Slipping while sleeping? Trinucleotide repeat expansions in germ cells

Christopher E. Pearson

Program of Genetics and Genomic Biology, The Hospital for Sick Children, 555 University Avenue, Elm Wing 11-135, Toronto, Ontario, Canada M5G 1X8

Trinucleotide expansions cause at least 30 diseases including Huntington’s disease (HD). Many are inherited predominantly through paternal transmissions, which are probably the result of germ-cell-specific mutations. A recent study of testicular germ cells in HD patients revealed that expansions occur in diploid cells before the completion of meiosis. Therefore, expansions are not limited to the late-haploid spermatids, in which the genome is ‘sleeping’. These results have implications both for research aimed at understanding the transmission of this serious mutation and for developing new therapies for the disease.

Expansion of gene-specific CAG-, CTG- and CGG-trinucleotide repeats cause multiple diseases, including Huntington’s disease (HD), myotonic dystrophy type 1 (DM1) and fragile X mental retardation (FRAXA). Many are inherited in a non-mendelian manner known as genetic anticipation, in which disease symptoms are more severe and/or are evident at an earlier age with successive generations. Genetic anticipation is thought to arise through dynamic mutations displayed by the repeats, in which larger repeat tracts are more susceptible to subsequent mutations than progenitor alleles. Each disease displays different levels of repeat instability in somatic tissues and during parent–offspring transmissions (instability in patients is reviewed in Ref. [1]).

Whereas expansions and contractions both during post-fertilization cell divisions and throughout life can contribute to somatic instability (Figure 1a), transmitted instability must involve germline mutations (Figure 1b). Many of these diseases are inherited predominantly through either paternal or maternal transmissions; however, the mechanism of this parent-of-origin effect for any one of the diseases has, until recently, remained elusive [2]. Huntington’s disease shows a paternal bias for transmission of repeat expansions. By contrast, in FRAXA, disease-associated expansions occur only through maternal transmissions. Myotonic dystrophy – like HD – shows a paternal bias for expansions of progenitor CTG/CAG lengths of 100 or less (in the range of HD alleles). For longer alleles of 200–600 repeats (which do not occur in HD), DM1 shows both a paternal and maternal bias for expansions; tract lengths of >600 are maternally transmitted to exceedingly large lengths of 1000–4000 repeats (reviewed in Ref. [1]). Each parent-of-origin mutation bias is driven by complex processes specific to sperm or oocyte development. Expansion mediated by DNA slippage between repeat strands [3] could occur in proliferating, arrested meiotic cells or in the relatively dormant haploid germ cells (Figure 2). Knowing when expansion events occur during spermatogenesis in HD patients is crucial to understanding this important mutation and its transmission in affected families. A recent study reported that expansion events occurred in pre-meiotic germ cells of male HD patients [2].

Cellular and DNA metabolism through human gametogenesis

Knowing when expansions occur during germ-cell development will suggest which DNA metabolic process mediates instability. Mitotic and DNA metabolic activities change throughout the stages of gametogenesis (Figure 1b). Many replication, repair or recombination genes show altered or specialized expression during gametogenesis [4–9]. If expansions occur during replication, the larger number of mitotic divisions in male gametogenesis could explain the paternal bias for HD mutation. Following spermatogonial proliferation and meiosis, the genome of the haploid spermatids is re-packaged with protamine, thereby entering a relatively inactive or ‘sleeping’ state, and eventually becoming spermatozoa. If the mismatch repair proteins MSH2 and MSH3 are required for CAGCTG expansions in humans (and possibly protection from deletions), as has been reported in transgenic mice [10–13], then expansion events must occur in the germ-cell stages that contain MSH2 and MSH3. In humans and mice, MSH2 levels are high in mitotically active spermatogonia and decrease from meiosis onwards to immunohistochemically undetectable levels in round and elongating spermatids [5,7]. Interestingly, the level of MSH3 expression was low in spermatogonia but increased during meiosis – peaking in pachytene spermatocytes – then decreased in later stages [7]. Together, these results suggest that expansion mutations occur in replicating spermatogonia.

CAG expansions in germ cells of HD patients

Analysis of testicular cells of human HD patients with (CAG)50 (presumed the progenitor allele) in the blood revealed that high frequencies of repeat expansions of considerable magnitude were present in mitotic diploid...
Spermatozoa

Ovum

blocked at metaphase II

Somatic instability (mitotic and non-mitotic tissues) via replication, genome maintenance and repair

Male embryo

16 divisions

~ Day 5 - 12, germ cell segregation primordial germ cells

17 divisions

In utero

# of germ cells (~ 99 million)

Replication errors, repair associated with DNA replication and genome maintenance

Female embryo

16 divisions

~ Day 5 - 12, germ cell segregation primordial germ cells

13 divisions

Max. # of germ cells (~ 6.8 million)

Primary oocytes (2 x 2C) arrested at dictyotene stage of prophase

Birth

Puberty

Male embryo

1 division

~ 74 days

Primary spermatocyte (2 x 2C)

Secondary spermatocyte (2 x 1C)

Spermatogonia (2C)

Meiosis I

Meiosis II

Secondary spermatocyte (2C + pb) with polar body

Beginning of meiosis II

Beginning of meiosis I

Release from dictyotene arrest

Ovulation

Replication errors, repair associated with DNA replication and genome maintenance

Testis

Brain

Kidney

Blood

Skeletal muscle

Somatic instability (mitotic and non-mitotic tissues) via replication, genome maintenance and repair

Female embryo

1 division

~ Day 5 - 12, germ cell segregation primordial germ cells

13 divisions

Max. # of germ cells (~ 6.8 million)

Primary oocytes (2 x 2C) arrested at dictyotene stage of prophase

Birth

Puberty

Primary spermatocyte (2 x 2C)

Secondary spermatocyte (2 x 1C)

Spermatogonia (2C)

Meiosis I

Meiosis II

Secondary spermatocyte (2C + pb) with polar body

Beginning of meiosis II

Release from dictyotene arrest

Ovulation

Replication errors, repair associated with DNA replication and genome maintenance

Postmeiotic repair

Spermatozoa

Ovum blocked at metaphase II

2 pronuclei (2 x 1C)

Completion of maternal meiosis II

Envelope breakdown

Division

2 cell embryo

Gonemeric replication

Histone

Protamine

Protamine to histone

2 pronuclei (2 x 1C)

Completion of maternal meiosis II

Embryo (4C)
germ cells before meiosis [2]. Cells at defined stages of spermatogenesis were isolated by laser capture microdissection, which circumvents many of the difficulties of other isolation procedures [14]. Repeat lengths were assessed by the highly sensitive single-molecule PCR.

This instability could have occurred in at least two phases; expansion products were evident at similar frequencies in both pre-meiotic and post-meiotic cells. However, the proportion of large expansions in post-meiotic spermatids or spermatoozoa was greater, that expansion events occurred before and possibly after (or during) meiosis. The frequency of expanded alleles in either pre-meiotic or meiotic spermatogonia was higher for the HD mouse than for mice, or in a mixture of pre-meiotic and meiotic cells was 87% and 88%, respectively. Post-meiotic spermatids and spermatoozoa showed a similar expansion frequency of 75%. (The slight reduction might indicate a failure of some pre-meiotic or meiotic expansions to progress.) The progenitor (CAG)50 showed a range of mutant length changes: pre-meiotic cells showed expansion products of up to 125 repeats and tracts as short as 43 repeats, with a median of 77 repeats. Post-meiotic cells showed a relatively continuous range of expansions products of up to 127 repeats and tracts as short as 43 repeats, with a median of 87 repeats. Thus, in HD male germ cells, expansions occur before and possibly after meiosis.

Data from HD patients revealed that extensive expansions of a (CAG)50 occurred in diploid germ cells before meiosis. Similar to the results from HD patients, a DM1 transgenic mouse showed expansions arising in diploid spermatogonia [11]; however, these did not occur in post-meiotic spermatids or spermatoozoa (G. Gourdon, pers. commun.). Results from another DM1 mouse also showed that expansions occur in pre-meiotic proliferating spermatogonia [15]. Together with data from humans, these results contrast with conclusions reported for a transgenic HD mouse, which suggest that expansions occur only in late-haploid stages of spermatogenesis, specifically in elongating spermatids when the genome is becoming dormant [10,16]. In this case, a mouse with a (CAG)117 tract showed modest increases of 1–8 repeats. The discrepancy with results from the HD mouse can be explained in several ways [2,11,15]. In addition to DNA-sequence context differences between HD patients and mice, cellular and DNA metabolism during spermatogenesis differ between the two species [4,6]. For example, in humans, the greater number of mitotic divisions and the extended time before and during spermatogenesis increases the potential for mutations compared with mice. Inter-species differences in repair gene expression in germ cells might also contribute [4,5,7]. Moreover, humans could be exposed to very different mutagens than mice, many of which might affect either DNA or repair systems in gametes [17]. These inter-species differences might also explain the inability of any mouse model to recapitulate the large expansion increments observed in some human transmissions.

There are clear differences in transmitted repeat instability between the HD mice [10,16] and HD patients [2]; similar differences exist between the HD mice and other transgenic mice [11,15]. The differences between transgenic mice could reveal important cis- and trans-acting factors. Both within and among the various HD, DM1 and other transgenic mice, there are differences in transmitted parental mutation biases (if at all evident) [10–13,15]. Such differences might reflect context-specific (chromosomal integration site) differences [15], similar to the different human parental mutation biases between diseases such as HD and spinocerebellar ataxia (SCA8; both are associated with CAG repeats [1]). In addition, the length of the integrated repeat might determine the parent mutation bias pattern in transgenic mice, similar to the variable parent-of-origin effects displayed by different CTG lengths in human DM1 families (see earlier) [1]. Mouse-strain-specific differences might also account for variable murine mutation patterns: unsuspected mutations in repair genes within the supposed wild-type embryonic stem cells, such as those occurring in the polymerase iota (pol i) gene in the 129-derived strains [18], might unknowingly contribute to repeat instability. Importantly, high levels of the error-prone DNA polymerases \( \iota, \kappa, \lambda, \eta \) and \( \beta \), as well as pol \( \epsilon \), are present in mouse testes at various stages of spermatogenesis, before, during and after meiosis, and, in some cases, in round spermatids [8,9 and Refs therein]. Outside of mitotic spermatogonia, the replicative polymerases \( \alpha \) and \( \delta \) are present at higher levels coincident with meiotic DNA synthesis [8] and Refs therein. Strain-specific defects in polymerases or other repair genes might influence the interpretation of transmitted or somatic instabilities observed in transgenic mice.

Figure 1. Somatic and germline DNA metabolic processes in humans. (a) Repeat instability in patients suffering from a disease associated with a trinucleotide repeat can display both pre-natal and post-natal tissue-specific repeat instabilities. Tract-length heterogeneity is evident in both proliferating and terminally differentiated patient tissues (reviewed in Ref. [1]). Huntington’s disease patients show CAG-length mosaicism in sperm and brain tissues [26]. (b) Both male and female germ-cell segregation is thought to occur between day 5–12 post-fertilization, although the timing has not been rigorously established. Male germ-cell production involves – 30 in utero mitotic divisions as well as post-puberal spermatogonial cell divisions throughout life (1 every 16 days over a spermatogenetic cycle of 74 days). Assuming puberty begins at 14 years of age, the number of accumulated divisions (mitotic only) through spermatogenesis for a 30-year-old male is approximately 400 divisions (≈ 30 + (30 × 14 – 174 – 368) ÷ 16 divisions per year)). There are various possible sources of error in the estimates of the number of replicative cell divisions: estimations of the number of stem cells and the number of divisions to produce them made no allowance for the extensive cell degeneration (apoptosis) known to occur. Therefore, the true number of replicative divisions might be higher. Female germ-cell production, which is complete before birth (5 months gestation), involves only 24–33 mitotic cell divisions. Oogenic meiosis begins in utero and is arrested in dictyotene for years, resuming only minutes before ovulation, after which meiosis I is complete. Meiosis II then begins and is not completed until fertilization. In contrast to post-meiotic male germ cells, arrested and/or resting oocytes are primed with DNA repair and recombination activities [4,26], which can lead to age-related CAG instability in SCA1 transgenic mice [33]. (c) Fertilization and the steps leading to the two-cell embryo, which includes the formation of two pronuclei, the completion of maternal meiosis II, decondensation of the paternal genome, DNA repair, gonocemic DNA replication within two haploid pronuclei, the breakdown of pronuclear envelopes, syncyngium and cleavage [34]. Gonomic replication is the only haploid DNA replication in the diploid metazoan life cycle. DNA metabolic processes occurring during each developmental stage are indicated by graded shading. Abbreviations: HD, Huntington’s disease; DM1, myotonic dystrophy type 1; DRPLA, dentatorubral-pallidoluysian atrophy; SCA, spinocerebellar ataxia; FRAXA, fragile X mental retardation; p, ploidy; pb, polar body; gm-repair, genome maintenance repair; rec-repair, recombination-associated repair; rep-, replication errors and replication-associated repair. Information compiled from Refs [4,6,17 and 34] and references therein. These processes can vary greatly between different species.
Mechanistic implications

Results from HD patients suggest various DNA metabolic processes that could mediate instability. Huntington CAG expansion events must occur at some point between segregation of the primordial germ cells, development to puberty, or during the life-long post-pubertal spermogonial stem-cell divisions (Figure 2a). A second phase of expansion might occur after meiosis and possibly through to elongating spermatids. Primordial germ cells and spermatogonia undergo genome replication and are subject to DNA replication errors, repair associated with replication, and genome maintenance repair (Figures 1b and 2a). The mitotically inactive post-natal pre-pubertal stages are subject to genome maintenance repair. Later stages undergo meiotic recombination, genome maintenance repair and, ultimately, chromatin decondensation (coincident with dormancy) (Figure 1b). The absence of considerable instability in non-germ somatic tissues of HD patients suggests various DNA metabolic processes that could mediate instability. Huntington CAG expansion events must occur at some point between segregation of the primordial germ cells, development to puberty, or during the life-long post-pubertal spermogonial stem-cell divisions (Figure 2a). A second phase of expansion might occur after meiosis and possibly through to elongating spermatids. Primordial germ cells and spermatogonia undergo genome replication and are subject to DNA replication errors, repair associated with replication, and genome maintenance repair (Figures 1b and 2a). The mitotically inactive post-natal pre-pubertal stages are subject to genome maintenance repair. Later stages undergo meiotic recombination, genome maintenance repair and, ultimately, chromatin decondensation (coincident with dormancy) (Figure 1b). The absence of considerable instability in non-germ somatic tissues of
HD patients argues against the occurrence of expansion events in early embryos and favors their occurrence during late-fetal, post-natal or post-pubertal stages.

If HD CAG expansion events occur only during the post-pubertal cycling of spermatogonia, sperm produced at advanced ages would have larger expansions than sperm produced at the onset of puberty. A limitation of CAG expansions to cycling spermatogonia would be expected to show a paternal age effect on the progress of CAG expansions in HD families, as has been observed in several transgenic mice [11,15,19]. Analysis of pre-meiotic male germ cells in a DM1 mouse, over the course of seven weeks to 11 months, revealed continuous expansions, suggesting that expansions occur in cycling spermatogonia (G. Gourdon, pers. commun.). An effect of paternal age on transmitted CAG instability has only been reported for dentaturation-pallidalolysium (DRPLA) families [19]. In HD, such an association has only been suspected based upon the inverse relationship between age of disease onset and paternal age (this study was performed before the discovery that HD was caused by an unstable CAG tract [20]). However, subsequent analyses failed to reveal such an association at the molecular level [21,22]. The inability to detect a paternal age effect in HD-affected fathers with CAG instability (through analysis of sperm CAG instability or parent-to-child mutations) is likely to be the result of: (i) a paucity of transmissions by older HD parents, (ii) the paucity of sperm samples from the same individual taken over long enough periods of time, or (iii) other complex variables. Thus, although clinical data [20–22] indicate a probable paternal age effect on HD CAG instability, it is difficult to prove its association with mutations occurring during spermatogonial cycling.

Rather than narrowing the potential processes that might mediate CAG expansions, the results from HD patients [2] have broadened the spectrum relative to some previous murine studies [10,16]. As expansion products are evident in human pre-meiotic diploid cells, the initial mutation event must involve DNA replication-fork errors, replication-associated repair or genome maintenance repair. Further expansions might involve recombination or post-meiotic genome maintenance. As shown in Figure 2b, CAG expansions can arise through various processes. Instability via replication can be complex and could vary between different developmental stages and/or tissues [1,23]. There is evidence to suggest that instability involves DNA repair proteins, which might be associated with replication fork errors, genome maintenance, or recombination [24]. Many aspects of genome maintenance repair are active in mitotic, inter-mitotic and non-proliferating cells. All models of expansion involve slipped-strand DNA structures that are formed by repeat tracts as the putative mutagenic intermediates of instability (reviewed in Ref. [3]). Deciphering the formation and processing of these mutagenic intermediates will be crucial to understanding instability (Figure 2b).

Caution should be taken when extrapolating or generalizing the results from HD patient to other HD patient tissues (such as the brain [25]) or to other trinucleotide-associated diseases. Notably, tissue-specific somatic instability in HD and other diseases (Figure 1a) might occur by processes that are unrelated to germ cells [23,26,27]. In DM1 and SCA8, which, like HD, are caused by an unstable CTG or CAG repeat, families show a maternal bias rather than a paternal bias for large expansions. Furthermore, in FRAXA, maternal oocyte-specific CCG expansions probably occur soon after germ-cell segregation, whereas, in males, the inherited expansion undergoes germ-cell-specific CCG contractions and selective growth, resulting in spermatooza with only pre-mutation tract lengths [28]. A similar process leading to deletions in sperm might occur in SCA8 patients [29]. Thus, it is probable that HD instability mechanisms are similar in some respects to certain disease loci, but vastly different to others.

Future studies

Although we are far from curing diseases caused by trinucleotide expansions, understanding the mechanism(s) tissue specificity, and developmental timing of repeat expansions and contractions is crucial to achieving such goals. The following sections cover only those avenues directed towards DNA.

Transmission

If HD CAG expansion events occur only during the post-pubertal cycling of spermatogonia, sperm samples isolated at the onset of puberty would have smaller expansions than samples isolated later in life. If this is true, early pubertal semen could be stored for the eventual purposes of in vitro fertilization, which would lead to transmissions with minimal expansions or no expansions for those fertilizations derived from sperm harboring the normal HD allele. Such measures could limit the contribution of germline CAG instability to genetic anticipation.

Somatic

In HD, like most trinucleotide repeat diseases, the mutant gene product (protein and/or RNA) exerts a clinically dominant effect mediated by a gain-of-function. Thus, at the somatic level, a cure or treatment for the individual might require replacing the mutant gene with a wild-type gene. One way of ‘replacing’ the mutant gene with a normal gene would be to modulate specifically the instability of the expanded allele in cells to harbor only contractions to shorter non-diseased repeat lengths. Treatment of cells either in vitro or within the patient with a compound that might modulate instability towards contractions could effectively transmute a diseased allele to become a normal allele. An important criterion of such a compound is that it should affect instability only at the expanded diseased locus and not have deleterious effects upon the rest of the genome. Ablating a repair gene or system could have dire consequences, including genome-wide instability and induction of tumorigenesis [13,30]. Targeting the activity of a particular gene product involved in DNA metabolism to both a specific tissue and genic locus (the expanded disease locus) should be a goal of future studies [31]. Another approach would be to isolate and/or selectively grow somatic (stem) cells that have repeat contractions, and these could be transplanted into patients. Incidentally, the process of repeat contraction and selective growth advantage appears to occur naturally in the
germ cells of FRAXA [28] and SCA8 males [29], and in the cerebellar cortex of HD, DM1, SCA1, SCA2, SCA3 and DRPLA patients (for review, see Ref. [1]).

**Concluding remarks**

In the case of Huntington CAG expansion mutation, Arnheim and colleagues have revealed both the window of human sperm development and the breadth of the DNA metabolic processes, which contribute to the CAG dynamic mutations that drive genetic anticipation. Huntington CAG expansion events in male germ cells occur before meiosis at a point following segregation of the primordial germ cells, through to the life-long post-pubertal spermatogonial stem-cell divisions. Future research should focus on delimiting both the specific timing of instability and the DNA metabolic process(es), as well as identifying the development- and tissue-specific factors and mechanisms that are involved in disease-associated instability. Hopefully, such studies will keep pace with the advances of modulating repeat instability and stem-cell transplantation.

**Note added in proof**

Incidentally, somatic CAG instability in the brains of HD patients might not occur by the assumed process of genome maintenance repair [27], as recent evidence has revealed increased cell proliferation and neurogenesis in HD brains [32]. Also, spontaneous locus-specific CTG expansions in cultured DM1 patient fibroblasts occurred only during proliferation, and these were specifically enhanced by chemicals that perturb replication-fork progression [31].

**Acknowledgements**

I thank John D. Cleary and many of my colleagues for preparation of Figures 1 and 2. Research in the Pearson laboratory is supported by grants from the Canadian Institutes of Health Research (CIHR), the Muscular Dystrophy Association USA, and the Fragile X Research Foundation Canada. C.E.P. is a CIHR Scholar and a Canadian Genetic Disease Network Scholar.

**References**

18. McDonald, J.P. et al. (2003) 129-derived strains of mice are deficient in DNA polymerase (i) and have normal immunoglobulin hypermutation. J. Exp. Med. 196, 843–848

http://tmm.trends.com