptor dysfunction. *PLoS Biol* 4(3), e67 (2006)
144. J.Bilen, N.Liu, B.G.Burnett, R.N.Pittman, and N.M.Bonini: MicroRNA pathways modulate polyglutamine-induced neurodegeneration. *Mol Cell* 24(1), 157-163 (2006)
145. A. Albrecht and S. Mundlos: The other trinucleotide repeat: polyalanine expansion disorders. *Curr Opin Genet Dev* 15(3), 285-293 (2005)
146. K. Han and J. L. Manley: Functional domains of the Drosophila Engrailed protein. *EMBO J* 12(7), 2723-2733 (1993)
147. H. Lavoie, F. Debeane, Q. D. Trinh, J. F. Turcotte, L. P. Corbeil-Girard, M. J. Dicaire, A. Saint-Denis, M. Page, G. A. Rouleau and B. Brais: Polymorphism, shared functions and convergent evolution of genes with sequences coding for polyalanine domains. *Hum Mol Genet* 12(22), 2967-2979 (2003)
148. J. Amiel, D. Trochet, M. Clement-Ziza, A. Munnich and S. Lyonnet: Polyalanine expansions in human. *Hum Mol Genet* 13 Spec No 2, R235-243 (2004)
149. L. Y. Brown and S. A. Brown: Alanine tracts: the expanding story of human illness and trinucleotide repeats. *Trends Genet* 20(1), 51-58 (2004)
150. F. R. Goodman, S. Mundlos, Y. Muragaki, D. Donnai, M. L. Giovannucci-Uzielli, E. Lapi, F. Majewski, J. McGaughran, C. McKeown, W. Reardon, J. Upton, R. M. Winter, B. R. Olsen and P.J. Scambler: Synpolydactyly phenotypes correlate with size of expansions in HOXD13 polyalanine tract. *Proc Natl Acad Sci U S A* 94(14), 7458-7463 (1997)
151. A. N. Akarsu, I. Stoilov, E. Yilmaz, B. S. Sayli and M. Sarfarazi: Genomic structure of HOXD13 gene: a nine polyalanine duplication causes synpolydactyly in two unrelated families. *Hum Mol Genet* 5(7):945-952 (1996)
152. F. R. Goodman: Limb malformations and the human HOX genes. *Am J Med Genet* 112(3), 256-265 (2002)
153. K. R. Johnson, H. O. Sweet, L. R. Donahue, P. Ward-Bailey, R. T. Bronson and M. T. Davisson: A new spontaneous mouse mutation of Hoxd13 with a polyalanine expansion and phenotype similar to human synpolydactyly. *Hum Mol Genet* 7(6), 1033-1038 (1998)
154. A. P. Davis and M. R. Capecchi: A mutational analysis of the 5' HoxD genes: dissection of genetic interactions during limb development in the mouse. *Development* 122(4), 1175-1185 (1996)
155. J. Zakany and D. Duboule: Synpolydactyly in mice with a targeted deficiency in the HoxD complex. *Nature* 384(6604), 69-71 (1996)
156. S. Bruneau, K. R. Johnson, M. Yamamoto, A. Kuroiwa and D. Duboule: The mouse Hoxd13 (spdh) mutation, a polyalanine expansion similar to human type II synpolydactyly (SPD), disrupts the function but not the expression of other Hoxd genes. *Dev Biol* 237(2), 345-353 (2001)
157. F. R. Goodman, C. Bacchelli, A. F. Brady, L. A. Brueton, J. P. Fryns, D. P. Mortlock, J. W. Innis, L. B. Holmes, A. E. Donnenfeld, M. Feingold, F. A. Beemer, R. C. Hennekam and P. J. Scambler: Novel HOXA13 mutations and the phenotypic spectrum of hand-foot-genital syndrome. *Am J Hum Genet* 67(1), 197-202 (2000)
158. L. C. Post, E. H. Margulies, A. Kuo and J. W. Innis: Severe limb defects in Hypodactyly mice result from the

expression of a novel, mutant HOXA13 protein. *Dev Biol* 217(2), 290-300 (2000)
159. B. Brais, J. P. Bouchard, Y. G. Xie, D. L. Rochefort, N. Chretien, F. M. Tome, R. G. Lafreniere, J. M. Rommens, E. Uyama, O. Nohira, S. Blumen, A. D. Korczyn, P. Heutink, J. Mathieu, A. Duranceau, F. Codere, M. Fardeau and G. A. Rouleau: Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat Genet* 18(2), 164-167 (1998)
160. U. Kuhn and E. Wahle: Structure and function of poly(A) binding proteins. *Biochim Biophys Acta* 1678(2-3), 67-84 (2004)
161. C. Messaed, P. A. Dion, A. Abu-Baker, D. Rochefort, J. Laganiere, B. Brais and G. A. Rouleau: Soluble expanded PABPN1 promotes cell death in oculopharyngeal muscular dystrophy. *Neurobiol Dis* 26(3), 546-557 (2007)
162. A. Calado, F. M. Tome, B. Brais, G. A. Rouleau, U. Kuhn, E. Wahle and M. Carmo-Fonseca: Nuclear inclusions in oculopharyngeal muscular dystrophy consist of poly(A) binding protein 2 aggregates which sequester poly(A) RNA. *Hum Mol Genet* 9(15), 2321-2328 (2000)
163. R. Schober, W. Kress, F. Grahmann, S. Kellermann, P. Baum, S. Gunzel and A. Wagner: Unusual triplet expansion associated with neurogenic changes in a family with oculopharyngeal muscular dystrophy. *Neuropathology* 21(1), 45-52 (2001)
164. M. D. Briggs, S. M. Hoffman, L. M. King, A. S. Olsen, H. Mohrenweiser, J. G. Leroy, G. R. Mortier, D. L. Rimoin, R. S. Lachman and E. S. Gaines: Pseudoachondroplasia and multiple epiphyseal dysplasia due to mutations in the cartilage oligomeric matrix protein gene. *Nat Genet* 10(3), 330-336 (2005)
165. E. Delot, L. M. King, M. D. Briggs, W. R. Wilcox and D. H. Cohn: Trinucleotide expansion mutations in the cartilage oligomeric matrix protein (COMP) gene. *Hum Mol Genet* 8(1), 123-128 (1999)
166. C. E. Pearson, K. Nichol Edamura and J. D. Cleary: Repeat instability: mechanisms of dynamic mutations. *Nat Rev Genet* 6(10), 729-742 (2005)
167. R. R. Sinden, V. N. Potaman, E. A. Oussatcheva, C. E. Pearson, Y. L. Lyubchenko and L. S. Shlyakhtenko: Triplet repeat DNA structures and human genetic disease: dynamic mutations from dynamic DNA. *J Biosci* 27(1 Suppl 1), 53-65 (2004)
168. R. D. Wells, R. Dere, M. L. Hebert, M. Napierala and L. S.Son: Advances in mechanisms of genetic instability related to hereditary neurological diseases. *Nucleic Acids Res* 33(12), 3785-3798 (2005)
169. C. J. Cummings and H. Y. Zoghbi: Trinucleotide repeats: mechanisms and pathophysiology. *Annu Rev Genomics Hum Genet* 1, 281-328 (2000)
170. J. P. Jakupciak and R. D. Wells: Genetic instabilities of triplet repeat sequences by recombination. *IUBMB Life* 50(6), 355-359 (2000)
171. R. R. Sinden: Biological implications of the DNA structures associated with disease-causing triplet repeats. *Am J Hum Genet* 64(2), 346-353 (1999)
172. J. D. Cleary and C. E. Pearson: Replication fork dynamics and dynamic mutations: the fork-shift model of repeat instability. *Trends Genet* 21(5), 272-280 (2005)
173. A. J. Hannan: Trinucleotide-repeat expansions and neurodegenerative disease: a mechanism of pathogenesis. *Clin Exp Pharmacol Physiol* 23(12), 1015-1020 (1996)
174. X. Lin and T. Ashizawa: Recent progress in spinocerebellar ataxia type-10 (SCA10). *Cerebellum* 4(1), 37-42 (2005)
175. J. D. Cleary, K. Nichol, Y. H. Wang and C. E. Pearson: Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. *Nat Genet* 31(1), 37-46 (2002)
176. B. S. Balakumaran, C. H. Freudenreich and V. A. Zakian: CGG/CCG repeats exhibit orientation-dependent instability and orientation-independent fragility in Saccharomyces cerevisiae. *Hum Mol Genet* 9(1), 93-100 (2000)
177. G. M. Samadashwily, G. Raca and S. M. Mirkin: Trinucleotide repeats affect DNA replication *in vivo*. *Nat Genet* 17(3), 298-304 (1997)
178. P. J. White, R. H. Borts and M. C. Hirst: Stability of the human fragile X (CGG)(n) triplet repeat array in Saccharomyces cerevisiae deficient in aspects of DNA metabolism. *Mol Cell Biol* 19(8), 5675-5684 (1999)
179. S. J. Gray, J. Gerhardt, W. Doerfler, L. E. Small and E. Fanning: An origin of DNA replication in the promoter region of the human fragile X mental retardation (FMR1) gene. *Mol Cell Biol* 27(2), 426-437 (2007)
180. K. Nichol Edamura, M. R. Leonard and C. E. Pearson: Role of replication and CpG methylation in fragile X syndrome CGG deletions in primate cells. *Am J Hum Genet* 76(2), 302-311 (2005)
181. S. Kang, A. Jaworski, K. Ohshima and R. D. Wells: Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in E. coli. *Nat Genet* 10(2), 213-218 (1995)
182. R. Pelletier, M. M. Krasilnikova, G. M. Samadashwily, R Lahue and S. M. Mirkin: Replication and expansion of trinucleotide repeats in yeast. *Mol Cell Biol* 23(4), 1349-1357 (2003)
183. J. K. Schweitzer and D. M. Livingston: The effect of DNA replication mutations on CAG tract stability in yeast. *Genetics* 152(3), 953-963 (1999)
184. J. K. Schweitzer and D. M. Livingston: Expansions of CAG repeat tracts are frequent in a yeast mutant defective in Okazaki fragment maturation. *Hum Mol Genet* 7(1), 69-74 (1998)
185. T. Nenguke, M. I. Aladjem, J. F. Gusella, N. S. Wexler, N. Arnheim and Venezuela HD Project: Candidate DNA replication initiation regions at human trinucleotide repeat disease loci. *Hum Mol Genet* 12(9), 1021-1028 (2003)
186. M. Bzymek and S. T. Lovett: Instability of repetitive DNA sequences: the role of replication in multiple mechanisms. *Proc Natl Acad Sci U S A* 98(15), 8319-8325 (2001)
187. A. Azaiez, E. F.Bouchard, M. Jean and F. J. Belzile: Length, orientation, and plant host influence the mutation frequency in microsatellites. *Genome* 49(11), 1366-1373 (2006)
188. D. L. Daee, T. Mertz and R. S. Lahue: Postreplication repair inhibits CAG.CTG repeat expansions in Saccharomyces cerevisiae. *Mol Cell Biol* 27(1), 102-110 (2007)

189. L. Kennedy and P. F. Shelbourne: Dramatic mutation instability in HD mouse striatum: does polyglutamine load contribute to cell-specific vulnerability in Huntington's disease? *Hum Mol Genet* 9(17), 2539-2544 (2000)
190. C. Jankowski, F. Nasar and D. K. Nag: Meiotic instability of CAG repeat tracts occurs by double-strand break repair in yeast. *Proc Natl Acad Sci U S A* 97(5), 2134-2139 (2000)
191. J. P. Jakupciak and R. D. Wells: Genetic instabilities in (CTG.CAG) repeats occur by recombination. *J Biol Chem* 274(33), 23468-23479 (1999)
192. J. P. Jakupciak and R. D. Wells: Gene conversion (recombination) mediates expansions of CTG.CAG repeats. *J Biol Chem* 275(51), 40003-40013 (2000)
193. Z. H. Zhou, E. Akgun and M. Jasin: Repeat expansion by homologous recombination in the mouse germ line at palindromic sequences. *Proc Natl Acad Sci U S A* 98(15), 8326-8333 (2001)
194. U. Mittal, A. K. Srivastava, S. Jain, S. Jain and M. Mukerji: Founder haplotype for Machado-Joseph disease in the Indian population: novel insights from history and polymorphism studies. *Arch Neurol* 62(4), 637-640 (2005)
195. S. T. Warren: Polyalanine expansion in synpolydactyly might result from unequal crossing-over of HOXD13. *Science* 275(5298), 408-409 (1997)
196. C. T. McMurray: DNA secondary structure: a common and causative factor for expansion in human disease. *Proc Natl Acad Sci U S A* 96(5), 1823-1825 (1999)
197. S. M. Mirkin: DNA structures, repeat expansions and human hereditary disorders. *Curr Opin Struct Biol* 16(3), 351-358 (2006)
198. P. L. James, T. Brown and K. R. Fox: Thermodynamic and kinetic stability of intermolecular triple helices containing different proportions of C+*GC and T*AT triplets. *Nucleic Acids Res* 31(19), 5598-5606 (2003)
199. J. Petruska, N. Arnheim and M. F. Goodman: Stability of intrastrand hairpin structures formed by the CAG/CTG class of DNA triplet repeats associated with neurological diseases. *Nucleic Acids Res* 24(11), 1992-1998 (1996)
200. M. M. Alba and R. Guigo: Comparative analysis of amino acid repeats in rodents and humans. *Genome Res* 14(4), 549-554 (2004)
201. L. Mularoni, R. A. Veitia and M. M. Alba: Highly constrained proteins contain an unexpectedly large number of amino acid tandem repeats. *Genomics* 89(3), 316-325 (2007)
202. J. D. Cleary and C. E. Pearson: The contribution of cis-elements to disease-associated repeat instability: clinical and experimental evidence. *Cytogenet Genome Res* 100(1-4), 25-55 (2003)
203. S. Choudhry, M. Mukerji, A. K. Srivastava, S. Jain and S. K. Brahmachari: CAG repeat instability at SCA2 locus: anchoring CAA interruptions and linked single nucleotide polymorphisms. *Hum Mol Genet* 10(21), 2437-2346 (2001)
204. A. Hoekema A, R. A. Kastelein, M. Vasser and H. A. de Boer: Codon replacement in the PGK1 gene of Saccharomyces cerevisiae: experimental approach to study the role of biased codon usage in gene expression. *Mol Cell Biol* 7(8), 2914-2924 (1987)
205. G. Caponigro, D. Muhlrad and R. Parker: A small segment of the MAT alpha 1 transcript promotes mRNA decay in Saccharomyces cerevisiae: a stimulatory role for rare codons. Mol Cell Biol 13(9), 5141-5148 (1993)
206. D. B. Carlini: Context-dependent codon bias and messenger RNA longevity in the yeast transcriptome. *Mol Biol Evol* 22(6), 1403-1411 (2005)
207. T. Ikeuchi, S. Igarashi, Y. Takiyama, O. Onodera, M. Oyake, H. Takano, R. Koide, H. Tanaka and S. Tsuji: Non-Mendelian transmission in dentatorubral-pallidoluysian atrophy and Machado-Joseph disease: the mutant allele is preferentially transmitted in male meiosis. *Am J Hum Genet* 58(4), 730-733 (1996)
208. Y. Trottier, V. Biancalana and J. L. Mandel: Instability of CAG repeats in Huntington's disease: relation to parental transmission and age of onset. *J Med Genet* 31(5), 377-382 (1994)
209. M. A. Pujana, J. Corral, M. Gratacos, O. Combarros, J. Berciano, D. Genis, I. Banchs, X. Estivill and V. Volpini: Spinocerebellar ataxias in Spanish patients: genetic analysis of familial and sporadic cases. The Ataxia Study Group.*Hum Genet* 104(6), 516-522 (1999)
210. L. Pianese, F. Cavalcanti, G. De Michele, A. Filla, G. Campanella, O. Calabrese, I. Castaldo, A. Monticelli and S. Cocozza: The effect of parental gender on the GAA dynamic mutation in the FRDA gene. *Am J Hum Genet* 60(2), 460-463 (1997)
211. I. V. Kovtun, T. M. Therneau and C. T. McMurray: Gender of the embryo contributes to CAG instability in transgenic mice containing a Huntington's disease gene. *Hum Mol Genet* 9(18), 2767-2775 (2000)
212. I. V. Kovtun, G. Welch, H. D. Guthrie, K. L. Hafner and C. T. McMurray: CAG repeat lengths in X- and Y-bearing sperm indicate that gender bias during transmission of Huntington's disease gene is determined in the embryo. *J Biol Chem* 279(10), 9389-9391 (2004)
213. A. Eriksson and B. Mehlig: On the effect of fluctuating recombination rates on the decorrelation of gene histories in the human genome. *Genetics* 169(2), 1175-1178 (2005)
214. A. Kong, D. F. Gudbjartsson, J. Sainz, G. M. Jonsdottir, S. A. Gudjonsson, B. Richardsson, S. Sigurdardottir, J. Barnard, B. Hallbeck, G. Masson, A. Shlien, S. T. Palsson, M. L. Frigge, T. E. Thorgeirsson, J. R Gulcher and K. Stefansson: A high-resolution recombination map of the human genome. *Nat Genet* 31(3), 241-247 (2002)
215. D. P. Depledge, R. P. Lower and D. F. Smith: RepSeq--a database of amino acid repeats present in lower eukaryotic pathogens. *BMC Bioinformatics* 8, 122 (2007)
216. D. A. Parry: Structural and functional implications of sequence repeats in fibrous proteins. *Adv Protein Chem* 70, 11-35 (2005)
217. J. D. Wren, E. Forgacs, J. W. 3rd Fondon, A. Pertsemlidis, S. Y. Cheng, T. Gallardo, R. S. Williams, R. V. Shohet, J. D. Minna and H. R. Garner: Repeat polymorphisms within gene regions: phenotypic and evolutionary implications. *Am J Hum Gene*t 67(2), 345-356 (2000)
218. P. I. Missirlis, C. L. Mead, S. L. Butland, B. F. Ouellette, R. S. Devon, B. R. Leavitt and R. A. Holt: Satellog: a database for the identification and prioritization of satellite repeats in disease association studies. *BMC Bioinformatics* 6, 145 (2005)

219. V. Guryev, E. Berezikov and E. Cuppen: CASCAD: a database of annotated candidate single nucleotide polymorphisms associated with expressed sequences. *BMC Genomics* 6(1), 10 (2005)
220. T. Boby, A. M. Patch and S. J.Aves: TRbase: a database relating tandem repeats to disease genes for the human genome. *Bioinformatics* 21(6), 811-816 (2005)
221. L. Mularoni, R. Guigo and M. M. Alba: Mutation patterns of amino acid tandem repeats in the human proteome. *Genome Biol* 7(4), R33 (2006)
222. Y. Kashi, D. King and M. Soller: Simple sequence repeats as a source of quantitative genetic variation. *Trends Genet* 13(2), 74-78(1997)
223. Y. Kashi and D. G King: Simple sequence repeats as advantageous mutators in evolution. *Trends Genet* 22(5), 253-259 (2006)
224. J. W. Fondon 3rd H. R. Garner HR: Molecular origins of rapid and continuous morphological evolution. P*roc Natl Acad Sci U S A* 101(52), 18058-18063 (2004)
225. E. A. Hammock and L. J. Young: Microsatellite instability generates diversity in brain and sociobehavioral traits. *Science* 308(5728), 1630-1634 (2005)

Note: Further details on the gene, mutations, and disease can be obtained from the OMIM link (http://www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM), using the MIM number provided

Abbreviations: AD: autosomal dominant, AR: autosomal recessive, XR: X-linked recessive, NR: not reported, De novo: De novo mutation in a germ cell, M: maternal, P: paternal, LOF: loss-of-function, GOF: gain-of-function

Key Words: Tandem repeat sequences, repeat dynamism, PolyQ disorders, PolyA Disorder, Ubiquitin-Proteaosome System, neuronal dysfunction, cell death, Review

Send correspondence to: Dr Subramaniam Ganesh, Department of Biological Sciences and Bioengineering, Indian Institute of Technology, Kanpur 208016, India, Tel: 91-512-259-4040, Fax: 91-512-259-4010, E-mail: sganesh@iitk.ac.in

http://www.bioscience.org/current/vol13.htm