Microsatellite or Simple Sequence Repeat (SSR) instability depends on repeat characteristics during replication and repair

Om Prakash CHOUDHARY and Seema TRIVEDI*

Computational Biology Lab., Department of Zoology, JN Vyas University, Jodhpur (Raj.), India
(*author for correspondence; svtrivedi@hotmail.com)

Received: 29 June 2010; Accepted: 23 December 2010

Abstract

Microsatellites or simple sequence repeats (SSRs) are ubiquitous in genomes studied so far. SSR mutations or microsatellite instability (MSI) can result in either expansion or contraction of repeat units. These alterations can lead to disorders like cancer and neurodegeneration. As gleaned from studies, it appears that SSR mutations are dependent on motif size, nucleotide content and SSR length. Besides these, position of SSRs in genome (coding or non-coding sequences), presence on leading or lagging strand and distance from origin of replication also affect SSR mutations. Further, fidelity of replication and repair machinery as well as epigenetic factors may influence repeat stability. This review summarizes some of these aspects of SSR mutations that occur especially during DNA replication/repair besides epigenetic factors that influence SSR contraction or expansion.

Keywords: Simple sequence repeats, microsatellites, DNA polymerase, DNA repair, strand slippage

Mikrosatelit veya Basit Dizi Tekrar (SSR) kararsızlıkları replikasyon ve tamir sırasında tekrar karakteristiklerine bağlıdır

Özet


Anahtar Sözcükler: Basit dizi tekrarları, mikrosatelitler, DNA polimeraz, DNA tamiri, zincir kayması
Introduction

Microsatellites or Simple Sequence Repeats

SSRs or microsatellites are tandem repeats of 2-8 nt units of DNA and are ubiquitous in all genomes studied so far. SSR have varying density and motif distribution that may be species specific even in genomes with low SSR density. For example, among Eukaryotes, bivalves have one of the lowest SSR densities but mollusecan specific repeat density variations are reported (Cruz et al., 2005). Similarly, variations are seen in SSRS in mitochondrial DNA of Procyris rabaudi (Tchang) (an endemic fish species in China), (Zhang et al., 2009). These species specific variations can be explained on the basis of differences in repair systems (Pérez et al., 2005) or mutation rates.

SSRs have functional roles in gene regulation, chromatin modeling, recombination, evolution, development of new genes, evading host immune response in pathogens, adaptation, resistance to environmental stresses, binding of regulatory proteins to RNA structure by formation of hairpin loop, affect efficiency of exon splicing, protein functions etc. whether they are present in coding or non-coding or intragenic regions (Amador et al., 2004; Kashi and King, 2006; Mrázek, 2006; Sreenu et al., 2007; Coil et al., 2008). In primates, repeats like ATn could be involved in segmental duplications that can mediate evolutionary rearrangements (Kehrer-Sawatzki and Cooper, 2008). AT-rich repeats affect replication dynamics by possibly blocking replication, decrease the efficiency of nucleosome assembly and DNA supercoiling (Yamakoshi et al., 2005; Mrázek, 2006; Lukusa and Fryns, 2008). However, all SSRS may not function as contingency loci or recombination hot spots (Mrázek, 2006).

Despite several functions attributed to repeats, it is also known that mutations in SSRs can be source of disorders including cancer and neurodegeneration (Hancock et al., 2001; Hancock and Santibanez-Koref, 1998; Bacolla and Wells, 2009). SSRs have high mutation rates which are more in intergenic sequences compared to coding sequences. These mutation rates within a genome may differ depending not only on the motif types but also on length of repeats (Jacob and Eckert, 2007; Eckert and Hile, 2009). Repeat expansion or deletion can happen due to replication slippage which is considered as one of the main reasons for SSR mutations. The other reasons that cause repeat instability have been attributed to faulty repair/recombination processes. However, as gleaned from other studies, repeats also expand or contract independent of errors during replication, repair or recombination events. This review focuses on some of the reasons for SSR mutations that occur due to replication or repair process which may depend on not only the motif size but also the nucleotide composition of each motif as well as orientation of repeats or position with reference to replication origin. This review does not focus on details of mechanisms of SSR mutations.

Repeat motifs and lengths

SSR mutations are affected by nucleotide contents of the motifs, lengths of the motifs and total length of the repeats. Regarding motif lengths, studies so far have concentrated on trinucleotide repeat instability (mainly in exons) possibly because of their implication in disorders (Wang and Vasquez, 2009; Shah et al., 2010). However, other repeat types like di-, tetra- and penta-nucleotide can be unstable and result in diseases (Mirkin, 2004). Studies have shown that mutations are more frequent in repeats with larger motifs like tetra- and penta-nucleotide compared to shorter repeats like di- and tri-nucleotide repeats. However, this may at times depend on disease associated or non-associated repeat motifs (Chakraborty et al., 1997; Schug et al., 1998).

Nucleotide content of repeats is also an important factor in repeat stability as CG rich repeats show higher rates of mutation. Further, repeat instability is also length dependent as longer SSRs are more unstable compared to shorter repeats (Callahan et al., 2003; Wang and Vasquez, 2009). On the other hand some shorter repeat lengths undergo higher rates of mutation compared to longer repeats but this length dependence also has a limit (Paiva and Sheardy, 2004; Wells et al., 2005). However, repeat
expansion does not result in disorders until the increase in length exceeds threshold that falls within the range of 100–200 bps (Mirkin, 2004).

Nucleotide content and length of motifs as well as total SSR lengths are important due to their tendency to form bulges, non-B DNA or secondary structures. This is proved in studies that have found frequent misalignments in pyrimidine templates but guanine templates are less prone to strand slippage upon misincorporation (Chi and Lam, 2006; Chi and Lam, 2007; Chi and Lam, 2008). Paradoxically, it has also been reported that in guanine templates stable mismatches like G.G, T.G, and A.G are formed. Further, studies show that bulges due to purine are more destabilizing than pyrimidine bulges (Chi and Lam, 2009). Mispairing is enhanced by the complementary strand especially when dTTP, dATP, and dGTP are opposite 5’-AG, 5’-TG, and 5’-CG templates. Further, misalignment is seen only in 5’-CG templates with a downstream purine. However, mismatched structures are also formed at 5’-CG templates that allow formation of stable terminal G.C base pair. Nucleotide downstream of the template guanine also facilitates formation of strand slippage (Chi and Lam, 2009). These structures can affect DNA replication/repair and result in expansion or contraction of repeats (discussed in the following sections).

**Repeat position and mutation**

Mutations in SSRs depend on the position of repeats in genomes besides the aforementioned factors. Repeats in introns have higher mutation rates compared to repeats in exons (Shah et al., 2010). Moreover, instability is higher in repeats located near origin of replication besides their presence in lagging strand compared to leading strand (Schweitzer and Livingston, 1999; Callahan et al., 2003; Wells et al., 2005; Wang and Vasquez, 2009). In absence of the hairpin nuclease SbcC, if repeat length is less than the length of Okazaki fragment there is repeat expansion. On the other hand, if repeat length is equal to or greater than Okazaki fragment there is saltatory amplifications of repeats (Sarkar et al., 1998). However, there are cases where slippage is independent of orientation but deletions of repeats become orientation dependent during correction process. This is seen in *Escherichia coli* where CAG x CTG slippage is independent of orientation but deletion of repeat becomes orientation dependent during processing of this slipped strand. This is because CTG repeat containing strands can misfold and proofreading by subunit of DNA polymerase III (DnaQ) in the presence of SbcCD (Rad50/Mre11) is inefficient on such folds. Moreover, SbcCD processing is dependent on proofreading and repeat tract orientation. Further, repeat instability through this system is also dependent on nucleotide composition of SSRs (Sarkar et al., 1998; Connelly et al., 1999; Zahra et al., 2007). If GAA serves as the lagging strand template there is an increase in instability in RecA-deficient (deficient in components of the RecFOR and RecBCD pathways) but not RecA-proficient *E. coli* strains. It is possible that genetic stability of the (GAA*TTC)n may require efficient RecA-dependent recombination restart of stalled replication forks. It has also been found that transcription and recombination do not influence repeat instability in this system (Zahra et al., 2007). Orientation dependent replication fork stalling is also explained by experiment in *E. coli* where study shows that binding of two proteins Tus (inhibits helicase activity) and Ter (replication termination) result in a complex that disrupts DNA protein interaction of DnaB in an orientation-specific manner, thus blocking DnaB function (Mirkin and Mirkin, 2007). Distance between the origin of replication and repeat motifs like (CTG.CAG)n sequence may affect repeat instability in mammals where stability order of repeats is CGG > CTG > CAG > CCG (M Rindler et al., 2006). However, in case of GA or TC repeats replication stalling is not dependent on position of repeat near origin of replication in *E. coli*. Replication fork stalling in this case is due to TraY protein (required for bacterial conjugation) binding at these sequences (Krasilnikova et al., 2001). Therefore, it appears that formation of different non-B DNA structures depend on various factors including repeat type besides their position on leading or lagging strand as well as distance form origin of replication.
Non-B DNA structure or secondary structure

During replication DNA is unwrapped from nucleosome cores and becomes single stranded, providing possibilities of formation of non-B DNA structures, particularly in the lagging strand. Some of these non-B DNA structures are more stable than the others (Mirkin, 1999 and 2001; Sinden et al., 2002; Wells et al., 2005; Veytsman and Akhmadeyeva, 2006; Völker et al., 2008) and hence, increase the possibility of stalling of DNA polymerase (Callahan et al., 2003; Wang and Vasquez, 2009) or evoke repair mechanisms. Since non-B DNA structures are thermodynamically more stable, the replication and repair machinery or proteins that assist in rectifying this structure would require more energy to unfold (Völker et al., 2008). This causes polymerase stalling or signaling to repair mechanisms and can result in either expansion or contraction of SSR lengths. Alternatively, mistakes made by replication or repair machineries during recognition, binding, and/or processing of some or all of the non-B DNA structure substrates may facilitate expansion of the triplet repeat sequences (Völker et al., 2008). Non-histone proteins, DNA binding proteins like ribosomal barrier proteins, origin recognition complex, telomeric binding proteins etc. also cause replication fork stalling (Mirkin and Mirkin, 2004). Fork stalling due to non-B DNA structure can lead to template switching (not unequal crossing over) that results in repeat expansion (Cocquempot et al., 2009).

It is also known that repeats like (C-A)n,(T-G)n can lead to formation of Z-DNA structure under very high ionic strength, or methylation of all cytosines in the Z-forming repeat in linear DNA. There is un-pairing of complementary strands during the B-to-Z transition (Mirkin, 2001). This un-pairing can permit formation of loops either before or during replication process. In either case this would result in stalling of DNA polymerase and thus increasing the chances of repeat mutations.

However, not all non-B DNA structures cause expansion and instead they may be helpful in replication as proposed for H form DNA that may be required at origin of replication to help in replication initiation or may be involved in homologous recombination (Mirkin, 1999). Paradoxically, in vivo validation of hairpin structure formation responsible for SSR instability has been difficult (Zahra et al., 2007).

DNA polymerase and mutations

Errors during replication either due to aforementioned factors or inefficiency of the replication process can result in SSR mutations. Repeats expand particularly when repeat is on lagging strand and if the replication temperature is low, but repeats undergo deformation if replication temperature is high (Sarkar et al., 1998). Requirement of higher temperatures to resolve non-B DNA structures and association with repeat expansion supports this work (Völker et al., 2008).

DNA polymerase itself can be cause of SSR mutation which is seen when both polymerase kappa and alpha-primase pause at mononucleotide repeats (T)11 but repeat is interrupted by Polymerase kappa by inserting dGMP and dCMP within repeat (Hile and Eckert, 2008). The reason why these polymerases pause at these sequences is not known. Mutations in genes involved in DNA replication, e.g., ligase, primase, rad27 (thus defects in flap processing), FEN-1 or DNA polymerases (delta, DnaQ or epsilon), PCNA (proliferating cell nuclear antigen) can affect repeat stability (Sarkar et al., 1998; Bzymek et al., 1999; Schweitzer and Livingston, 1999; Iyer et al., 2000; Callahan et al., 2003; Wells et al., 2005; Wang and Vasquez, 2009). Mutation in dnaQ49 (proofreading function of DNA polymerase III) either alone or along with mutation in mutD5 enhance repeat instability (Saveson and Lovett, 1997; Bzymek et al., 1999; Iyer et al., 2000; Zahra et al., 2007). There may be competition between DNA ligase I and Fen1 during CTG flap processing in addition to displacement synthesis by polymerase delta that could increase the size of the flap (Callahan et al., 2003). It is proposed that simple replication misalignment events are sensitive to DNA polymerase III exonuclease, whereas sister chromosome exchange (SCE) associated events are sensitive to exonuclease I (Bzymek et al., 1999). If repair at this stage fails, post replication repair
mechanisms step in which may also affect repeat stability (Callahan et al., 2003).

**DNA Repair and SSR mutation**

Post-replication repair system and cell cycle checkpoints affect SSR abundance and distribution. Paradoxically, DNA repair process itself may be one of the causes of non-B DNA structure formation at or near the repetitive sequences (Wang and Vasquez, 2009). Trinucleotide repeats expansions in yeast activate the DNA damage response mediated by MEC1, RAD9, or RAD53, RAD17 or RAD24. However signaling down the Rad9 pathway and not the Rad17-Rad24 pathway plays a major role in sensing and repairing breaks in triplet repeats like CAG. Further, repeat contractions can increase due to Mec1, Ddc2, Rad17, Rad24, or Rad53. It has also been seen that repeat expansion of repeats like (CAG,CTG) is restricted by Rev1 (independent of dCMP transferase activity of pol zeta) and Srs2 if there is a 3’ slippage. This is done in association with post-replication repair system (Collins et al., 2007; Daee et al., 2007; Yang and Freudenberg, 2007). Anc1 (a member of seven multi-protein complexes involved in transcription) may also be involved in post-replication repair like SRS2 and RAD5. This protein may prevent expansion of repeats like CAG at least in yeast (Erlich et al., 2008). Therefore, both chromosomal integrity and repeat stability are maintained by components of the checkpoint machinery (Lahiri et al., 2004).

**Double stranded break and repair:** Double stranded breaks (DSBs) repair system through homology directed repair induce instability of (GAA,TTC)n or (CAG,CTG)n repeats (Pollard et al., 2007; Pollard et al., 2008; Mittelman et al., 2009; Wang and Vasquez, 2009). RecA and/or RecBC might play a role in DSB repair within the CTG,CAG repeats in an orientation-dependent manner in E. coli. If the DSBs are located within CTG or CGG repeats in RecA or RecBC-deficient E. coli there is an increase in deletions of the repeats (Hebert et al., 2004; Wang and Vasquez, 2009). However, expansions or deletions of repeats are not seen if the nick is not in centre of repeats (Hebert et al., 2004; Wang and Vasquez, 2009). Contrary to this view, other studies suggest that repair-mediated instability (particularly deletions) is dependent on the sequence of the triplet repeat but not on length and orientation of repeat (Pollard et al., 2007; Pollard et al., 2008). Mechanisms involving these triplet repeats in double stranded breaks and the repairs that may result in instability are explained by Hebert and Wells, 2005.

**Mismatch repair (MMR) system:** This system corrects mispaired or unpaired bases which were missed by the proofreading activity of DNA polymerases. MMR system deploys proteins that vary in number and nature in different organisms but basic aspects are conserved throughout evolution. MMR can be biased towards repair or maintaining stability of some sequences. For example, human postmeiotic segregation 2(hPMS2) proteins protect tetranucleotide expansions rather then deletions and displays a sequence bias, wherein (TTCC/AAGG) sequences are stabilized more than (TTTC/AAAG) (Shah and Eckert, 2009).

Rates of mutations increase in absence of MMR genes or defective genes or CpG islands hypermethylation (Spampinato et al., 2009). Further, DNA repair is specific for errors due to mononucleotide repeats. MutSa repairs only errors due to mononucleotide repeats, whereas MutSβ repairs mono-, di-, tri-, and tetra-nucleotide repeat mutations (Shah et al., 2010). It has also been found that ATPase domain mutation of MSH2 (mismatch repair protein) affects the formation of CTG expansions and leads instead to transmitted contractions (Tomé et al., 2009). However, MSI stability for both mono- and di-nucleotide sequences may be achieved due to involvement of hMRE11 through physical interaction with hMLH1 (Vo et al., 2005).

**Base/nucleotide excision repair:** SSR mutations can also be affected by the base or nucleotide excision repair. In E. coli, stability of (CTG/CAG)68 tracts is influenced by mutations in nucleotide excision repair (NER) genes. Further, absence of the uvrC or uvrD gene products enhance the instability of trinucleotide repeats, but lack of the functional UvrA or UvrB proteins causes substantial
stabilization of (CTG/CAG) tracts (Szwarocka et al., 2007). The stability of loops formed due to replication slippage can be altered by creation of abasic site near or within the loop especially in case of triplet repeats. This can result in alteration in base excision repair function that can result in repeat expansion (Völker et al., 2009).

At times SSR repair/mutation process may not have association with post-replication repair system. It is seen that slip-outs of (CAG)(20) or (CTG)(20) can undergo correct repair, escaped repair or error-prone repair depending on nick location and composition of CAG or CTG. Further, error-prone repair can incompletely excise excess repeats, which may generate expansions but not deletions. However, for this mismatch repair (MMR) and nucleotide excision repair (NER) proteins hMSH2, hMSH3, hMLH1, XPF, XPG or polymerase beta are not required. This study emphasizes differential processing of slipped repeats that may explain the differences in mutation patterns between various disease loci or tissues (Panigrahi et al., 2005). Mosaics of repeats in different cells (Takano et al., 1996) as explained by the mutational mode (Veytsman and Akhmadeyeva, 2006) support possibilities of such processes.

In non-dividing cells, knocking-out of the base excision repair enzyme, 7,8-dihydro-8-oxoguanine-DNA glycosylase (OGG1), results in stabilization of CAG repeats in mouse brains. This indicates role of BER in repeat expansion (Wang and Vasquez, 2009). But what happens to the damaged repeat in such cases? In addition to these processes, there are prokaryotes that lack post-replicative DNA repair enzymes like MutL, MutH and MutS (e.g. Mycobacterium tuberculosis, Helicobacter pylori, Campylobacter jejuni, Mycoplasma genitalium, and Mycop. pneumoniae) (Eisen and Hanawalt, 1999).

In Mycobacterium avium, M. leprae, M. bovis and the two strains of M. tuberculosis (CDC1551 and H37Rv) SSRs distribution is on an average of 220-230 SSR tracts per kb but lack long repeats (Sreenu et al., 2007). It is possible that in M. tuberculosis proteins coding context-specific codon choice partially compensates for the lack of a mismatch repair system. Hence genome integrity is not compromised in this pathogen (Wanner et al., 2008). In E. coli genome, despite higher rates of mutations, SSRs are short and not abundant. This low abundance has been attributed to absence of F-plasmid which acts independent of repair system pathway (Schlötterer et al., 2006). Further, in E. coli triplet repeats are stabilized by SOS repair but increases with increase in negative supercoiling of DNA (Majchrzak et al., 2006).

**SSRs in replication and repair genes**

Mutations in replication and repair genes can affect genomic fidelity. Chances of mutations in these genes may increase if the replication and repair genes have SSRs. Studies have shown presence of repeats not only in replication and repair genes of Mycoplasmas but also in higher eukaryotes ( Strand et., 1993; Chang et al., 2001; Loire et al., 2009; Guo and Mrázek, 2008; Trivedi, 2003 and 2010; Matsubara et al., 2000; Komatsu et al., 2000). Some of these genes related directly or indirectly to cell cycle are known causes of diseases especially cancer due to microsatellite instability (Table-1).

**Other mechanism besides replication and repair**

Repeats can also expand/contract by mechanisms that are independent of replication but can at times depend on repair process (Takano et al., 1996; Zahra et al., 2007). Such mechanisms operate in cells that seldom replicate like neurons in brains, muscles etc. (Takano et al., 1996; Zahra et al., 2007; Wang and Vasquez, 2009) and can happen despite replication and repair efficiency of Werner's syndrome helicase (WRN) and TRP53 genes or p53-mediated error correction mechanism. This is despite the fact that WRN and TRP53 facilitate replication of the FXS repeat and enhances Okazaki fragment processing. Therefore, it is possible that there are other mechanisms that are responsible for repeat expansion or contraction (Fleming et al., 2003) which may include metabolic pathways (Pearson et al., 2005).

It also possible that types of ions/concentrations present at time of DNA replication, repair or recombination may affect processing of repeat sequences. It is known that concentration of MgCl2...
or other ions as well as water (Radhakrishnan et al., 2006) affect fidelity of DNA replication, repair and recombination process. Further, Taq polymerase fidelity is affected by changes in dNTP concentration, pH, and the concentration of MgCl2 (Eckert and Kunkel, 1990). Further, in presence of K+ during recombination may be excluded from the process. This is particularly observed in repeats that are located towards the 5'-side of the structure like G-quadruplex motif that may be part of recombination (Barros et al., 2009). Further, increase in strand slippage is reported due to inefficient concentrations of ions (particularly Mg ions) during DNA replication (Viguera et al., 2001a and 2001b). The deficit in ion concentrations may be due to either blocking of channel proteins (due to unknown process) or mutations in these proteins. This may also be due to the total potential difference between intracellular and extracellular milieu which can be attributed to disease/physiological status of the individual.

However, in hairpin formation the effect of buffers containing NaCl or NaPi did not show any difference in formation of hairpin structures at normal pH (Chi and Lam, 2009). It is suggested that Z-DNA structure is possible only under conditions such as very high ionic strength, or methylation of all cytosines repeats in linear DNA. However, Hr-DNA is stable under neutral pH in presence of divalent metal ions like Mg2+ whereas Hyr DNA form is stable under mild pH in presence of divalent metal ions. There are some isoforms of these alternate structures that can also be stabilized by monovalent ions. However, among factors favoring non-B DNA structure formation, nucleotide content of the tracts is important (Mirkin, 1999 and 2001).

**Transcription-coupled repeat instability:** Several studies have indicated that repeat instability can occur during transcription, modulated by DNA repair proteins like MMR and transcription-coupled nucleotide excision repair (TC-NER). This can possibly happen by transcription facilitated secondary structure formation in repeat-specific manner which in turn can trigger DNA repair

<table>
<thead>
<tr>
<th>Genes with MSI</th>
<th>Functions</th>
<th>Disease/Cancer</th>
<th>Cause/Mutation</th>
<th>Position of Repeat</th>
<th>Repeat Type</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>APC</td>
<td>Wnt pathway</td>
<td>Colorectal cancer</td>
<td>Frame shift</td>
<td>Exon 10</td>
<td>(A)7</td>
<td>Tougeron et al., 2009</td>
</tr>
<tr>
<td>ATR</td>
<td>Cell cycle checkpoint, response to DNA damage</td>
<td>Colon cancer, Sporadic stomach cancer</td>
<td>Slippage</td>
<td>Exon 3</td>
<td>(G)8</td>
<td>Yamamoto et al., 1997; Rampino et al., 1997; Cao et al., 2003;</td>
</tr>
<tr>
<td>BAX</td>
<td>Proapoptotic factor</td>
<td>Colon, Stomach, Endometrium cancer, Hereditary nonpolyposis colon cancer (HNPPC)</td>
<td>Frame shift</td>
<td>Exon 3</td>
<td>(A)9</td>
<td>Calin et al., 2000</td>
</tr>
<tr>
<td>BLM</td>
<td>DNA repair and helicase</td>
<td>Colon Cancer, stomach cancer</td>
<td>Frame shift</td>
<td>Coding region</td>
<td>(A)9</td>
<td>Menoyo et al., 2001</td>
</tr>
<tr>
<td>CHK-1</td>
<td>Cell cycle checkpoint, response to DNA damage, G2M arrest</td>
<td>Colon cancer, sporadic stomach cancer</td>
<td>Frame shift</td>
<td>Coding region</td>
<td>(A)9</td>
<td>Akiyama et al., 2001</td>
</tr>
<tr>
<td>MLH3</td>
<td>MMR</td>
<td>Colon cancer, sporadic stomach cancer</td>
<td>Frame shift</td>
<td>Coding region</td>
<td>(A)9</td>
<td>Malkhosyan et al., 1996; Mori et al., 2001</td>
</tr>
<tr>
<td>MSH3</td>
<td>DNA Repair mechanism</td>
<td>Colon cancer, sporadic stomach cancer</td>
<td>Frame shift</td>
<td>Coding region</td>
<td>(A)8</td>
<td>Malkhosyan et al., 1996; Mori et al., 2001</td>
</tr>
<tr>
<td>MSH6</td>
<td>DNA Repair mechanism</td>
<td>Colon cancer, sporadic stomach cancer</td>
<td>Frame shift</td>
<td>Coding region</td>
<td>(C)8</td>
<td>Malkhosyan et al., 1996; Mori et al., 2001</td>
</tr>
<tr>
<td>PRDM2</td>
<td>Cell cycle and apoptotic protein</td>
<td>Colon cancer</td>
<td>Frame shift</td>
<td>Coding region</td>
<td>(A)8</td>
<td>Tougeron et al., 2009</td>
</tr>
<tr>
<td>PRDM2</td>
<td>Cell cycle, Wnt pathway</td>
<td>Colon cancer</td>
<td>Frame shift</td>
<td>Exon 8</td>
<td>(A)9</td>
<td>Tougeron et al., 2009</td>
</tr>
<tr>
<td>RAD50</td>
<td>Response to DNA damage</td>
<td>Colon, stomach cancer</td>
<td>Frame shift</td>
<td>Exon 7</td>
<td>(A)6</td>
<td>Tougeron et al., 2009</td>
</tr>
<tr>
<td>TCF4</td>
<td>Transcription factor (Wnt pathway)</td>
<td>Colon cancer</td>
<td>Frame shift</td>
<td>Exon 6</td>
<td>(A)9</td>
<td>Tougeron et al., 2009</td>
</tr>
</tbody>
</table>
mechanisms resulting in repeat instability due to change in repeat length (Lin et al., 2009). Transcription slippage can also result in repeat expansion (poly A or T or GAA.TTC) (Krasilnikova et al., 2007; Werngreen et al., 2010). Such events can cause variations in reading frame resulting at times in functional alternate proteins as seen in prokaryotes. This may rescue functionality of genes possibly by restoring correct reading frame if genes have undergone mutations. Such tracts are conserved in many prokaryotes indicating their evolutionary importance (Werngreen et al., 2010). Ewing's sarcoma can develop due to EWS/FLI an aberrant ETS-type transcription factor that interacts with GGAA-microsatellites to regulate some of its target genes resulting in their dysregulation (Gangwal and Lessnick, 2008). However, some studies contradict involvement of transcription in repeat instability (Krasilnikova et al., 2001; Kelkar et al., 2008).

**Epigenetic factors:** As mentioned in the non-B DNA structure section, methylation of specific repeat sequences in addition to ionic concentrations can lead to Z-DNA formation which may also lead to repeat instability. It is also possible that slippages are more common in case of trinucleotide repeats especially in CG rich repeats because of epigenetic factors especially methylation. Hypermethylation may cause weaker recognition for these repeats in polymerase DNA binding domains thus affect proofreading of both the polymerase and the repair system. However, it is not known whether fidelity of DNA polymerase or repair system is not affected in AT rich repeat regions. Further, if AT rich regions are processed with more efficiency, the mechanism remains elusive. It is interesting to note that expansion of repeats (at least some) is depended on inheritance form either mother or father. Smaller expansions have been found in case of CGG.CCG-repeat tract in the 5' UTR of the FMR1 gene if inherited from father but longer expansions from mother (Entezam and Usdin, 2008).

Besides the aforementioned reasons, mutagen evoked repair pathways can also cause repeat instability especially in non-dividing cells. These mutagens can be UV radiations, oxidative damage, monofunctional alkylating agent MNNG etc. This is supported by observation of contraction of tetranucleotide repeats AAAG(16) slippages due to DNA damaging agent t-butyl hydrogen peroxide but γ-irradiation have little effect (Wang and Vasquez, 2009).

The reasons for SSR mutations as gleaned from the above reports still leave some questions unanswered.

1. Alternate B-DNA structure formations (including stem loop) are considered the cause of slippage and stalling of DNA polymerase. How do these structures form when the SSBP are bound to template? In case of nascent strand there may be delay in SSBP binding due to hemimethylated state of DNA. This can lead to possibility of stem loop formation on the nascent strand but not on the template strand.

2. It is not known why slippage is common on lagging strand. Is it possible that delay (if any) in SSBP binding on the lagging strand is sufficient for stem loop formation especially in regions of CG rich repeats? If this is possible, it leads to another question about why there would be delay in SSBP binding on lagging strand except for time for repair process during replication.

3. It is known that stalling of polymerase at replication fork is one of the reasons for repeat expansion or deletion. However, it is not known how long DNA polymerase will remain stalled to result in repeat mutation.

From the above listed reasons, it can be summarized that DNA replication, repair and recombination processes are not the exclusive reasons for repeat mutations (Mirkin, 2004). Comparisons within species like human and chimpanzee have given indications regarding a sum of aforementioned reasons for mutability of SSRs and heterogeneity regarding mutation rate variations among loci within a species. Rate of SSR mutations and slippage can be predicted on basis of repeat number, length, and motif size. Further, faulty repair, add to mutations as there is a non-uniform increase in mutability with length. It has also been found that mutation rates vary with SSR motifs
(nucleotide composition) which may be due to dissimilarities in secondary DNA structure formed due to slippage. Even presence of some repeats on specific chromosomes and position on chromosome can affect SSR mutation rates. This is seen in case of mononucleotide SSRs located on sex chromosomes verses autosomes and inside verses outside of Alu repeats. In the former case replication process is considered important and in the latter a role for gene conversion (Kelkar et al., 2008).

Conclusion

Though there are several reasons for SSR mutations, characteristics of SSRs appear to be the key factor for repeat instability. These characteristics are the motif size, nucleotide content, length of SSRs, position in genome and distance from replication origin. Replication slippage or failures of repair system to correct SSR mutations are dependent on these characteristics of SSRs. The non-B DNA structure or stem loop formation also depend on CG richness of repeats and their methylation state. Even influences of concentration of ions and temperature that may lead to either secondary structure formation or affect fidelity of DNA polymerase are apparently enhanced in repetitive regions (SSRs) of genome. Despite several disorders that are attributed to repeat instability, it appears that SSRs are necessary evils in any genome due to their functional roles not only in gene regulation but also in providing genome plasticity.

Acknowledgments

We are grateful to Department of Biotechnology (DBT), India for funding part of this work.

References


Cao SN, Chang KH, Luthra R and Liu J. Frameshift mutations in the bax gene are not involved in development of ovarian endometrioid carcinoma. Mod Pathol. 16(10): 1048-52, 2003;


Chi LM and Lam SL. NMR investigation of DNA


Hile SE and Eckert KA. DNA polymerase kappa


Mirkin SM. DNA topology: Fundamentals.


Sinden RR, Potaman VN, Oussatcheva EA, Pearson


