CAG•CTG repeat instability in cultured human astrocytes

Brian T. Farrell and Robert S. Lahue*

Eppeley Institute for Research in Cancer and Allied Diseases and Department of Pathology and Microbiology, University of Nebraska Medical Center, Box 986805, Omaha, NE 68198-6805, USA

Received July 11, 2006; Revised August 3, 2006; Accepted August 6, 2006

ABSTRACT

Cells of the central nervous system (CNS) are prone to the devastating consequences of trinucleotide repeat (TNR) expansion. Some CNS cells, including astrocytes, show substantial TNR instability in affected individuals. Since astrocyte enrichment occurs in brain regions sensitive to neurodegeneration and somatic TNR instability, immortalized SVG-A astrocytes were used as an ex vivo model to mimic TNR mutagenesis. Cultured astrocytes produced frequent (up to 2%) CAG•CTG contractions in a sequence-specific fashion, and an apparent threshold for instability was observed between 25 and 33 repeats. These results suggest that cultured astrocytes recapitulate key features of TNR mutagenesis. Furthermore, contractions were influenced by DNA replication through the repeat, suggesting that instability can arise by replication-based mechanisms in these cells. This is a crucial mechanistic point, since astrocytes in the CNS retain proliferative capacity throughout life and could be vulnerable to replication-mediated TNR instability. The presence of interruptions led to smaller but more frequent contractions, compared to a pure repeat, and the interruptions were sometimes deleted to form a perfect tract. In summary, we suggest that CAG•CTG repeat instability in cultured astrocytes is dynamic and replication-driven, suggesting that TNR mutagenesis may be influenced by the proliferative capacity of key CNS cells.

INTRODUCTION

Trinucleotide repeat (TNR) expansion causes at least 15 debilitating and often fatal neurodegenerative and neuromuscular diseases such as Huntington’s disease (HD), myotonic dystrophy (DM1), dentatorubropallidoluysian atrophy (DRPLA) and many of the spinocerebellar ataxias (SCAs) [reviewed in Refs (1–3)]. Each disease has a unique clinical manifestation, but the underlying genetic mutation is the same—an expansion in the length of a TNR alters the function or expression of a nearby gene. The link between this unique form of mutagenesis and human disease makes understanding the mechanisms of TNR instability both medically relevant and scientifically important.

It is clear that new DNA must be synthesized to expand a TNR, but there are differing models as to the nature of the synthesis. One school of thought is that non-replicational mechanisms, such as error-prone DNA repair, are the major sources of instability in key affected tissues. Evidence from several model systems supports this contention. For example, intergenerational instability in male HD transgenic mice was traced to post-mitotic, haploid germ cells (4). Instability in cell lines derived from DM1 mice did not correlate with cell division (5). Moreover, a DM1 mouse line that showed remarkable somatic instability showed the greatest instability in organs that were not highly proliferative, namely kidney and brain (6). In addition, a transcription-based assay for instability in cultured human cells found that large contractions were independent of proliferation, occurred equally in confluent and non-confluent cultures, and were reduced by knockdowns of repair proteins Msh2, Msh3 and XPA (7).

In addition to post-mitotic mechanisms, a second idea is that TNR mutations can arise as a result of aberrant DNA replication in proliferating cells. Evidence to support replication-mediated instability again stems from several sources. Analysis of CAG repeat mutations in sperm tissue from HD men showed instability in proliferating pre-meiotic cells, and instability continued through meiosis into post-meiotic cells (8). Other studies have shown that TNR mutability can be modulated up to 10-fold by altering the direction of DNA replication through a repeat allele in human cells (9), human cell extracts (10) and nonhuman primate cells (9,11). Also, a repeat tract’s proximity to the origin of replication appears to influence the bias for expansions or contractions in nonhuman primate cells (11) and replication inhibitors can modulate instability in primary human DM1 fibroblasts (12). When this evidence is placed in context with many microbial studies that invoke DNA replication as a major source of TNR expansions and contractions (13–16), these findings indicate...
that TNR instability likely has more than one mechanistic source and that repair, transcription and replication need to be evaluated for their contribution to TNR mutagenesis.

Since the brain is a target organ in many of the TNR diseases, an important consideration is that terminally differentiated neurons are outnumbered by other critical cell populations such as astrocytes, oligodendrocytes, and microglia that retain proliferative capacity throughout life. This consideration is important when interpreting studies in humans and transgenic mice showing age-dependent increases in TNR length in the central nervous system (CNS) [reviewed in Ref. (3)]. A main conclusion in some of these studies (6,17–19) is that non-replicative repair must be the key mechanism for triggering TNR expansion because neurons are non-dividing and thus do not replicate their DNA. A concern with this interpretation is that the affected brain regions include neurons and non-neuronal cells. Since these PCR analyses included DNA from cells such as glia that retain proliferative capacity, one cannot rule out replication as a source of the expansions. In fact, several labs have proposed that replication-derived TNR mutations in glial cells account for a large portion of the somatic mutations seen in autopsy tissue from TNR diseased patients (20–22). Moreover, a laser-capture microdissection study showed that cortical and cerebellar glial cells from DRPLA patients exhibit greater somatic TNR instability than surrounding neurons (23). This is especially significant because the cerebellum very rarely exhibits somatic TNR instability (18), implying that the cerebellum’s rare TNR mutations are enriched in cells capable of proliferation. Taken together with recent evidence that glia help protect neurons from the degenerative effects of expansions in animal models and that glia themselves are sensitive to the toxic effects of TNR expansion (24,25), a clearer picture of the factors that promote or prevent somatic TNR instability in astrocytes and other glial cells is needed.

To better understand TNR instability in key CNS cells such as astrocytes, we took advantage of an immortalized human astrocytic cell line, SVG-A, and a previously described shuttle vector assay for CAG\textsuperscript{CTG} repeat contractions (9). The results from this model system indicate several striking similarities with human TNR genetics and provide new mechanistic insights into the fascinating but complex problem of triplet repeat mutagenesis. We investigated TNR contractions for several reasons. First, contractions usually outnumber expansions at allelic levels below the threshold for the CAG\textsuperscript{CTG} repeats at the HD locus (26), the androgen receptor gene and the DM1 locus (27), and the MJD1 gene (28). This observation also holds for subthreshold CGG\textsuperscript{CCG} repeats at the fragile X locus (29,30), suggesting that contractions help avoid attainment of the threshold. Second, a recent study found that HD transgenic animals produce somatic TNR deletions during development, dependent on cell proliferation, leading the authors to suggest that contractions may precede expansions in some circumstances (31). Third, several lines of evidence also suggest that at least some cellular factors are different in expansions versus contractions (31,32). If so, the contraction process might be exploited as a means to selectively shorten TNR alleles.

MATERIALS AND METHODS

Cell lines and cell culture

SVG-A cells, an astrocytic subclone of the astroglial SVG cell line (33,34), were obtained as a generous gift from Dr Ashok Chauhan, Johns Hopkins University. These cells are immortalized with SV40 and express replication-competent SV40 Large T Antigen. Cells were grown adherently in DMEM (Mediatech, Inc.) supplemented with 10% fetal bovine serum, 50 U/ml Penicillin/Streptomycin and 2.5 µg/ml Amphotericin B (all from Invitrogen Corporation).

Plasmids and cloning

All shuttle vectors (Figure 1A and B) were produced as described previously (9). Briefly, these plasmids are derivatives of pRS313 (35). The presence of the SV40 viral origin of replication (ori) enables plasmid replication in SVG-A cells. Yeast markers include a selectable promoter-TNR-URA3 reporter construct for identification of TNR contractions (9). The ARS/CEN sequence enables low-copy plasmid replication and segregation in yeast while the HIS3 gene allows selection for plasmid molecules independent of TNR length.

Two different shuttle vector backbones were used in this study. The first backbone, pBL67 (Figure 1A), contains the SV40 origin of replication and large T antigen gene cloned into the BamHI site of pRS313. The promoter-TNR-URA3 cassettes were transferred into the vector backbone in two steps. First, complementary oligonucleotides containing TNRs and SpH\textsuperscript{i} sticky ends were annealed and cloned into the SpH\textsuperscript{i} site of pBL94 (16,36). This produced promoter-TNR-URA3 constructs which were then digested with EcoRI and Xho\textsuperscript{i} (New England Biolabs, Inc.) and transferred en bloc to the EcoRI/Xho\textsuperscript{i} sites of pBL67 (9). The second vector used here is pBL245 (Figure 1B). This backbone contains only the SV40 origin of replication (and not the large T antigen gene), cloned in the Apal site of pRS313 to create pBL185 (9). SVG-A cells constitutively express SV40 large T antigen, allowing the elimination of large T sequences from the pBL67 vector. Another advantage of this approach was that it eliminated an SpH\textsuperscript{i} site from the vector backbone that was present in the large T antigen gene, making unique the SpH\textsuperscript{i} site in the promoter-URA3 cassette. The empty promoter-Sph\textsuperscript{i}-URA3 cassette, containing no TNRs, was transferred into pBL185 after digestion with the EcoRI and BamHI, producing pBL245. SpH\textsuperscript{i} digestion (New England Biolabs, Inc.) of pBL245 enabled annealed TNR oligonucleotides containing SpH\textsuperscript{i} sticky ends to be cloned into the shuttle vector backbone directly. Since cloning the SV40 origin of replication into the Apal site of pRS313 resulted in its placement on the opposite side of the repeat tract versus pBL67, a standard system of nomenclature was adopted to ensure consistency between experiments. In all experiments, the named repeat, such as (CAG)\textsuperscript{m}, refers to a plasmid molecule where CAG repeats comprise the template for lagging strand synthesis. As a further precaution, comparisons between repeats of different sequence, length or orientation were only made between plasmids derived from the same vector backbone.
Shuttle vector assay

The shuttle vector assay (Figure 1C) was slightly modified from the previous study (9). Briefly, the day before transfection ~2.5 × 10^5 SVG-A cells were seeded in triplicate in 60 mm tissue culture dishes. The next day, cells were transfected using Lipofectamine reagent (Invitrogen Corporation), according to the manufacturer’s protocol using 5 μg of shuttle vector DNA and 20 μl of Lipofectamine per dish. After transfection, cells were incubated for 72 h to enable plasmid replication. The cells were lysed and plasmid DNA was extracted using Hirt’s technique (37). Lysates were concentrated by centrifugation through a Centricon 50 kDa MWCO filter (Millipore Corporation). Plasmid DNA was precipitated with ethanol and digested with DpnI (New England Biolabs, Inc.) DpnI digestion eliminates any plasmid molecules that did not replicate in human cells. DpnI-resistant DNA was then transformed into yeast strain S1502B (MATα leu2-3, leu2-112 his3-D trp1-289 ura3-52) using the lithium acetate method (38). A fraction of each transformation mixture (typically 0.5%) was plated onto SC-His plates (synthetic complete, lacking histidine), and the remainder onto SC-His-Ura (synthetic complete, lacking histidine and uracil) to score for contractions. Colonies on each plate were counted after 3 days of growth at 30°C. The frequency of contraction was determined as the number of colonies obtained on SC-His-Ura divided by the total number of transformants on SC-His, with appropriate correction for dilution factors. Statistical analyses were performed by applying a two-tailed Student’s t-test with equal variance to the relevant datasets. Comparisons between datasets observed in SVG-A cells were made with no correction for background. Full details of the assay results are given in Table 1.

A critical control for this assay involves determination of the background contraction frequency, defined as the percentage of contracted TNR tracts in the plasmid population that did not occur in human cells. The background was determined by transforming stock plasmid directly into yeast, bypassing replication in SVG-A cells (Figure 1C). The background provides quantitation of any TNR contractions that occur in Enterobacteriaceae during preparation stock.
plasmid, or in yeast before selection became effective. Background frequencies were calculated from at least three independent measurements.

Identification of TNR contractions

Yeast was used as a biosensor (9) for contractions arising in SVG-A cells (Figure 1D). Briefly, gene expression driven by the Schizosaccharomyces pombe adh1 promoter is dependent on the proper spacing between the TATA box and the downstream transcription initiation site (labeled ‘I’ in the figure). When a TNR-containing 33 triplet repeats is inserted between the TATA box and ‘I’, yeast transcription initiates upstream of ‘I’ and this causes an out-of-frame ATG codon to be encoded and a nonsense URa3 gene product to be produced. However, if a contraction to a TNR length ≤28 repeats occurs in human cells, upon transfer to yeast transcription will initiate at ‘I’, translation begins at the URa3 ATG codon, and the cells become phenotypically Urα. Thus Urα yeast colonies reflect contraction events that occurred in SVG-A cells. Note that in some plasmids, the initial TNR tract was <33 repeats. In these instances, genetically inert, randomized sequences were appended to make the total DNA length 99 nt.

**PCR analysis**

To authenticate contractions and to determine their size, individual Urα yeast colonies were resuspended in water and heated to 95°C for 10 min. Subsequent PCR amplification

### Table 1. TNR contraction data in SVG-A cells

<table>
<thead>
<tr>
<th>1</th>
<th>Repeat sequence</th>
<th>2</th>
<th>Plasmid name</th>
<th>3</th>
<th>Vector backbone</th>
<th>4</th>
<th>SVG-A contractions observed (His+ Ura+) colonies</th>
<th>5</th>
<th>Plasmid population (% of population × dilution)</th>
<th>6</th>
<th>N</th>
<th>7</th>
<th>SVG-A contraction frequency (±SEM) (%)</th>
<th>8</th>
<th>Background contraction frequency (±SEM) (%)</th>
<th>9</th>
<th>Corrected contraction frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>(C,A,G)33</td>
<td>pBL200</td>
<td>pBL67</td>
<td>0</td>
<td>6000</td>
<td>3</td>
<td>0 (±0)</td>
<td>0 (0±0)</td>
<td>0 (0±0)</td>
<td>0</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>(CAG)15±18</td>
<td>pBL213</td>
<td>pBL67</td>
<td>24</td>
<td>392400</td>
<td>3</td>
<td>0.0079 (±0.0039)</td>
<td>0.0073 (±0.0036)</td>
<td>0.0002</td>
<td>0.19</td>
<td>0.0004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>(CAG)25±8</td>
<td>pBL176</td>
<td>pBL67</td>
<td>2923</td>
<td>789000</td>
<td>5</td>
<td>0.34 (±0.09)</td>
<td>0.16 (±0.02)</td>
<td>0.19</td>
<td>0.0004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>(CTG)25±8</td>
<td>pBL207</td>
<td>pBL179</td>
<td>10</td>
<td>29800</td>
<td>3</td>
<td>0.04 (±0.02)</td>
<td>0.02 (±0.04)</td>
<td>0.02</td>
<td>0.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>(CAG)30±3</td>
<td>pBL227</td>
<td>pBL67</td>
<td>204</td>
<td>70800</td>
<td>6</td>
<td>0.82 (±0.28)</td>
<td>0.43 (±0.12)</td>
<td>0.39</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>(CAG)33</td>
<td>pBL217</td>
<td>pBL67</td>
<td>622</td>
<td>37600</td>
<td>4</td>
<td>2.3 (±0.4)</td>
<td>0.31 (±0.11)</td>
<td>1.9</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>(CTG)25±8</td>
<td>pBL199</td>
<td>pBL67</td>
<td>8</td>
<td>22200</td>
<td>2</td>
<td>0.03 (±0.006)</td>
<td>0.015 (±0.009)</td>
<td>0.019</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>(CTG)50</td>
<td>pBL175</td>
<td>pBL67</td>
<td>37</td>
<td>12800</td>
<td>3</td>
<td>0.34 (±0.06)</td>
<td>0.055 (±0.014)</td>
<td>0.24</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>(CAG)25±8</td>
<td>pBL245</td>
<td>pBL245</td>
<td>49</td>
<td>34600</td>
<td>4</td>
<td>0.17 (±0.01)</td>
<td>0.003 (±0.002)</td>
<td>0.17</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>(TAG)25±8</td>
<td>pBL246</td>
<td>pBL245</td>
<td>4</td>
<td>6000</td>
<td>2</td>
<td>0.067 (±0.009)</td>
<td>0.013 (±0.013)</td>
<td>0.054</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>(CTA)50</td>
<td>pBL216</td>
<td>pBL67</td>
<td>173</td>
<td>274000</td>
<td>5</td>
<td>0.09 (±0.05)</td>
<td>0.06 (±0.01)</td>
<td>0.03</td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>(CTG)6(AGT)2(CTG)25</td>
<td>pBL248</td>
<td>pBL245</td>
<td>576</td>
<td>60000</td>
<td>6</td>
<td>0.87 (±0.07)</td>
<td>0.051 (±0.004)</td>
<td>0.82</td>
<td>0.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>(CAG)33</td>
<td>pBL249</td>
<td>pBL245</td>
<td>50</td>
<td>12800</td>
<td>3</td>
<td>0.39 (±0.01)</td>
<td>0.12 (±0.02)</td>
<td>0.27</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>(CAG)33</td>
<td>pBL247</td>
<td>pBL245</td>
<td>2421</td>
<td>118400</td>
<td>3</td>
<td>2.0 (±0.5)</td>
<td>0.14 (±0.01)</td>
<td>1.9</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Column 1 lists the trinucleotide repeat sequence comprising the template for lagging strand synthesis (as defined in Materials and Methods). Column 2 gives each plasmid’s name for reference. Column 3 indicates the vector backbone in which each TNR sequence was cloned. For (CAG)25±8 and (CAG)33, the repeats were cloned in both vector backbones—the pBL67 derivatives (pBL176 and pBL217) were used in the threshold study while the pBL245 vectors (pBL250 and pBL247) were used in the hairpin and orientation dependence studies. The pBL67 derivative, pBL217, was also used as the (CAG)33 plasmid vector in the orientation dependence study. Column 4 indicates the total number of contractions observed in triplicate transfections in SVG-A cells (the number of colonies counted on SC-His-Ura plates). Column 5 lists the total number of plasmids in the population (the number of colonies counted on SC-His plates multiplied by the appropriate dilution factor). Column 6 lists the number of samples recovered after transfection. Column 7 lists the contraction frequency observed for plasmids replicated in SVG-A cells (%) ±1 SEM (%). The corrected contraction frequency is used to produce the graphs shown in the figures. Column 10 lists the P-value for each sequence versus the background as determined by two-tailed student’s t-test.

**Restiction analysis of contractions of interrupted TNR alleles**

Interrupted TNR alleles were created by introducing tandem ATG interruptions in a CTG repeat tract, resulting in a TNR tract with the sequence (CTG)6(AGT)2(CTG)25. The presence of either one or two ATG interruptions in a CTG repeat sequence introduces a recognition site for the restriction enzyme SfaNI (New England Biolabs). After passage through SVG-A cells and transformation into yeast, PCR products from contracted interrupted and perfect TNR alleles were isolated by the QIAquick PCR Purification Kit (QIAGEN Inc.) and digested with SfaNI. The presence or absence of the interruptions (as determined by SfaNI sensitivity) was visualized after separation on a 6% sequencing gel.

**Immunofluorescent detection of GFAP**

The day before staining, 10^5 SVG-A cells were seeded in each well of a 24-well tissue culture plate. Cells were washed with PBS and fixed in ice-cold acetone/methanol. After fixation, cells were permeabilized and blocked with Triton-X 100/BSA for 20 min. Half of the wells were treated overnight with a 1:500 dilution of mouse anti-human GFAP.
(Dako Denmark A/S) in PBS and the other half treated with PBS alone. After staining, cells were washed with PBS and treated with a 1:100 dilution of fluorescein-tagged goat-anti-mouse IgG (Invitrogen Corporation) in PBS in the dark for 30 min. Cells were washed again with PBS and stained for 2 min with a 1:1000 dilution of DAPI (Sigma-Aldrich) in sterile deionized H2O. Cells were visualized under an inverted fluorescence microscope through either DAPI or fluorescein specific filters and the images merged. The primary and secondary antibodies as well as the DAPI reagent were provided as a gift from Dr Anuja Ghorpade, University Nebraska Medical Center.

RESULTS

In previous work (9), we developed a shuttle vector assay that quantitatively measures CAG•CTG repeat contractions (Figure 1). The assay allows facile analysis of contraction frequency as a function of key cis-elements known from human TNR genetics, such as the repeat length, the direction of DNA replication, the structure-forming capacity and the purity of the repeat. This assay utilizes shuttle vector plasmids (Figure 1A and B) that are passed through SVG-A astrocytes and transferred into yeast, where genetic selection reveals the frequency of TNR contractions occurring during replication in cell culture. An outline of the assay is shown in Figure 1C and the mechanism of the yeast selection process is shown in Figure 1D. Yeast colonies, harboring a contracted TNR-containing shuttle vector plasmid, are lysed and the repeat tract is amplified with primers flankling the repeat to measure contraction size (Materials and Methods). The shuttle vector assay enables quantitative measurement of contraction frequencies (9), and is, to our knowledge, the only cell culture assay with suitable sensitivity to readily monitor TNR instability for alleles near the crucial threshold length of 30–40 repeats.

Replication and CAG•CTG repeat contractions occur readily in SVG-A cells

To confirm the CNS nature of the SVG-A cells under study, cells were stained for expression of glial fibrillary acidic protein (GFAP), an intermediate filament restricted to glial cells like astrocytes. Figure 2A shows that SVG-A cells express GFAP. More than 1000 cells were examined by fluorescence microscopy and 100% stained positively for GFAP. Control experiments (data not shown) proved that positive staining required the anti-GFAP primary antibody. Additionally, the GFAP stain in Figure 2A demonstrates a morphology consistent with that observed in another study of cultured astrocytes maintained in serum-supplemented media (39). Together, these findings support the astrocytic nature of SVG-A cells used here.

Since the role of replication in TNR mutagenesis is controversial, it is of particular importance to evaluate replication in cultured astrocytes. Shuttle vectors (Figure 1A and B) are predicted to replicate in SVG-A cells due to expression of large T antigen plus the presence of the SV40 ori. The SV40 system is a reasonable model because, once the origin is unwound, viral replication occurs with high fidelity using the host-cell replication machinery (40), suggesting our

plasmid system will likely mimic replication events on chromosomal DNA. Furthermore, incubation with nuclear extracts causes SV40-derived plasmids to fold into chromatin structures in vitro, and this nucleosome folding is enhanced after plasmid replication (41). Following transfection and recovery from SVG-A cells, replication was assessed by challenging DNA with the dam methylation-sensitive restriction enzyme DpnI. DNA that has replicated at least once in SVG-A cells will be resistant to digestion. A representative DpnI analysis is shown in the left panel Figure 2B. Lanes 1–2 show that DpnI digestion proceeds to completion for background determinations (lane 1 = uncut, lane 2 = cut) and plasmid DNA that has been passed through SVG-A cells (lane 3 = uncut, lanes 4–6 = cut). Sample wells between lanes 2 and 3 and lanes 3 and 4 were intentionally left empty. The right panel shows the analysis of individual contraction events from SVG-A cells. PCR amplification of starting plasmid DNA indicates the size of the 33-repeat parental allele (lane 7). Individual contractions are shown in lanes 8–13. Fragment sizes, stated as repeat units, were deduced by comparison with molecular weight standards (data not shown).

Figure 2. Analysis of SVG-A cells and transfected TNR plasmid DNA. (A) SVG-A cells stained with anti-GFAP antibodies (green false color) and counterstained with DAPI (blue false color) at 400× magnification. (B) DpnI resistance assay after recovery of plasmid pBL248 from SVG-A cells and subsequent PCR analysis of contracted alleles after yeast selection. The left panel shows the result of DpnI digestion of stock plasmid used for background determinations (lane 1 = uncut, lane 2 = cut) and plasmid DNA that has been passed through SVG-A cells (lane 3 = uncut, lanes 4–6 = cut). Sample wells between lanes 2 and 3 and lanes 3 and 4 were intentionally left empty. The right panel shows the analysis of individual contraction events from SVG-A cells. PCR amplification of starting plasmid DNA indicates the size of the 33-repeat parental allele (lane 7). Individual contractions are shown in lanes 8–13. Fragment sizes, stated as repeat units, were deduced by comparison with molecular weight standards (data not shown).
plasmid DNA, consistent with substantial levels of DNA replication. DpnI-digested plasmid preparations were transformed into yeast and putative CAG repeat contractions were identified as Ura⁺ yeast colonies (Figure 1C and D). Contractions were confirmed by PCR amplification of the CAG repeat locus, as shown in Figure 2B, right panel. The starting plasmid gave a PCR product consistent with 33 repeats, whereas plasmid DNA from the Ura⁺ colonies yielded shorter PCR products corresponding to ultimate repeat tracts of 11–25 repeats (contractions of –8 to –22). Together, these results indicate that the glial SVG-A cells support shuttle vector replication and show instability of CAG•CTG repeat tracts.

Length and orientation dependence of CAG•CTG contractions in SVG-A cells

A signature characteristic of TNR instability in humans is a sharp increase in mutagenesis as the tract reaches or exceeds a threshold of ~35 repeats. Consistent with the literature, we define a threshold as a demarcation between genetically stable and unstable alleles over a narrow range of repeat length. We sought to evaluate a potential threshold in astrocytic SVG-A cells. To accomplish this, CAG alleles of 0–33 repeats were created as shown in Figure 3A. Non-repeating sequence was used to normalize the length of all CAG tracts to 99 nt. Shuttle vectors harboring these repeat constructs were passaged through SVG-A cells and subsequent analysis in yeast produced the plot shown in Figure 3B. The frequency of CAG contractions was indistinguishable from background at 0 or 15 repeats (Table 1, entry 9B), but then rose sharply between 25 and 33 repeats. For the longest reporter, ~2% of all the CAG alleles underwent contraction during a transient three-day passage in SVG-A cells. Furthermore, lengthening the CAG tract from 25 to 33 repeats (+32%) resulted in a 1000% increase in contraction frequency, consistent with a threshold at or near 35 repeats. The observed contraction frequency at 30 repeats did not achieve statistical significance over background, but the 25 and 33 repeat results were clearly distinguishable from background (Table 1, entries 10C, E and F). Thus CAG contraction frequencies suggest a threshold-like response in SVG-A cells.

Given that the distance of the SV40 origin from the CAG•CTG tract is known to modulate repeat instability in primate cells (11), the position of the appended random sequence [e.g. in (CAG)₂₅₋₆₄] is a potential complicating factor. However, the data in Table 1 show that (CAG)₂₅, whether tested in pBL67 or pBL245, had indistinguishable mutation frequencies of 0.19 and 0.17%. The results for (CAG)₁₅ were also consistently ~10-fold higher in the two plasmids (1.9% for both). These data indicate that the appended random sequence does not affect the results, at least for these two locations relative to the SV40 origin.

Replication-dependent TNR contractions in microbes often exhibit an orientation effect, where contractions are more frequent when CTG occupies the lagging template strand rather than the complementary sequence CAG. In a prior study using the shuttle vector assay, the opposite was true in human 293T and 293 kidney cells; contractions were ~10-fold more frequent when CAG comprised the lagging strand template (9). This prompted us to examine orientation effects in SVG-A cells, as depicted schematically in Figure 4A. The only difference in the two plasmids (left and middle panels in Figure 4A) is the cloning orientation of the CAG•CTG tract. To add to the findings in our previous study (9), we examined the orientation dependence of (CAG•CTG)₁₅ alleles, in addition to alleles containing 25 repeats. The result in Figure 4B indicates that (CAG)₁₅ sequences contract ~7-fold more often than (CTG)₁₅ and the orientation effect was statistically significant ($P = 0.02$). Moreover, we took advantage of another series of shuttle vectors—with the SV40 ori relocated to the opposite side of the repeat tract—to test whether switching the direction of replication through otherwise identical repeats
The plasmid, called (CAG)$_{33}$–new ori, harbored repeats in the same sequence context as the (CTG)$_{33}$ vector, but movement of the replication origin resulted in fork movement through the repeat in the opposite direction, such that CAG repeats occupied the lagging template strand. This effect is depicted pictorially in the middle and right panels of Figure 4A. Reassignment of CAG repeats to the lagging template resulted in full restoration of instability in the (CAG)$_{33}$–new ori vector (Figure 4B and Table 1 entry 9F). This restoration suggests that the orientation of the repeat with respect to replication is the most important factor, not the distance from the origin nor flanking DNA sequences. Furthermore, restoration also held true for (CAG)$_{25}$–8 versus (CTG)$_{25}$–8 repeat tracts (Table 1, entries 9C, 9G and 9I). To address explicitly the role of DNA replication in producing CAG•CTG contractions in SVG-A cells, the contraction frequency was indistinguishable from background (Table 1, pBL207, entries 9D and 10D), confirming that astrocytes experience replication-mediated TNR instability in this model system, with similar orientation effects as kidney-derived cells (9).

The relatively low frequency of (CTG)$_{25}$–8 and (CTG)$_{33}$ contractions raised the issue of whether CTG repeats require longer lengths to be mutable, or whether—for some unknown reason—they are largely immutable in our system. This question was answered by increasing the tract length to (CTG)$_{50}$. Figure 4C shows that the contraction frequency for (CTG)$_{50}$ increased significantly compared to (CTG)$_{25}$–8 (Table 1, entries 9H and 9G). The absolute frequency for the (CTG)$_{50}$ reporter approaches the level seen for (CAG)$_{25}$–8, and is significantly above background ($P = 0.002$). However, the observed (CTG)$_{50}$ contraction frequency likely underestimates the true contraction frequency in SVG-A cells because the 50-repeat allele must delete at least 22 repeats to register as Ura$^+$ in the yeast selection step of our assay.

Figure 4. The repeat orientation relative to the direction of replication influences CAG•CTG contractions in SVG-A cells. (A) Complementary CAG and CTG repeats can differentially affect TNR instability based on the orientation of the repeat relative to the direction of replication. In the left panel, CAG repeats reside on the lagging strand template. In the middle panel, the reverse orientation places the CTG repeats on the lagging template strand. On the right, switching the position of the SV40 ori reverses the direction of replication and swaps the assignment of leading and lagging strand. The direction of replication was deduced from the position of the S40 ori relative to the TNR. (B) Contraction frequency (%) is plotted as given in Table 1 (column 9) for the indicated starting alleles. Error bars indicate ±1 SEM. (C) Contraction frequency (%) is plotted versus CTG repeat length. All values in (B and C) represent the corrected contraction frequency in SVG-A cells.
(Figure 1D). Therefore our assay does not register contrac-
tions of <22 repeats from the (CTG)30 reporter. By compari-
sion, a 25 + 8 allele must lose only five repeats to score as a
contraction in our assay. Nonetheless, we can conclude that
CTG runs are mutable in SVG-A cells in a length-dependent
manner (P = 0.04), but they require a longer tract than CAG
repeats to achieve similar contraction frequencies.

Structure-forming capacity promotes TNR contractions
in SVG-A cells

One of the seminal observations in the field of TNR muta-
genesis was that disease-causing repeat sequences have the
ability to form aberrant DNA secondary structures (42–44).
This observation led to models in which TNR instability,
either expansions or contractions, results from error-prone
replication, recombination or repair of DNA structures that
form at TNR loci. To determine whether structure-forming
capacity influences the propensity for TNR contractions in
SVG-A cells, we created shuttle vectors harboring either a
(CAG)25+8 allele, which is predicted to form a stable hairpin
under physiological conditions, or a (TAG)25+8 allele, which
is not predicted to form stable secondary structure. Even with
allele lengths below the putative threshold, hairpin-forming
capacity stimulated contractions by ~3-fold (P = 0.04;
Table 1, entries 9I and 9J). The (CAG)25+8 allele contracted
at a frequency that is statistically significantly above back-
ground (P = 0.006), while the (TAG)25+8 contraction frequency
compared to background is only borderline signifi-
cant (P = 0.05; entries 10I and J in Table 1). Hairpin-forming
capacity also appears to be important for CAG•CTG contrac-
tions arising from longer 50-repeat alleles. The (CTG)30
run contracted more frequently than hairpin-incompatible
(CTA)30 alleles (Table 1, entries 9H and 9K, P = 0.002).
From these studies, it appears that TNR sequence compatible
with hairpin formation is critical for promoting maximal TNR
instability in CNS cells.

CTG repeat contractions arise from
an interrupted allele

In humans, interrupted alleles harboring 1–3 bp substitutions
do not expand, whereas perfect alleles do. The inference is
that interruptions inhibit expansions, and that the interrupting
nucleotide(s) must be mutated or deleted to yield a perfect,
expandable repeat. However there is little or no contraction
data available in human cells for interrupted alleles. We
used the shuttle vector assay to investigate the contractibility
of an interrupted allele in cultured human astrocytes.
Shuttle vectors were created with an interrupted repeat,
(CTG)6(ATG)2(CTG)25, and a perfect (CTG)33 repeat as a
control. After passage of interrupted and perfect repeat harboring shuttle vectors through SVG-A cells, there was a
3-fold increase in contraction frequency for interrupted repeat
alleles compared to perfect repeat controls (Figure 5A).
The contraction frequency for the interrupted allele was sta-
tistically significant (P = 0.0001; entry 10L in Table 1) while
the perfect (CTG)33 contraction frequency was borderline
significant (P = 0.05; entry 10M in Table 1). The difference
in contraction frequency between the interrupted allele
and perfect allele, depicted in Figure 5A, was statistically
significant (P = 0.005). Thus the interruption in this particu-
lar allele resulted in more contractions, not fewer.

Though contractions occurred more frequently with the
interrupted allele, PCR analysis (Figure 5B) revealed that
the median contraction size was smaller for an interrupted
allele (median contraction size ~16 repeats) compared to
the perfect repeat control (median contraction size ~20
repeats). Consistent with a smaller median contraction size,
the spectrum of contraction mutations for the interrupted
allele is skewed towards smaller contractions, and the two
spectra are statistically different by two-tailed Student’s
t-test (P < 0.0001). Retention or loss of the interruptions
was deduced by treating PCR-amplified repeat tracts from
contracted alleles with the restriction enzyme SfaNI.
Contracted alleles that retain at least one ATG interruption
will be sensitive to SfaNI, but deletions that remove both
interruptions become enzyme-resistant. A representative gel
obtained from this assay is shown in Figure 5C. SfaNI sensi-
tivity was retained in 14 out of 20 contractions tested, sug-

gestions that disease-causing repeat sequences have the

a 25 repeat block, and none of the contractions extended

from the SfaNI sensitive and resistant categories. All
contracted alleles that retain at least one ATG interruption
were smaller and more frequent than a perfect repeat control,
and the interruptions were usually retained.

DISCUSSION

The cells of the CNS are among the most prone to the devas-
tating consequences of TNR expansion, and some CNS cells,
including astrocytes, show substantial TNR instability in
affected individuals. Furthermore, astrocytes are enriched in
brain regions that are sensitive to neurodegeneration and
somatic instability in TNR diseases (45,46). Based on this
biological relevance, we used immortalized SVG-A astro-
cytes as an ex vivo model to mimic TNR mutagenesis. The
culture model takes into account the proliferative capacity
of astrocytes, and therefore enables us to address questions
regarding replication-dependent sources of TNR instability
in an important CNS cell type. Our results show that astro-
cytes in culture have substantial levels of CAG•CTG repeat
contractions, and the genetic elements controlling this instabil-
ity mimic signature features from human genetics. Together
these findings suggest that replication-based models of TNR
contractions are relevant to CNS cells with proliferative
capacity, such as astrocytes, and that the molecular mecha-
nisms of instability are likely retained in these cultured
cells. To our knowledge, this study comprises the first genetic
assay for TNR instability performed in cells derived from
the human CNS. The data also suggest that the mechanistic
features that govern TNR instability are consistent between
CNS cells and other tissue types, such as the kidney (9),
that show somatic instability in DM1.
A genetic hallmark of TNR disease is sequence specificity, such that sequences capable of aberrant secondary structure in vitro correlates strongly with instability and disease in human patients (42). In cultured astrocytes, hairpin-forming capacity was critical for producing statistically distinguishable TNR contractions. Furthermore, the unique sensitivity and flexibility of the shuttle vector assay enabled us to determine that this holds true in CNS cells for both short and disease-causing CAG•CTG alleles. Another signature characteristic of TNR instability in humans is the threshold, where mutagenesis increases strongly over a narrow range of repeat lengths, typically 30–40 repeats (1). SVG-A cells exhibit a striking 1000% increase in CAG tract contraction frequency between 25 and 33 repeats, a length change of only 32%. CTG sequences also showed a strong increase in contraction frequencies between 25 and 50 repeats. This is the first genetic evidence that TNR contractions display a threshold-like behavior in cells derived from the human nervous system. A similar result was reported earlier for COS-1 cells (9). Moreover, our observations are consistent with a contraction threshold in cultured astrocytes that is similar to the clinically relevant germline expansion threshold of ~35 repeats (1). Perhaps as TNR length approaches the threshold in CNS cells, contractions can occur to reduce the allele size back to a more stable length and thereby reduce the risk of disease-causing expansions. We note that our astrocyte system models somatic instability moreso than germline events, and that much less is known about genetic thresholds in somatic tissues. Nonetheless the contraction frequency curve (Figure 3B) strongly suggests a threshold-like response.

The requirement for the SV40 origin of replication for contractions and the orientation effect observed during replication of the shuttle vector in SVG-A cells suggests that CAG•CTG contractions in this system are replication-dependent. This is important in light of the fact that astrocytes and other glia retain proliferative capacity throughout life, unlike terminally differentiated neurons. Therefore, replication could be a key source of somatic instability in glial cells. One indication of the important role of somatic instability in the disease process is a correlation between the amount of somatic instability in specific brain regions and susceptibility to neurodegeneration in HD (46). Since expression of expanded huntingtin in cultured astrocytes decreased...
their protection of co-cultured neurons against excitotoxicity (25), then replication-mediated somatic instability in glial cells could hasten neurodegeneration in susceptible brain regions.

In our astrocyte model, we observed a ~10-fold orientation effect such that TNR contractions were strongly favored when CAG repeats form the template for lagging strand synthesis. The magnitude of this effect was constant for allele lengths of both 25 and 33 repeats. Similar findings were observed in 293T and 293 cells, using the same shuttle vector assay (9). However these results are somewhat surprising because contractions in most model systems are favored when CTG runs populate the template strand (13,15,16,47). Perhaps differential binding by trans-acting factors might impart greater contractibility to repeats that are replicated in the CAG orientation. For example, an unknown protein might bind preferentially to the CAG hairpin and stabilize it. Alternatively, RPA or another protein might bind preferentially to single-stranded CTG repeats and reduce their ability to form a hairpin. Whatever the mechanism, our findings highlight this unique feature of TNR instability in human cells that is not observed in yeast or E.coli.

We found that an interrupted CTG repeat allele is approximately three times more contraction prone than a perfect repeat control. Although our observations run contrary to the current belief that interruptions act in a wholesale fashion to prevent TNR instability, it is important to note that ours is the first study of the effect of interruptions on TNR contractions in a human system. In yeast, genetic studies indicated that interruptions prevent contractions primarily by destabilizing putative hairpin intermediates (48), yet our data implies a different mechanism. Since the majority (70%) of contractions from interrupted TNRs retained at least one of the interruptions, it suggests that the interruptions are excluded from any mutagenic hairpins, but somehow promote contraction in the nearby DNA. The 30% of contractions that delete interruptions also suggest that creation of a shortened but perfect TNR allele via a contraction also represents a potentially significant event. Consistent with this idea, when disease-free alleles at the SCA1 locus were analyzed, 98% contained interruptions, but the normal alleles that lacked interruptions were among the shortest observed (49), perhaps the byproduct of a TNR contraction. If interruptions indeed promote contraction of TNR alleles, this would further highlight their critical importance in maintaining short, expansion-averse TNRs in the human genome.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Dr Raisa Persidsky and Dr Anuja Ghorpade for assistance in immunostaining of SVG-A cells, and Dr Richard Pelletier for instruction (to B.T.F.) in the procedures for the shuttle vector assay and for cloning of some of the shuttle vectors used here. This work was supported by National Institutes of Health grant GM61961 (to R.S.L.), by a graduate fellowship from the University of Nebraska Medical Center and National Institutes of Health fellowship F30 NS046135 (to B.T.F.), and by National Cancer Institute Cancer Center Support Grant P30 CA36727 (to the Eppley Institute). Funding to pay the Open Access publication charges for this article was provided by NIH GM61961.

Conflict of interest statement. None declared.

REFERENCES


