Neurobiology of Disease 76 (2015) 98–111

Contents lists available at ScienceDirect

Neurobiology of Disease

journal homepage: www.elsevier.com/locate/ynbdi

Age-, tissue- and length-dependent bidirectional somatic CAG•CTG repeat instability in an allelic series of R6/2 Huntington disease mice

Eloise Larson [a,b], Ian Fyfe [a], A. Jennifer Morton [a], Darren G. Monckton [b,*]

[a] Department of Physiology, Development and Neuroscience, University of Cambridge, Tennis Court Road, Cambridge CB2 3DY, UK
[b] Institute of Molecular, Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Davidson Building, University Avenue, Glasgow G12 8QQ, UK

ARTICLE INFO

Article history:
Received 31 August 2014
Revised 30 November 2014
Accepted 25 January 2015
Available online 3 February 2015

Keywords:
Huntington disease (HD)
Polyglutamine expansion
Trinucleotide repeat expansion
Somatic instability
Mutation
R6/2 mice

ABSTRACT

The expansion of simple sequence CAG•CTG repeats is associated with a number of inherited disorders including Huntington disease (HD), myotonic dystrophy type 1 and several of the spinocerebellar ataxias. Inherited disease-associated alleles usually exceed 40 repeats and may be in excess of 1,000 repeats in some disorders. Inherited allele length is inversely proportional to age at onset, and frequent germline expansions account for the striking anticipation observed in affected families. Expanded disease associated alleles are also somatically unstable via a pathway that is age dependent and tissue specific, and also appears to be expansion biased. Somatic expansions are thought to contribute toward both tissue specificity and disease progression. Here we have examined the somatic mutational dynamics in brain and peripheral tissues from an allelic series of R6/2 HD transgenic mice inheriting from 52 to >700 CAG repeats. We found age-dependent, tissue-specific somatic instability, with particularly large expansions observed in the striatum and cortex. We also found a positive increase in somatic instability with increasing allele length. Surprisingly, however, the degree of somatic variation did not increase in a linear fashion, but leveled off with increasing allele length. Most unexpectedly, the almost exclusive bias toward the accumulation of expansions observed in mice inheriting smaller alleles was lost, and a high frequency of large somatic contractions was observed in mice inheriting very large alleles (>500 repeats). These data highlight the bidirectional nature of CAG•CTG repeat instability and the subtle balance that exists between expansion and contraction *in vivo*. Defining the dynamics and tissue specificity of expansion and contraction is important for understanding the role of genetic instability in pathophysiology and in particular the development of novel therapies based on suppressing expansions and/or promoting contractions.

Introduction

The expansion of simple sequence DNA repeats is associated with a growing number of inherited human disorders including Huntington disease (HD), myotonic dystrophy (DM), several of the spinocerebellar ataxias and, more recently, amyotrophic lateral sclerosis (Gomes-Pereira and Monckton, 2006; Orr and Zoghbi, 2007; Lopez Castel et al., 2010; Orr, 2011). The repeat tract is polymorphic in the general population with alleles typically in the range of 5–30 repeats. Disease-associated alleles have expanded beyond this length (mostly >40 repeats) and inherited alleles may exceed 1,000 repeats in some disorders. Typically, longer inherited alleles are associated with a greater disease severity and an earlier age at onset. Once into the expanded disease-associated range, the repeat tracts become highly genetically unstable in both the germline and soma (Gomes-Pereira and Monckton, 2006; Lopez Castel et al., 2010). Germline mutation rates can approach 100% and a major bias toward expansions explains the striking anticipation observed in many of the associated disorders. Similarly, somatic instability appears to be highly biased toward further expansions. Somatic instability is also highly tissue specific and expansion sizes are frequently much higher in the affected tissue. The dynamics of somatic mosaicism are consistent with a major role for genetic instability, both in driving genotype–phenotype relationships and in the tissue specificity and progressive nature of the symptoms (Gomes-Pereira and Monckton, 2006; Kaplan et al., 2007; Swami et al., 2009; Lopez Castel et al., 2010; Morales et al., 2012).

Most of the disorders caused by the expansion of a simple sequence repeat are associated with the expansion of trinucleotide repeats, most frequently CAG•CTG (Gomes-Pereira and Monckton, 2006; Orr and Zoghbi, 2007; Lopez Castel et al., 2010; Orr, 2011). Of these disorders, HD is one of the most frequent and best studied. HD is a very severe neurological disorder associated primarily with neurodegeneration in the striatum (Bates et al., 2002; Shoulson and Young, 2011) and is one of at least nine disorders caused by the expansion of a polyglutamine-

Abbreviations: HD, Huntington disease; DM, myotonic dystrophy; DM1, myotonic dystrophy type 1; PCR, polymerase chain reaction; YAC, yeast artificial chromosome; BAC, bacterial artificial chromosome; SP-PCR, small pool-PCR; SCA1, spinocerebellar ataxia type 1; SCA7, spinocerebellar ataxia type 7.

* Corresponding author.
E-mail addresses: e.larson.1@research.gla.ac.uk (E. Larson), ian@ianfyfe.co.uk (I. Fyfe), ajm41@cam.ac.uk (A.J. Morton), darren.monckton@glasgow.ac.uk (D.G. Monckton).

Available online on ScienceDirect (www.sciencedirect.com).

http://dx.doi.org/10.1016/j.nbd.2015.01.004
0969-9961/

encoding CAG•CTG repeat (Gusella and MacDonald, 2000). In HD, the polyglutamine tract lies in exon 1 of the *HTT* gene (The Huntington's Disease Collaborative Research Group, 1993) and varies in length from 10 to ~34 repeats in the general population and is expanded beyond 40 repeats in the majority of HD patients (Duyao et al., 1993; Snell et al., 1993; The Huntington's Disease Collaborative Research Group, 1993). Most adult onset patients inherit 40–50 repeats, and in this range there is a very dramatic genotype–phenotype relationship with allele length increases of a single CAG repeat precipitating a reduction in age at onset by 2.6 years (Andresen et al., 2007). Alleles >50 repeats are frequently associated with the severe juvenile form of the disorder, although additional increases in inherited allele length have a less dramatic impact on age at onset (Andresen et al., 2007). Nevertheless, there are at least five patients reported who have inherited >200 repeats (Nance et al., 1999; Milunsky et al., 2003; Seneca et al., 2004; Reynolds et al., 2008; Nicolas et al., 2011). The expanded HD allele is highly unstable with a maternal germline mutation frequency of ~80% with a ~2:1 expansion bias (Duyao et al., 1993). Expansions are both more frequent and tend to be larger in size upon transmission from males (Duyao et al., 1993). Longer alleles have a higher mutation frequency, approaching 99% with a >95% expansion bias for alleles >50 repeats in the male germline (Leeflang et al., 1999). The differing germline mutational dynamics account for the large excess (~80%) of paternal transmission of juvenile HD.

As well as being unstable in the germline, the expanded HD repeat is also unstable in somatic tissues. Most notably, the HD repeat is particularly unstable in the brain, with the largest expansions observed in the striatum and cortex (Telenius et al., 1994; De Rooij et al., 1995), the brain areas primarily affected in this disorder. Indeed, sensitive small pool-PCR analyses have revealed somatic expansions in some cells in the striatum and cortex of >1,000 repeats in patients who inherited alleles in the range of 40–50 repeats (Kennedy et al., 2003). Importantly, laser capture micro-dissection of single cells has been used to demonstrate that these very large somatic expansions are primarily observed in striatal and cortical neurons as opposed to the supporting glial cells (Shelbourne et al., 2007; Gonitel et al., 2008). As well as favoring a cell-division-independent mechanism for generating somatic mutations, these data strongly implicate somatic instability in the tissue and cell-type specificity of the disease. In further support of this link, it has been recently demonstrated that patients presenting with an extremely young age at onset relative to their inherited repeat size presented with a more positively skewed repeat length distribution in the cortex relative to patients with similar inherited allele lengths who developed the symptoms much later in life (Swami et al., 2009). Together, these data not only directly implicate somatic expansion in the disease pathway but also highlight somatic instability as a potential therapeutic target in this disorder (Gomes-Pereira and Monckton, 2006; Lopez Castel et al., 2010). Thus, understanding the dynamics and mechanisms of repeat instability in HD remains a research priority. However, further research in this area is compromised by the inaccessibility of the affected tissue and the reliance on post mortem end-stage tissue in which a high proportion of the vulnerable populations of cells have already been lost (Vonsattel et al., 1985). Indeed, evidence suggests that the repeat distribution observed in end-stage striatal tissue does not reflect the true level of variation observed earlier in the disease course (Kennedy et al., 2003). In contrast to disorders such as myotonic dystrophy type 1 (DM1) in which patients typically inherit many hundreds of repeats (Morales et al., 2012), the level of somatic instability in the peripheral tissue of HD patients is usually very low (Leeflang et al., 1995; Veitch et al., 2007). Although analysis of somatic mosaicism in buccal cells has been used to confirm a major role for allele length in driving somatic instability, the mutation frequencies observed were so low as to preclude a high throughput analysis of repeat length variation in a large cohort of HD patients (Veitch et al., 2007).

To overcome the limitations of tissue availability and further understand pathogenesis and the mechanisms and role of genetic instability in HD and related disorders associated with the expansion of CAG•CTG repeats, numerous mouse models have been created (Heng et al., 2008). These comprise various knock-in HD models (White et al., 1997; Shelbourne et al., 1999; Wheeler et al., 1999; Lin et al., 2001) and a variety of HD transgenic mice, including short gene fragment models (Mangiarini et al., 1996), as well as cDNA (Reddy et al., 1998; Schilling et al., 1999), YAC (Hodgson et al., 1996), BAC (Gray et al., 2008) and conditional models (Yamamoto et al., 2000). Such mice have been used to reveal a wide range of important insights, including the identification of neuronal intranuclear inclusions in HD (Davies et al., 1997); the confirmation that expanded polyglutamine tracts are inherently toxic (Ordway et al., 1997); the identification of transcriptional (Luthi-Carter et al., 2000), mitochondrial (Bogdanov et al., 2001), proteosomal (Martin-Aparicio et al., 2001), intracellular Ca^{2+} (Deckel et al., 2002), autophagic (Kegel et al., 2000) and axonal transport dysfunction (Trushina et al., 2004); and an increasingly recognized role for protein phosphorylation (Pardo et al., 2006) and interactions with the wild-type function of the associated protein (Kratter and Finkbeiner, 2010). Mouse models have also been used to great effect to investigate CAG•CTG repeat instability, supporting a cell-division-independent mechanism for repeat expansion (Kaytor et al., 1997; Lia et al., 1998; Fortune et al., 2000; Kennedy and Shelbourne, 2000; Gonitel et al., 2008), an important role for *cis*-acting modifiers of instability (Mangiarini et al., 1997; Monckton et al., 1997; Zhang et al., 2002; Libby et al., 2003; Libby et al., 2008) and key functions for a variety of DNA repair genes (Manley et al., 1999; van den Broek et al., 2002; Gomes-Pereira et al., 2004b; Kovtun et al., 2007; Hubert et al., 2011). Most notably, mouse models have been instrumental in revealing a critical role for the *Msh2*, *Mlh1*, *Mlh3*, *Msh3* and *Pms2* DNA mismatch repair genes in mediating both somatic and germline expansions (Manley et al., 1999; Kovtun and McMurray, 2001; van den Broek et al., 2002; Savouret et al., 2003; Gomes-Pereira et al., 2004b; Foiry et al., 2006; Dragileva et al., 2009; Pinto et al., 2013; Tome et al., 2013). Providing further support for a disease model in which somatic expansion drives downstream pathology, it has been demonstrated that a number of aspects of pathology are delayed in an HD knock-in mouse model on either an *Msh2* or *Mlh1* null background in which somatic instability is suppressed (Wheeler et al., 2003; Pinto et al., 2013).

To date, the most widely used CAG•CTG repeat mouse models are the R6 HD series, in particular the R6/1 and R6/2 lines (Mangiarini et al., 1996; Mangiarini et al., 1997); these two lines having played a major role in many of the observations discussed above (Heng et al., 2008). The R6 series of mice were created by random integration into the mouse genome of a 1.9 kb fragment of the human *HTT* gene incorporating ~1 kb of 5′-untranslated region, exon 1 with ~130 CAG repeats and the first 262 bp of intron 1. In the R6/1 line, the transgene integrated into mouse chromosome 3 (Chiang et al., 2012) and the original founder animal retained a CAG tract of ~116 repeats (Mangiarini et al., 1996; Mangiarini et al., 1997). In the R6/2 line, the transgene integrated into mouse chromosome 4 (Cowin et al., 2011; Chiang et al., 2012) and the original founder animal retained a CAG tract of ~144 repeats (Mangiarini et al., 1996; Mangiarini et al., 1997). Although originally created to model CAG•CTG repeat instability, these mice also develop a progressive neurological phenotype that mimics aspects of HD pathophysiology (Mangiarini et al., 1996; Carter et al., 1999; Lione et al., 1999). The phenotype is particularly acute in the R6/2 line with mice with ~140 repeats presenting with overt symptoms at 9–11 weeks and dying by 12–13 weeks. The phenotype in the R6/1 line is much less severe with symptoms not observed until 4–5 months of age. These phenotypic differences have meant that the R6/1 line has been most commonly used to investigate genetic instability, while the R6/2 line has been more widely used to explore aspects of the downstream pathology. Interestingly, the repeat is unstable in the R6/2 line with both somatic and germline instability reported in the initial analyses of these mice (Mangiarini et al., 1997). Indeed, selective breeding of mice inheriting germline mutations has enabled the generation of an