Somatic deletion events occur during early embryonic development and modify the extent of CAG expansion in subsequent generations

I.V. Kovtun¹, A.R. Thornhill² and C.T. McMurray¹,³,⁴,*

¹Department of Molecular Pharmacology and Experimental Therapeutics, ²Department of Obstetrics and Gynecology, ³Department of Biochemistry and Molecular Biology and ⁴Molecular Neuroscience Program, Mayo Clinic and Foundation, Rochester, MN 55905, USA

Received July 2, 2004; Revised September 29, 2004; Accepted October 11, 2004

Alterations in trinucleotide repeat length during transmission are important in the pathophysiology of Huntington’s disease (HD). However, it is not well understood where, when and by what mechanism expansion occurs. We have followed the fate of CAG repeats during development in mice that can [hHD(−/+)/Msh2(+/−)] or cannot [hHD(+/−)/Msh2(−/−)] expand their repeats. Here we show that long repeats are shortened during somatic replication early in the embryo of the progeny. Our data point to different mechanisms for expansion and deletion. Deletions arise during replication, do not depend on the presence of Msh2 and are largely restricted to early development. In contrast, expansions depend on strand break repair, require the presence of Msh2 and occur later in development. Overall, these results suggest that deletions in early development serve as a safeguard of the genome and protect against expansion of the disease-range repeats during transmission.

INTRODUCTION

At least eight hereditary and progressive neurodegenerative disorders have been identified, in which the underlying mutation is a CAG expansion within the coding sequence of the gene (1–3). Once the CAG length exceeds a disease threshold, expansion becomes the predominant change and exceeds contraction events by 3–175-fold (4–7). Increase in repeat length from generation to generation is a critical feature in pathophysiology, yet little is known about how and where the mutation occurs in vivo. Human data and experiments in mouse models for disease have revealed that instability is present in the germ cells and may account for the growth of the triplet repeat tract in subsequent generations.

Developmental aspects of expansion in coding regions are not as well characterized. Expansions in the sperm are found in HD patients (10,11) and hHD transgenic mice (14). However, alterations in size of the CAG tracts found in coding sequences are small relative to either fragile X or myotonic dystrophy, and they are most often transmitted through the male line (4,20,21). Single cell analysis of testes from HD patients has revealed that mutations occur in all testicular cell types including those that proliferate and those that differentiate into mature sperm (22). In hHD mice, expansion is present in haploid germ cells that are post-mitotic and post-meiotic and limit timing of expansion to late testicular development and differentiation (14). While the mechanism of expansion is not resolved from these studies, data from humans and mice agree that instability is present in the germ cells and may account for the growth of the triplet repeat tract in subsequent generations.

Less is known about events post-fertilization. However, several pieces of evidence indicate that there is somatic instability. For example, the degree of expansion observed in HD pedigrees does not correspond to that in the sperm of

*To whom correspondence should be addressed. Tel: +1 5072841597; Fax: +1 5072849111; Email: mcmurray.cynthia@mayo.edu

Human Molecular Genetics, Vol. 13, No. 24 © Oxford University Press 2004; all rights reserved
transmitting parent (10). Studies in hHD transgenic mice have demonstrated that CAG repeats deleted in female and expanded in male progeny from identical fathers, indicating that alterations of repeat length can arise from events in the early embryo (23,24).

The developmental aspects of expansion are further complicated by the as-yet unknown role of Msh2 (25–29). In all mouse models for expansion, loss of a mismatch repair complex involving Msh2 attenuates expansion (14,30,31). Thus, Msh2, normally present in cells, is a causative factor. This is puzzling since a stabilizing role of Msh2 on microsatellites is well described. Small expansions can be observed at microsatellite loci in MMR defective background (32,33). However, loss of Msh2 in yeast (34), bacteria (35–37), fly (33), mice (31,38,39) and humans (40–43) most often results in deletions. The direction of change may depend on initial length of the repetitive sequence. Additionally, the effects of Msh2 during development may be tissue specific.

To better understand how repeats are altered during transmission and development, we have followed intergenerational alterations in hHD mice which can \( hHD^{(+/-)/Msh2^{(+/+)}} \) or cannot \( hHD^{(-/-)/Msh2^{(-/-)}} \) expand their repeats. We specifically focused on answering three questions: (1) where and (2) when CAG expansion occurs during mammalian transmission and development and (3) how Msh2 contributes to the mutational process. We report here that \( hHD^{(-/-)/Msh2^{(-/-)}} \) animals and their progeny not only lost the ability to expand, but they consistently deleted their repeats during parent–offspring transmission. During intergenerational transmission, deletion events occurred at a different time in development and by a different mechanism than expansion. Both expansion and deletions depended on the length of the CAG repeats, yet deletions arose during replication, were independent of Msh2 and were largely restricted to early embryo development. These features contrasted with those of expansion which depended on strand break repair, required Msh2 and occurred later in development. Deletions and expansions, together, determined the net intergenerational change in repeat size upon transmission.

RESULTS

In mice harboring exon 1 of the hHD gene, the CAG repeats are known to expand during parent–offspring transmission (23,44). Loss of Msh2 in these animals abrogates expansion in both somatic and germ cells suggesting that Msh2 serves as a causative factor (14,30,31). To better understand the role of Msh2 in causing expansion, we followed the progeny of hHD mice with and without their complement of Msh2. We bred hHD males with heterozygous Msh2\((-/-)\) female mice to produce two distinct lines from the same founder (an example is shown in Fig. 1A).

The size of the repeats that are transmitted to the progeny is expected to arise from alterations in the sperm of the father. Expansions of CAG repeats at hHD allele in \( hHD^{(-/-)/Msh2^{(+/+)}} \) animals have been found in their sperm (14). Consistent with this, the \( hHD^{(-/-)/Msh2^{(+/+)}} \) males displayed intergenerational expansion (Fig. 1B). In the same lineage, the \( hHD^{(-/-)/Msh2^{(-/-)}} \) progeny showed deletion of the repeats during transmission (Fig. 1A and B). These deletions, however, were not present in their germ cells (14). Using both a Wilcoxon Mann–Whitney test (Materials and Methods) (23) \( P < 0.01 \) and a Student’s \( t \)-test \( P < 0.015 \), we found that deletions and expansions in the Msh2\((-/-)\) and Msh2\((+/+)\) lines were statistically different.

It is accepted that length changes measured by Genescan analysis do not reflect rare alterations in repeat lengths (45). However, the deletions events in Msh2\((-/-)\) mice occurred during most transmissions and, therefore, were non-random. Thus, rare changes in repeat length would not account for the consistent heritable changes observed in the transmission pattern of \( hHD^{(-/-)/Msh2^{(-/-)}} \) mice (Fig. 1A). Genescan, which is known to be accurate in showing the most prevalent tract sizes (14,30,31,44), indicated that no alterations in repeat lengths were found in the major pool from which transmission occurred in these animals. Since events in sperm did not account for the net intergenerational deletions in \( hHD^{(+/+)\}/Msh2^{(-/-)} \) animals (Fig. 1B) (14), these data raised the possibility that deletions arose somatically after transmission and by a different mechanism than expansion. Expansions appeared to depend on Msh2 as they are associated largely with \( hHD^{(+/-)/Msh2^{(+/-)}} \) mice. In contrast, deletions are independent of Msh2 since they occur in both \( hHD^{(+/-)/Msh2^{(-/-)}} \) and \( hHD^{(-/-)/Msh2^{(+/-)}} \) mice (Fig. 1B) (23). Moreover, lack of Msh2 appeared to promote deletions as the frequency of deletion increases extensively in \( hHD^{(+/-)/Msh2^{(-/-)}} \) mice.

Absence of Msh2 leads to repeat deletion

To further examine how absence of Msh2 promoted deletions, we evaluated instability in somatic cells post-fertilization. Mouse embryonic fibroblasts (MEF) were isolated from \( hHD^{(+/-)/Msh2^{(-/-)}} \) and \( hHD^{(-/-)/Msh2^{(+/-)}} \) animals at E13. Using these cells, we tested whether their inherited repeat length was altered during early somatic development and whether instability depended on the status of Msh2. Included in this analysis were also proliferating fibroblasts from adult HD patients. In both, MEFs from \( hHD^{(+/+)\}/Msh2^{(+/-)} \) animals (Fig. 2B) and HD fibroblasts (Fig. 2A), we found no change in repeat number with increasing cell passage (P) for any microsatellite tested (hHD, ACTC, D1Mit24). Moreover, stability during proliferation was independent of tract length (19 CAG, 41 CAG or 130 CAG). Thus, long repeat tracts in \( hHD^{(+/-)/Msh2^{(+/-)}} \) embryonic cells (Fig. 2B, top) remained stable during proliferation in culture despite the fact that the inherited repeat tract was expanded. As expected, expansion of short normal-sized repeat tracts did not arise during proliferation (Fig. 2A, top and B, bottom).

In contrast to cells containing Msh2, all inherited alleles tested in \( hHD^{(+/-)/Msh2^{(-/-)}} \) embryonic cells were deleted relative to their parental size (Fig. 2C). Moreover, deletions occurred at all loci independent of their repeat lengths (Fig. 2C). As microsatellite loci are known to be very stable in mice of different strains (shown is D1Mit24 locus), these data indicated that deletion had occurred earlier in the embryogenesis, before day E13 when the fibroblasts were isolated (see Materials and Methods). Further, deletions appeared to
arise as the result of Msh2 deficiency. Deletion at hHD allele in Msh2(-/-) cells was obvious when the number of the repeats was compared with those of the transmitting parent (133 and 135 CAG, respectively, Fig. 2C, top). Proliferation studies, thus, indicated that absence of functional Msh2 promoted deletion at both normal-sized and expanded repeats. We did not observe expansion in Msh2(-/-) cells.

Expansion of CAG repeats can occur in somatic cells during break repair when Msh2 is present

In embryonic cells from hHD mice, expansion and deletion appeared to depend differently on Msh2. Expansions have been proposed to occur during replication (22,46,47). However, we did not observe expansion in proliferating MEFs from hHD(+/-)/Msh2(-/-) and hHD(-/-)/Msh2(+-) mice. Other models have suggested that expansions can occur at DNA strand breaks and require the presence of Msh2 (14). Therefore, we tested whether repeats that were normally stable could expand in proliferating cells if breaks were introduced. We challenged the hHD(+/-)/Msh2(+-) MEF cultures by treating them with sub-lethal doses of hydrogen peroxide (H2O2) to create strand breaks. Despite the fact that repeats were stable during proliferation, we found that repeated treatment with low doses of H2O2 led to expansion when Msh2 was present in both human HD fibroblasts (Fig. 3A, left) and in MEFs from hHD(+/-)/Msh2(+-) animals (data not shown). Expansion was observed both in cells attached to the plate and in cells released into the culture media (Fig. 3A, left). Remarkably, when Msh2 was absent, DNA damage induction did not result in any change in CAG repeat number. Thus, Msh2 appeared to be necessary for expansion in the presence of strand breaks (Fig. 3A, right).

We have previously proposed that Msh2 may cause expansion by stabilizing CAG loops that form during repair of DNA strand breaks (14). To test whether such a mechanism was plausible, we performed chromatin immunoprecipitation following oxidative DNA damage. Human HD fibroblasts that harbored one normal and one mutated HD allele received a single acute dose (10 mM) of H2O2 to induce DNA breakage, and cross-linked DNA-protein complexes were then captured on beads coupled with anti-Msh2 antibody. We established
Figure 2. Microsatellites are stable during in vitro proliferation but are deleted relative to their parents. (A) Stability of repeats in human fibroblasts from HD patients containing functional Msh2 protein (HD allele, top; two microsatellite loci, bottom). P4 and P10 are cell passage numbers. Repeat type and repeat numbers are as indicated. (B) Stability of repeats in primary mouse embryonic fibroblasts containing functional Msh2 protein (hHD transgene, top; microsatellite D1Mit24, bottom). Cell passage, repeat type and repeat numbers are as indicated. (C) Repeat deletion in cells deficient in Msh2 protein upon proliferation. Primary mouse embryonic fibroblasts are shown (hHD transgene, top; microsatellite D1Mit24, bottom). Cell passage, repeat type and repeat numbers are as indicated.

Figure 3. Expansion of CAG in the HD locus is associated with Msh2 in the presence of breaks after damage induction. (A) Normal (top) and disease (bottom) alleles for human HD fibroblasts are shown on the left. HD locus for LoVo cells deficient in Msh2 protein is shown on the right. Cells were treated with indicated concentrations of H2O2 multiple times according to the experimental procedure. CAG repeat sizes corresponding to midpoints of the distributions are shown. (B) Chromatin immunoprecipitation with anti-Msh2 antibody following DNA damage induction in human HD fibroblasts [Msh2(+/-)]. Cross-linked chromatin from treated or untreated fibroblasts was immunoprecipitated with anti-Msh2 antibody and analyzed by PCR using primers specific for region adjacent to CAG repeats in HD gene (left) or for kallikrein (right). Genescan traces for PCR products are shown in blue, red peaks correspond to size standards. Top panels represent precipitation on Dynabeads (see Materials and Methods) alone without antibody as a negative control. Bottom panels correspond to the amplification of genomic DNA before immunoprecipitation. The time shown indicates period between the end of the treatment and start of cross-linking procedure.
whether the hHD allele was present in the precipitate by PCR amplification. The kallikrein gene was included as a negative control. This gene contained no repeats, therefore no alternative structure could be formed at induced DNA strand breaks.

We found that Msh2 protein was specifically bound to HD locus 20 min after DNA damaging agent was removed (Fig. 3B, left). On the other hand, no specific binding to kallikrein gene was observed in the chromatin immunoprecipitates with anti-Msh2 antibody (Fig. 3B, right). Taken together, these data were consistent with the model in which Msh2 binds to the CAG repeats in the process of DNA repair. In the absence of strand breaks and/or Msh2, proliferation resulted only in deletions.

Deletion does not occur during replication of adult tissue in vivo

Experiments in mice and MEFs (Figs 1 and 2), taken together, suggested that deletion and expansion might occur by different mechanisms, and possibly at different times in development. To directly test these hypotheses, we established the timing of deletion and expansion events during normal mouse development, in vivo.

To test the fate of repeat tracts during replication in adult mice, we surgically removed a portion of the liver and allowed regeneration to occur (Fig. 4A) (48,49). These experiments were performed in 6- to 7-week-old hHD(−/+)/ Msh2(+/+) animals, well before age-dependent expansion occurs (14,30,31,44). In multiple animals, replication of liver cells during regeneration resulted in no observable alterations of the hHD CAG tract (Fig. 4A). Replication during liver regeneration is limited to a few rounds. More extensive replication can be forced in the spleen by treating animals with a Staphylococcal enterotoxin B that induces proliferation (50,51). Again, induced replication in the spleen did not give rise to any observable alterations in the CAG tract of the hHD allele (Fig. 4B), indicating that deletion events either do not arise from replication in mammals or are restricted to earlier developmental stages.

We next evaluated repeat stability in adults, postnatal and prenatal animals, as well as embryos (Fig. 5A). Expansions were evident in the germ cells of 7- to 11-week-old animals as previously reported. However, we found no other alterations in repeat tracts beyond the inherited number in any somatic tissue of postnatal animals (Fig. 4A). Similarly, in neonatal or prenatal animals at E13, all tissues tested showed identical distributions, indicating that repeat tracts were stable (Fig. 5A). Amplification of cells from eight-cell embryos, on the other hand, clearly revealed shifted distributions and asymmetries due to differences in repeat length (Fig. 5A and C). These data were confirmed in multiple PCR reactions from multiple animals (Fig. 5C). The asymmetries were specific to the embryos since the probability of contamination with extraembryonic cells was minimized by visual monitoring of blastomere dissection through video camera connected to stereo microscope (Fig. 5B).

In the analysis of the CAG distributions in each blastomere (a single cell of the eight-cell embryo), the size of the major peak within the distributions (peak with the largest area) was taken as the correct CAG length. The sizes of major peaks were clearly different in each blastomere. Heterogeneity within a distribution, on the other hand, was an artifact that arose from nested PCR (see Materials and Methods). To confirm that the major peaks accurately reflected the repeat length, we amplified mouse or plasmid DNA (data not shown) with known CAG repeat lengths and tested whether the midpoint peak corresponded to the known size (Fig. 5D). In all cases, the measured midpoint peak reproduced the known repeat length of the starting material. Further, the midpoint peak did not depend on the number of cycles in the PCR amplification or DNA concentration (Fig. 5D). Therefore, the shifts in the distributions in the embryo did not arise from the PCR method.

In further analysis, when superimposed (Fig. 5E, top and middle), additive scans from each blastomere formed a symmetrical distribution that reproduced the distribution for the entire eight-cell population. If instability events (deletion or expansion) take place during first divisions in early embryo, then we expect to see up to four discrete sizes in the eight-cell embryo due to three cell divisions that occur post-fertilization. Indeed, when major peaks of the scans from individual blastomeres are overlayed, three peaks corresponding to three different CAG sizes were evident (Fig. 5E, bottom). These analyses confirmed that somatic changes in repeat length during cell proliferation were restricted to early embryo development. The alterations in the eight-cell embryo were likely to contain some deletions. However, we could not unambiguously determine whether the shifted distributions were deletion or expansion events since the initial CAG length in the transmitting sperm was not known.

We therefore used a second approach. We directly tested changes in repeat length within the hHD allele in germ cells that undergo both proliferation and terminal differentiation.

Figure 4. Replication in adult tissue does not result in deletion of CAG triplets in animals. (A) Liver regeneration. Liver before (top) or after (bottom) regeneration; (right) Genescan analysis of repeat distributions from two R6/1 animals before (−) and after (+) regeneration compared with the tail. (B) (left) Spleen from R6/1 animal before (−) and after (+) Staphylococcal enterotoxin treatment; (right) Genescan analysis for tissues from two animals before (−) and after (+) toxin treatment. Tissue types are indicated.

Human Molecular Genetics, 2004, Vol. 13, No. 24 3061
Figure 5. Deletions in somatic cells are restricted to early embryo divisions and to proliferative stages of germ cell development. (A) (top) Representative examples of R6/1 [hHD(−/+)Msh2(+/+)] animals at the indicated stages in development, wks is weeks of age, P1 is postnatal day 1, E13 is embryo day 13, 8-cell is the eight-cell stage of embryo development. Five individual cells (1–5) from the same embryo and blank (B) are shown; (middle) Genescan analysis of the indicated tissues corresponding to stages of development. Each scan is representative of two to three reactions; numbers indicate the midpoint of the repeat distribution; (bottom) mean length changes in distribution of repeats from Genescan analysis. Each bar represents corresponding tissue shown in the middle panel. (B) (top) Images of the step-by-step dissecting procedure; eight-cell embryo (left), embryo without zona pellucida (middle) and disintegrated individual blastomeres (right). (C) Genescan traces of individual blastomeres for two eight-cell embryos. Top scans represent repeat distribution in the tail of the father. Blank indicates no DNA control. (D) Control experiments for single cell PCR analysis. Comparison of Genescan traces for amplification of different amounts of DNA. Two types of tissues from a mouse are shown: tail (right) and lymphocytes (left). (E) Repeat distribution construction based on scans from single cells of the eight-cell embryo. (Top) Overlay of the scans from each cell of the embryo, (middle) solid bell curve represents symmetrical distribution that single cell scans form when they are superimposed, (bottom) overlay of only mayor peaks from each single cell Genescan trace for two embryos shown in (C).
Thus, dividing and non-dividing cells within related cell populations could be directly compared. Repeat lengths were measured in non-dividing spermatooza (SZ) in epididymis and dividing spermatogonia (SG) in the testes in both hHD(−+/+)Msh2(+/+) and hHD(−+/+)/Msh2(−/−) animals and compared these to CAG lengths in somatic tissue (tails are shown).

As we have reported previously, repeats were expanded in the SZ of hHD(−+/+)Msh(+/+) mice (Fig. 6, top). However, we found that repeats in the testes of hHD(−+/+)/Msh2(−/−) mice were consistently deleted relative to either somatic cells (tail) or SZ within the same animal (Fig. 6, testes). Deletions were also observed in testicular cells of hHD(−+/+)/Msh2(−/−) but expansions were not present in their SZ (14). These data appeared to support a model in which expansions and deletions occurred by different mechanisms during mammalian development. Deletions, but not expansions, of long trinucleotide repeats were associated with proliferation phases in early development.

**DISCUSSION**

**CAG expansion and deletion occur by different mechanisms at disease–length tracts and depend differently on Msh2**

CAG expansion mutation is dynamic in mammalian development. However, it has been difficult, in humans, to definitively distinguish germ cells events from those in the early embryo. Here, we provide evidence that the length of the repeat tract in the developing organism is influenced by two cell-type–specific processes. These processes occur during different and defined stages of development and reveal a dual role for Msh2. We show that deletion events prevail in the progeny of hHD(−+/+)Msh2(−/−) mice, while the majority of repeat alterations were expansions in their hHD(−+/+)/Msh2(+/+) counterparts. Similarly, intergenerational deletions in Msh2(−/−) background were observed in a mouse model for DM (38). We show that both expansion and deletion depend on the length of the CAG repeats, yet deletions arise during replication, are independent of Msh2 and are largely restricted to early development. In contrast, expansions depend on strand break repair, require the presence of Msh2 and occur later in development. These data support a model in which deletions occur in the presence of Msh2, but they are promoted by its absence during proliferation.

That expansion and deletion can occur by different mechanisms is supported by key experiments in other systems. In yeast chromosomes that are engineered to contain a specific cutting site, single breaks can be introduced by HO endonuclease cleavage (52,53). When breaks occur on the chromosome, repair can occur by gene conversion from a plasmid containing homologous sequence (52,54). Expansions occurred exclusively at the break sites although deletions could also be observed (52). Within the same cells, however, only deletions occurred on the replicating plasmids (54). Expansion is known to arise in simple organisms during repair of double strand breaks (52,54,55) and at stalled replication forks in long CGG and CTG repeats (56,57).

The same phenomenon may account for the effects relating to origin of replication. Recently, it has been shown that the occurrence and frequency of expansions and deletions in replicating mammalian systems depend on the proximity of repeat tract to origin of replication (58). The mechanism by which expansions occur in these studies is not yet clear. However, unligated ends are created at bi-directional origins of replication proximal to the repeats at the sites where synthesis begins (58). The presence of proximal breaks and replication in the same system may explain the occurrence of both expansion and deletion events (52,54).

Deletions, in contrast to expansions, prevail in rapidly replicating systems. Numerous studies in bacteria and yeast have demonstrated that the deletion rate of long repeat tracts far exceeds that of expansion in wild-type repair background (32,46,59,60). All studies agree that loss of Msh2 greatly enhances the rate of deletions (32,34,36). We find that deletion events have already occurred by the E13 embryo, after which they were stably maintained over period of 21–25 doublings (Fig. 2A and B). Others have observed small expansions in proliferating HD fibroblasts cultured for longer times and in longer repeat tracts (61). However, replication and repair cannot be separated in these systems. DNA damage is known to accumulate in cultured cells with increasing...
passage, and studies in yeast have shown that long tracts of repeats tend to break during replication (52,56,57). Therefore, expansions in proliferating HD fibroblasts can arise in replicating cells, but from strand break repair process rather than replication per se.

We find that, in mammals in vivo, the somatic deletions have largely occurred by the eight-cell stage in the embryo in which cells are rapidly dividing. There may be two contributing factors that explain why deletions are restricted to early embryogenesis and do not occur in adult cells. First, checkpoints are not yet in place (Xenopus and Zebra fish); second, the replication rate in blastocysts is higher than in adult cells (Xenopus) (62–66). Recent data in yeast have confirmed that deletions at long triplet repeat sequences are substantially favored when checkpoint proteins are dysfunctional (67).

**Deletion bias during replication is likely to occur by a slippage mechanism**

Extensive data support the involvement of secondary DNA as an intermediate in the expansion process. The formation of secondary structure may explain why expansions and deletions have a different dependence on Msh2. In germ cells, we (14) and others (22) have observed that a significant fraction of expansions occur in the haploid stage of spermatogenesis, at which time neither a homologous chromosome nor a sister chromatid is available. On the basis of these data, we have previously proposed an intra-allelic simple gap-repair mechanism for expansion in germ cells (14). In this model, stable hairpins with mispaired bases can form at the site of breakage. The Msh2 complex recognizes the mispaired bases of the CAG hairpin in the HD gene and prevents re-annealing of the CTG or CAG flaps to its complementary strand. Stabilization of the hairpin loops creates a gap that, after repair, is ligated into the DNA and provides precursors for expansion (14) (Fig. 7A). Loops of DNA can also be introduced by slippage during the process of gap-repair synthesis (Fig. 7A)

For larger repeats, it is possible that gap repair can involve other mechanisms. For example, recent data in yeast have suggested that large expansions in DM and FRAX are likely to occur from breaks made in DNA during replication pausing, and repair may occur by recombination or incorporation of larger, more complex structures that form at the breaks. In other systems, double strand breaks are also known to be repaired by homologous recombination (55,68,69), which depends on Msh2 (70) (Fig. 7A). In this mechanism, gaps formed after resection of DNA ends are filled by template-directed, repair-dependent DNA synthesis. Although plausible (Fig. 7A, left), recombinational repair is not consistent with recent studies in mouse model for DM. For DM mice, loss of Rad52 and Rad54, both enzymes that are required for homologous recombination, had no effect on expansion (38). Thus, repair of single strand breaks by intra-allelic gap-filling mechanism appears to be the more likely mechanism for expansion at large or small repeats in mammals (Fig. 7A, right). Whatever the precise mechanism, we demonstrate here that expansion of CAG repeats requires strand breaks in cells and cannot occur if Msh2 is absent.

In contrast to expansion, deletion occurs during cell proliferation, most likely by a replication slippage mechanism. Deletions appear to depend on the ability of the repeat to form secondary structures, but deletions do not require the presence of Msh2 (Fig. 7B). During proliferation, the polymerase passes over the slipped structure and traps it (Fig. 7B). Thus, a looped structure is forced to be stable since there is no complementary strand for re-annealing. Msh2 binding can occur at trapped hairpin loops, but binding it is not required for loop stabilization (Fig. 7B).

**Deletion events may protect against expansion during early development**

Different mechanisms for expansion and deletion can account for intergenerational transmission patterns. In *hHD(−/+)/Msh2(−/−)* mice, expansion is absent in the sperm and...
cannot compensate for the somatic deletion events in the embryo. Consequently, in hHD(−/+)/Msh2(−/−) animals, there is net deletion during intergenerational transmission (Fig. 7C). In contrast, in hHD(−/+)/Msh2(+/+) animals, expansion occurs late in sperm development (14) and, in the majority of the cases (Fig. 1), overrides the deletion events that precede them (Fig. 7D). Thus, in hHD(−/+)/Msh2(+/--) animals, there is net intergenerational expansion during transmission (Fig. 7D). It has been puzzling why intergenerational changes in the germ cells of HD mice (+1−5 repeats) were smaller than those observed in age-dependent expansions (5−20 repeats). Our data indicate that expansions are likely to be larger than previously suspected. Since expansions in the sperm (SZ) must override the testicular and somatic deletion events, the size of the expansion after transmission is masked. We measure the size of the average intergenerational expansion to be between 8 and 10 repeats (Fig. 6).

Finally, it has never been clear why expansion mutations should be favored in the early embryo when these cell types hold the promise of a new organism. Our results suggest that replication may safeguard the genome against expansion of disease−length alleles. While large repeat tracts can have deleterious effects on gene expression, a deletion bias of replication during intergenerational transmission (Fig. 6) must override the testicular and somatic deletion events, the size of the expansion after transmission is masked. We measure the size of the average intergenerational expansion to be between 8 and 10 repeats (Fig. 6).

MATERIALS AND METHODS

Mouse lines and breeding

Transgenic male mice B6CBA-TgN R6/1 (71) were bred to negative B6CBA female partners. Litters were routinely screened for the presence of hHD transgene by PCR. Msh2 homozygous and heterozygous knockout animals harboring the hHD gene were generated by crossing B6CBA-TgN R6/1 and C57/B6 Msh2(−/−) mice (provided by R. Kucherlapati). Absence of Msh2 gene was confirmed by PCR (as described subsequently).

Tissue isolation and Genescan analysis for sizing repeat length

Various somatic tissues were collected from wild-type and Msh2 knockout mice carrying the hHD transgene. DNA was isolated, and the CAG repeat sizes were determined as described previously (71). Scanned traces of the PCR products resolved on polyacrylamide gel were obtained and evaluated using Genescan analysis V3 (Perkin Elmer). The peak with the largest area was taken as the midpoint of the peak distribution in Genescan traces after normalization with respect to internal standards.

Msh2 genotyping

Typically, 50 ng of DNA was amplified in a reaction with primers specific (10 pmol) to the wild-type allele (5′-CCC TCTGGTGTAGCCATCTTA-3′ and 5′-TTCTGCTGTTGTG C TCTGGAAAT-3′) or to the Msh2(−/−) allele (5′-GCCAGC TCATTCTCTCCTC-3′ and 5′-CCCTCTCCTGGAGCCAT CTTA-3′). In both cases, Taq polymerase (Perkin Elmer) was used in 10 mM Tris–HCl, 10 mM KCl, 0.2 mM dNTP and 1 mM MgCl2 (pH 8.9). The PCR products were resolved by electrophoresis on 1% agarose gel. Wild-type and mutant alleles were distinguished by the size of the product (180 and 300 bp, respectively).
Embryo isolation

Eight-cell embryos from pregnant B6CBA females bred with B6CBA-TgN R6/1 males were collected at 56–60 h post-coitum (p.c.). Both oviducts and uterine horns were dissected and flushed with phosphate buffered saline allowing the recovery of eight-cell stage embryos. Zona pellucida was removed from the embryo using acidified Tyrode’s solution (Sigma). Single blastomeres were obtained by disaggregating the entire embryo with a flame-polished finely pulled glass capillary under a dissecting stereomicroscope (Leica Mz12.5) at 40× magnification. Single blastomeres were each washed through three drops of PBS containing 5 mg/ml polynivylpyrolidine (PVP, Sigma) and lysed individually in 10 μl of proteinase K-based lysis buffer (Quantilyse, Hamilton Thorne Research, Inc.). Cell lysis was carried out at 50°C for 30 min, followed by 15 min incubation at 98°C to inactive the proteinase K. Aliquots of PBS from the wash drops served as contamination controls.

Single cell PCR

CAG repeat size in a single blastomere was determined by nested PCR. For primary PCR reaction with outer primers [(1) 5’-TTTACCTGGGCCCAGAGC-3’ (10 pmol) and (2) 5’-AGCGGCTCCTCAGCCACA-3’ (10 pmol)] DNA was amplified in 67 mM Tris–HCl (pH 8.8), 16.6 mM (NH4)2SO4, 2 mM MgCl2, 0.17 mg/ml BSA, 10 mM mercaptoethanol, 10% DMSO, 200 mM dNTPs with 0.5 U Taq polymerase. Cycling conditions were 5 min at 95°C, 20 cycles of 30 s at 95°C, 30 s at 58°C, 90 s at 72°C, followed by 5 min extension at 72°C. An aliquot of the primary PCR product (1 μl) was used in a secondary PCR round. forty cycles of amplifications were carried out using primers and conditions described by Serman et al. (73).

Liver regeneration analysis

Liver regeneration in B6CBA-TgN R6/1 mice was induced by partial hepatectomy. Animals that were 6- to 7-week-old were anesthetized with nembutal and subjected to surgery. Median hepatic lobule constituting 40–50% of liver mass was ligated with silk suture and removed through abdominal incision (74). Mice were sacrificed 2 weeks later after liver regeneration was complete. Samples of tissues before (normal) and after surgery (regenerated) were taken for DNA isolation. Repeat sizes were determined by Genescan analysis.

Spleen proliferation

Proliferation of spleens from hHD transgenic mice was induced by a single intra-peritoneal injection of 50 μg of Staphylococcal enterotoxin B. Animals were sacrificed 24–48 h later. DNA from spleen (proliferated) and tail (control tissue) was isolated and analyzed for repeat size by Genescan as described.

Statistical analysis

Non-parametric Wilcoxon Mann–Whitney test was used to compare repeat distributions of two groups of mice: Msh2(+/−) and Msh2(+/+) mice. The null hypothesis stating that medians of the two groups are the same was tested. All observations (repeat size for each individual animal) were converted to ranks from 1 (lowest) to n = nA + nB (nA and nB are sample sizes of the two groups) irrespective of group membership. The sums of the ranks in each group, RA and RB, were calculated and used in the equation for statistics calculations:

\[ U_A = R_A - \frac{n_A(n_A + 1)}{2} \quad \text{or} \quad U_B = R_B - \frac{n_B(n_B + 1)}{2}. \]

The smaller of the two U values was selected and compared to the critical U. The null hypothesis was rejected on the basis of the calculated U being smaller than U critical (P < 0.01 for a given sample number).

ACKNOWLEDGEMENTS

The authors would like to thank Drs C. Spiro, R.B. Dyer, B. Owen, E. Trushina, R.J. Maher and T.M. Therneau for critical reading of the manuscript, Dr R. Kucherlapati for providing Msh2 knockout mice, David Walker for technical expertise on embryo isolations and John D. Badger for assistance in photography. This work was supported by the Mayo Foundation, the Hereditary Disease Foundation and the National Institutes of Health (NS4073802S1 and GM066359).

REFERENCES


