Neurodegenerative triplet repeat expansion disorders: A review

Mitesh Patel, RK Patel, Tushar Chauhan, Jigar Suthar and Sanjay Dave

Abstract
Epigenetic alterations are the major causes of triplet repeat expansion. The repetitive DNA expands of its normal length results in sever neurodegenerative conditions. The common types of triplet repeat expansion (TNE) disorders are: Huntington disease, Friedreich ataxia, myotonic dystrophy, SBMA and SCA1 out of which Huntington disease, SBMA and SCA1 are categorized as a poly glutamine disorder due to the repeat of CAG. In contrast, the friedreich ataxia is occurred due to expansion of the GAA whereas myotonic dystrophy is due to the expansion of CTG. The triplet disease follows the mechanism of anticipation in which the onset of the disease increases with age. Conclusively, no clear mechanism can explain the origin of the disease. The pre mutation can be expanded in full mutation in successive generations and the number of repeats increased with each generation. TNE can observe in both somatic as well as germ line tissues.

Keywords: triplet repeat expansion disorder, trinucleotide repeats, Huntington disease, SBMA, SCA1, friedreich ataxia

Introduction
Since the early 1990s, a new class of molecular disease has been characterized based upon the presence of unstable and abnormal expansions of DNA-triplets (Trinucleotides). The first triplet disease to be identified was fragile X syndrome, which was mapped on the long arm of the X chromosome. At this point, there are from 230 to 4000 CGG repeats in the gene that causes fragile X syndrome in these patients, as compared with 60 to 230 repeats in carriers and 5 to 54 repeats in unaffected individuals. The chromosomal instability resulting from this trinucleotide expansion presents clinically as intellectual disability, distinctive facial features, and macroorchidism in males. The second, related DNA-triplet repeat disease, fragile X-E syndrome, was also identified on the X chromosome, but was found to be the result of an expanded CCG repeat. Identifying trinucleotide repeats as the basis of disease has brought clarity to our understanding of a complex set of inherited neurological diseases.

As more repeat expansion diseases have been discovered, several categories have been established to group them based upon similar characteristics. Category-I includes Huntington's disease (HD) and the spinocerebellar ataxias that are caused by a CAG repeat expansion in protein-coding portions of specific genes. Category-II expansions tend to be more phenotypically diverse with heterogeneous expansions that is generally small in magnitude, but also found in the exon of genes. Category-III includes myotonic dystrophy, two of the spinocerebellar ataxias, juvenile myoclonic epilepsy, and Friedreich's ataxia. These diseases are characterized by typically much larger repeat expansions than the first two groups, and the repeats are located outside of the protein-coding regions of the genes.

The present review discusses the general process of how the triplet repeat expansion is developed, the mechanism of TRE and its associated disease. Additionally, the role of epigenetic alterations in development of triplet repeat expansion disease. We are focusing on the neurodegenerative triplet repeat expansion disorders such as Huntington disease, Myotonic dystrophy, Spinal bulbar muscular atrophy, SCA1 and Friedreich ataxia.

Epigenetic alterations and triplet repeat expansion disorders
Triplet repeats are the type of epigenetic alteration in which the polymorphism is caused by the change in gene expression rather than mutation. The triplet repeats are located on either coding region or noncoding region of the genome. The coding regions are the DNA sequences which involve in protein formation, while the non coding regions are junk DNA which cannot codes for any type of protein. Despite that it plays an important role in maintaining gene expression.
The noncoding sequences present on 3' end of untranslated region works as the enzyme recognition site for completion of transcription and translation. Epigenetic alterations such as methylation, Histone modification and RNA processing involve in gene silencing (Sweatt et al., 2013) [36]. In methylation, once the methyl group is added to the DNA sequence, it become transcriptionally silent. Hence, it cannot undergo gene expression process. Methylation promotes tight wrapping of DNA and it cannot go through gene expression.

![Fig 1: The length of the different triplet repeats.](image)

The expression of triplet repeat DNA is highly protected by methylation (Sweatt et al., 2013) [36]. If the methyl group removed, the triplet repeat will expand and results in disease condition. The etiology of disease is different in both coding as well as noncoding region. In noncoding region the TNR alters the splicing of protein, hence alters the protein expression while in coding region TNR alters the function of the mutated protein (Budworth & Mcmurray, 2013) [2].

The noncoding or junk DNA of our genome is located in the satellite region of the chromosomes. Specifically, the TNR sequences are located at microsatellite consist of 30% portion of the human genome. Actually, the alteration in the length of the DNA provides additional diversity in the species. However, the rapid change in human genome is unaccepted. So it resists the rapid changes (as like TNR expansion) that could be deleterious (Budworth & Mcmurray, 2013) [2].

TNRs are largely divided into two regions: Coding and noncoding region. However, TNRs are more common is noncoding regions. Fragile X syndrome, Friedreich ataxia and spinocerebellar ataxia type 7 and type 1 are the type of TNR disorders common in non coding DNA sequences whereas the Huntington disease is the type of TNR disorder of the coding sequence (Pook, 2012) [25]. The trinucleotide repeats has two important properties, in each successive generation the repeat tends to expand. However, the contraction of repeat is never reported. Secondly, the repeat instability is observed in both somatic and germ line cells (Cummings & Zoghbi, 2000) [4].

The triplet repeat expansion disease of the coding region is located in exon while the triplet repeats in noncoding regions are located on 3' untranslated region, 5' untranslated region or introns. The complete list of triplet repeat disorder of the coding region and noncoding region are listed in the table below.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Gene</th>
<th>Location</th>
<th>Repeat</th>
<th>Normal range</th>
<th>Full mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myotonic Dystrophy type 1</td>
<td>DMPK</td>
<td>3' UTR</td>
<td>CTG</td>
<td>5-37</td>
<td>&gt;50</td>
</tr>
<tr>
<td>Myotonic Dystrophy type 2</td>
<td>CNBP</td>
<td>Intron 1</td>
<td>CCTG</td>
<td>&lt;30</td>
<td>75-11000</td>
</tr>
<tr>
<td>Fragile X mental retardation(FRAXA E)</td>
<td>AFF2</td>
<td>5' UTR</td>
<td>GCC</td>
<td>4-39</td>
<td>&gt;200</td>
</tr>
<tr>
<td>Friedreich ataxia</td>
<td>FXN</td>
<td>Intron 1</td>
<td>GAA</td>
<td>5-30</td>
<td>70-1000</td>
</tr>
<tr>
<td>Fragile X syndrome</td>
<td>FMR1</td>
<td>5' UTR</td>
<td>CGG</td>
<td>6-50</td>
<td>200-4000</td>
</tr>
<tr>
<td>Huntington disease like 2</td>
<td>JPH3</td>
<td>Exon 2A</td>
<td>CTG</td>
<td>6-27</td>
<td>35-57</td>
</tr>
<tr>
<td>SCA8</td>
<td>ATXN8OS</td>
<td>3' UTR</td>
<td>CTG</td>
<td>15-34</td>
<td>89-250</td>
</tr>
<tr>
<td>SCA12</td>
<td>PPP2R2B</td>
<td>5' UTR</td>
<td>CAG</td>
<td>7-28</td>
<td>66-78</td>
</tr>
</tbody>
</table>
Mechanism of triplet repeat expansion

Several mechanisms such as DNA structure in TRE, mismatch repair during TRE, strand slippage and DNA polymerase pausing in TRE are support the TRE. The best suitable model organism for studying TRE mechanism is E. coli.

The packaging of DNA starts with the formation of nucleosome assembly. The nucleosome assembly is a complex arrangement of DNA with the protein. The 146bp long stretch of DNA is wrapped on the histone octamere (having Histone H2A, H2B, H3 and H4) (Loscalzo & Handly, 2014) [16]. As the nucleosome assembly becomes tight, the transcription efficiency of DNA will decrease. CTG repeats more than 250 will suppress the activity of transcription (Wells 1996) [40]. The structure of DNA in TRE is different from the native B- form DNA. The non B form of triplet repeat DNA of CTG and CGG is called as toroids. The repeats of approximately 81 bp are present in helical structure of toroidal DNA. The DNA in a helical structure is curved and paired in duplex hence it is also called as H form DNA. The H form of the DNA is non prevalent in nature with having triplet strand. The formation of non B form toroidal DNA causes triplet repeat to expand by the mechanism of strand slippage (Sinden et al., 2002) [32].

Actually, the process of triplet repeat expansion follows the non Mendelian pattern of inheritance. Polymerase pausing strongly supports these findings. The strand slipping is another important mechanism which explains the expansion of triplet repeat expansion. The Okazaki initiation zone (OIZ) present on the lagging strand is the primary site for the expansion of the triplet. The process of triplet expansion is triggered by the strand separation during replication or by the DNA unwinding. The process of leading strand and lagging strand synthesis is different from each other, the coordination of strand synthesis of leading and lagging strand is important to control the coordination, the OIZ remain initially single stranded. Hence, the triple repeat present on the lagging stand fold into secondary DNA structures called as hairpin loop (Siyanova & Mirkin, 2001) [33]. This secondary structure is unusual and different than the B form of DNA. The secondary structure becomes more and more stable at lagging strand it slows down the process of replication by hindering at the replication fork. It also slows down the polymerase activity. However, the proteins are involved in DNA replication and repair may help in stopping expansion. The mechanism of strand slipping involves the process of polymerase pausing during replication. Here mesh up in polymerase activity leads to strand slippage. The DNA synthesis by polymerase is stopped at triplet repeat. Once it recognizes the triplet repeats are more than the normal one it pauses the activity. However, in vitro, the pause can be removed by heating at 70°C (Wells, 1996) [40]. When the triplet repeat in dsDNA is increased the pausing of polymerase is also increased. Hence, non B form “toroid DNA” (also called as H- DNA) is formed which blocks the polymerase progression. The entire strand misses the replication and mismatch repair, subsequently inherited into the next DNA replication cycle.

Table 2: Triplet repeats expansion disorders in coding region of the DNA: gene, location, repeats and range of disease.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Gene</th>
<th>Location</th>
<th>Repeat</th>
<th>Normal range</th>
<th>Full mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huntington Disease</td>
<td>IT15</td>
<td>Exon 1</td>
<td>CAG</td>
<td>6-29</td>
<td>38-180</td>
</tr>
<tr>
<td>DRPL</td>
<td>ATN1</td>
<td>Exon 5</td>
<td>CAG</td>
<td>6-35</td>
<td>49-88</td>
</tr>
<tr>
<td>SCA1</td>
<td>ATXN1</td>
<td>Exon 8</td>
<td>CAG</td>
<td>6-39</td>
<td>41-83</td>
</tr>
<tr>
<td>SCA2</td>
<td>ATXN2</td>
<td>Exon 1</td>
<td>CAG</td>
<td>&lt;31</td>
<td>32-200</td>
</tr>
<tr>
<td>SCA3</td>
<td>ATXN3</td>
<td>Exon 8</td>
<td>CAG</td>
<td>12-40</td>
<td>52-86</td>
</tr>
<tr>
<td>SCA6</td>
<td>CACNA1A</td>
<td>Exon 47</td>
<td>CAG</td>
<td>&lt;18</td>
<td>20-33</td>
</tr>
<tr>
<td>SCA7</td>
<td>ATXN7</td>
<td>Exon 3</td>
<td>CAG</td>
<td>4-17</td>
<td>36-440</td>
</tr>
<tr>
<td>SCA17</td>
<td>TBP</td>
<td>Exon 3</td>
<td>CAG</td>
<td>25-42</td>
<td>45-66</td>
</tr>
<tr>
<td>SMBA</td>
<td>AR</td>
<td>Exon 3</td>
<td>CAG</td>
<td>13-31</td>
<td>&gt;40</td>
</tr>
<tr>
<td>Synpolyactyly H</td>
<td>HOXD 13</td>
<td>Exon 1</td>
<td>GCG</td>
<td>1-15</td>
<td>&gt;22</td>
</tr>
<tr>
<td>Cleidocranial dysplasia</td>
<td>RUNX2</td>
<td>Exon 1</td>
<td>GCG</td>
<td>2-17</td>
<td>&gt;26</td>
</tr>
<tr>
<td>HFPG</td>
<td>HOXA13</td>
<td>Exon 1</td>
<td>GCG</td>
<td>12</td>
<td>&gt;17</td>
</tr>
<tr>
<td>XLMR</td>
<td>ARX</td>
<td>Exon 2</td>
<td>GCG</td>
<td>16</td>
<td>&gt;17</td>
</tr>
<tr>
<td>OPMD</td>
<td>PABPN1</td>
<td>Exon 1</td>
<td>GCG</td>
<td>10</td>
<td>&gt;11</td>
</tr>
</tbody>
</table>

Here, in strand slippage the loop created by the repeated nucleotide stops the activity of the polymerase therefore polymerase is temporarily released from the site of replication. Meanwhile, the repeats are paired with the repeats of another stand and replicated. Ultimately, the repeat numbers are increased in each replication cycle due to the polymerase pausing. Multiple slippages may form which are skipped by the polymerase and continues the triplet repeat expansion. The strand slippage mechanism has limited to some repeats, not all repeats can follow this mechanism (Siyanova & Mirkin 2001) [33].

Inadequate level of replication protein leads to expansion. The protein MSH2 and MSH3 complex can repair the mismatch bases in replication. Once the unusual hairpin structure is formed, immediately, the MSH2/MSH3 complex binds strongly to the loop. But the low level of ATPase activity results in non-functionality of this complex. Hence, instead of repairing it, the complex stabilizing the hairpin loop of triplet repeat (Siyanova & Mirkin, 2001) [33]. Mismatch repair during the TRE is another mechanism which is more prominently accepted. In vitro studies on E.coli suggest that during replication, the single stranded loop of triplet repeat is formed which creates the hairpin like structure (Sinden et al., 2002) [32]. This unstable hairpin structure is skipped by the polymerase during the proofreading of the mismatch repair. The loop is highly unstable and expands in subsequent generations. As the prokaryotic and eukaryotic mismatch proteins are same, similar mechanism is applicable for Hunan DNA as well. When plasmid containing CTG repeats more than 100 were grown for about 100 generation indicates that over 85% of the colonies contains the full mutation as compared to the parental strains. This result supports the mismatch repair mechanism involve in CAG repeat expansion (Wells, 1996) [40].
Huntington disease (HD)

Huntington disease is an autosomal dominant disorder (Paulsen et al., 2008) [21] previously, called as dancing disorder or epidermic dance (Molon, L Reyes, 2010) [18], the Huntington disease was first reported by George Huntington in 1872. However, the occurrence of the syndrome was seen since 1600. The prevalence of the disease was reported since 1374 (Molon, L Reyes, 2010) [18]. In 1993, with the help of the DNA studies, it was confirmed that the Huntington disease is occurring due to the abnormal growth of triplet nucleotide repeat of CAG. The condition earlier was called as a Huntington chorea or chorea (Molón, L Reyes, 2010; Rawlins et al., 2016) [18, 28]. The Huntington disease is prevalent 1 in 10,000 individual globally (Strobel, 2013) [35]. The disease is most prevalent in Caucasian population. It is a type of polyglutamine disorder.

The disease is originated because of the progression of intermediate CAG repeat into full mutation in germ-line cells (Paulsen et al., 2008) [21]. Hence it is inherited into the next generation (Rao et al., 2017) [27]. The Huntingtin protein encoded by the IT15 gene (Warren, 1993) [34] is located on the chromosome 4 (Pidgeon & Rickards, 2013; Strobel, 2013) [23, 35]. CAG repeat is located at 5’ untranslated region of the IT15 gene. Herein, HD disease the abnormal repeat number of CAG is responsible for the development of the disease. Normally, 5 to 15 CAG repeats are present in all individuals. 26 to 39 repeat numbers are intermediate which is more prone to the disease. In intermediate region, the range of 36 to 39 CAG repeats are called as incomplete penetrance in which the symptoms may be observed very late onset of age. The full mutation is observed in the repeat more than 40. The CAG codes for the amino acid glutamine hence, the disorder is often named as a polyglutamine track or polyQ disorder (Ashley and Warren, 1995) [31]. Generally, the repeats are present in the 5’ untranslated regions. Nonetheless, when the repeat numbers are exceeded more than 28, it may show the instability in replication (Walker, 2007) [30]. Replication leads expansion in 73% cases of unstable repeat numbers.

It is a type of neurodegenerative disorder, is mainly involved in nervous system and brain (Rao et al., 2017) [27]. Motor and cognitive disabilities are generally reported in HD patients so it is characterized as Neuro-degenerative disorder. Feeling sad, socially inactive, rigid muscle contraction, jerky body movements, slow and abnormal eye movements, difficulty in swallowing and talking, insomnia, fatigue or loss of energy, difficulty in focusing on a task, lack of impulse control, acting without thinking, lack of behavioral flexibility and other motor disorders are common in Huntington disease. The activity of mitochondrial enzyme is greatly reduced in HD patients especially, complex II and complex III activity, is decreased as compared to complex I and II. Further, the activity of enzyme pyruvate dehydrogenase is decreased in HD patients. The progression of disease is observed in both somatic as well as germ line cells, hence it can inherit into the next generation (Rao et al., 2017) [27].

Juvenile Huntington is reported in approximately 10% of the total disease population and the disease symptoms appear at the age between 10 to 20 years. In juvenile Huntington, learning and understanding difficulties at school in early life is very common (Roos, 2010) [31]. However, Juvenile HD disease is very rare because the disease strikes late onset of age of 35 to 50 or above. The full mutation is caused by the repeats more than 41 in juvenile HD, the repeat numbers may increase more than 55(Walker, 2007) [39]. The Huntington disease originated paternally (Reyes, 2010) [18]. The process of spermatogenesis is more unstable as compared with oogenesis so the prevalence of full mutation is very high in male as compared with females (Reyes, 2010) [18]. The most common causes of death in HD are pneumonia and suicide. The suicide tendency is reported very high in HD patient. Hypokiniesia, akinesia and breadykinesia are very commonly observed in HD patients along with the rigid slower pace of activities. Depression, apathy, weight loss is observed as well. Peoples of HD have usually low self-esteem with guilt feelings and anxiety. At the early stage, the hyper sexuality is a major problem associated with HD (Roos, 2010) [31]. The possible mechanism which eventually responsible for death are: mitochondria enzyme dysfunction, dysregulation of transcription, exitotoxicity, apoptosis (Pidgeon and Rickards, 2013) [23].
Myotonic dystrophy (MD)

Another type of TNR disorder is myotonic dystrophy. It is an autosomal dominant disorder caused by the expansion of CTG repeats (Cardoso and Marques, 2018) [39]. The prevalence of MD is 1 in 8000 individuals globally (Lesja, 2014) [19]. Prevalence of MD is reported in India as well. The CTG repeat is located on the 3rd untranslated region of the dystrophia myotonica protein kinase (DMPK) gene, again it is non-coding type of triplet repeat (Mankodi, 2014). The gene DMPK is located on the chromosome number 19 at 19q13.0 (Dahlqvist et al., 2014; Ray, 1989) [5, 8]. The two most common types of MD are reported: MD type I, named Steinert’s disease and MD type II, named proximal MD (Grunseich et al., 2014; Huang & Kuo, 2013) [9, 11]. MD type I is caused by the CTG repeat, whereas MD type II is caused by CCTG (Yum et al., 2017) [41]. Myotonic dystrophy is the most common type of dystrophy observed in humans. Several hypotheses which support the expansion of CTG repeat in MD are explained here. 

The scarcity of kinase which results in dystrophy is occurred by the blockage of DMPK gene, primary transcript. The RNA retains in nuclei therefore it cannot processed for post transcriptional modification at cytoplasm result in MD. However, experiments of knockout mice indicated that the proposed hypothesis is not always supported in case of heterozygous (Mankodi, 2014). In another hypothesis, CTG repeats expansion at the 3’ end result in repression of SIX5 gene. The gene SIX5 is located downstream to the gene DMPK, which regulates the muscle cell differentiation. Expansion of CTG repeat makes structural changes in surrounding chromatin structure, hence the SIX5 gene activity is affected by CTG expansion. Experiment on knockout mice indicates that dystrophy is observed in each case of homozygous and heterozygous. However, major features of dystrophy are not seen (Siyanova and Mirkin, 2001) [33].

The third hypothesis is widely and most promisingly accepted. The muscle specific RNA processing is altered by the CUG binding protein. Certainly, CUG repeats present on the RNA of DMPK gene promotes binding of several proteins like CUGBP1 which hinders in RNA processing, specifically in muscle specific tissue (Mankodi, 2014). MD characteristics are observed in transgenic mice containing more than 250 CUG repeats in DMPK RNA (Siyanova and Mirkin, 2001) [33].

<table>
<thead>
<tr>
<th>Gene</th>
<th>Location of chromosome</th>
<th>Inheritance pattern</th>
<th>Gender bias</th>
<th>Disease</th>
<th>Mutation type</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMPK</td>
<td>19q13</td>
<td>Autosomal dominant</td>
<td>Maternal</td>
<td>Myotonic dystrophy</td>
<td>Gain of function</td>
</tr>
<tr>
<td>X25</td>
<td>9q13-21.1</td>
<td>Autosomal recessive</td>
<td>Maternal</td>
<td>Friedriach ataxia</td>
<td>Loss of function</td>
</tr>
<tr>
<td>AR</td>
<td>X13q-21</td>
<td>X linked recessive</td>
<td>Not specified</td>
<td>SBMA</td>
<td>Gain of function</td>
</tr>
<tr>
<td>IT15</td>
<td>4p16.3</td>
<td>Autosomal recessive</td>
<td>Paternal</td>
<td>Huntington disease</td>
<td>Gain of function</td>
</tr>
</tbody>
</table>

The normal range of the CTG repeat is 5 to 37 (Huang & Kuo, 2013) [11], repeats range between 38 to 50 are the premutation condition which is responsible for the “anticipation” (Prendergast et al., 2010) [20]. The process in which the length of the repeats increased as compared to their parents is called as the process of anticipation (Yum et al., 2017) [41]. Myotonic dystrophy having the characteristic of muscular weakness, muscle wasting and muscular myotonia and dystrophy (Cardoso and Marques, 2018; Mankodi, 2014) [39]. Furthermore, cardiac defects (Hahn and Salajegheh, 2016) [10], premature balding in male and cataracts are observed in some cases. MD shows clear anticipation. The mechanism of inheritance of DM repeat expansion is still unclear because the expansion or contraction is not depends on the any of the parents (Warren, 1993) [16].

Friedreich ataxia

It is an autosomal recessive disorder observed without the anticipation. It is caused by the GAA repeat expansion on non-coding region of fraxatin gene also called as X25. Hypermethylation in X25 gene results in transcriptional silencing and reduced expression of X25 gene. Iron homeostasis and respiratory dysfunctions are commonly observed in Friedreich ataxia. Hence fraxatin deficiency results in oxidative stress, cellular degeneration and mitochondrial dysfunction at cellular level (Grunseich et al., 2014; Cummings & Zoghbi, 2000) [9, 4]. Friedreich ataxia is the most prominent trinucleotide repeat disorder observed worldwide. It is a form of ataxia with cardiomyopathy with diabetes mellitus. Areflexia and dysarthria are observed as well. Hypermethylation in the GAA repeats located in the first intron of the X25 gene is responsible for Friedreich ataxia. Transcriptional blockage, disturbance in enhancer activity and abnormal splicing are some of the possible mechanisms which cause the transcriptional silencing of X25 gene. In 1998, Bidichandani el al., first reported that the expansion of GAA trinucleotide repeats in Friedreich ataxia may be associated with an unusual DNA structure. Later on, several studies supports the finding that the expansion of GAA repeats of intron 1 X25 gene causes structural deformities in DNA. Interestingly, the variation in the expansion of GAA repeats is observed between somatic and germ line cells (Patel and Isaya, 2001) [29]. Fraxatin is a mitochondrial protein loss of the function leads to the accumulation of iron in mitochondria, which results in hypersensitivity in oxidative stress. Unregulated iron homeostasis causes cell damage. Yet, the role of iron deficiency is still questionable in Friedreich ataxia (Siyanova and Mirkin, 2001) [33].

Spinal-bulbar muscular atrophy

Kennedy’s disease often called as SBMA is a type of X linked recessive atrophy (Jordan & Lieberman, 2008) [13]. A type of atrophy which mostly occur in spinal muscle, proximal limb muscles and bulbar muscle in chewing and facial muscles. The disease was first explained by the Japanese neurologist, Kawahara. Later on the disease was completely described by the Kennedy in 1968 (Tanaka et al., 2012) [37]. The adult onset of the disease is seen between age of 30 and 50 years. Tremor, muscle cramps and fatigue are also commonly seen in SBMA (Lindsay E. Rhodes, 2012). Additionally, glucose intolerance, liver dysfunction is observed in some cases. The prevalence of SBMA has been estimated 1 in 40,000 in areas of high ascertainment (Masahisa Katsuno, 2003) [17].
Due to the origin of occurrence, individuals may feel weakness, cramping, problem in chewing, gynaecomastia and fasciculation. At a later stage of life individual may suffer from azoospermia and testicular atrophy. The gene AR (Grunseich et al. 2014) is located on the X chromosome at Xq11.0 to Xq12.0. CAG repeats present on the coding exonic region of AR gene (Jobsis et al., 1995) [12]. The AR gene is an androgen receptor gene, which can induce its gene expression after binding with the ligand. It is a type of receptor which cannot be activated without an appropriate ligand (Zelinkova et al., 2016) [42]. Dihydrotestosterone and testosterone are two major ligand for binding of AR gene (Freeman et al., 2012; Nelson, 2013). The AR gene is hormone receptor of the steroid hormones made up of DNA binding domain, ligand binding domain and amino terminus domain (Tanaka et al., 2012) [37]. Full mutation is observed in exonic region having CAG repeats between 40 to 60. CAG repeats between 12 to 30, is described as a premutation condition which is most prone to expand into full mutation. The chance of full mutation is higher in case of parental transmission because of the instability in spermatogenesis (Stephen T warren, 1993) [34]. Here in the case of SBMA, the muscle weakness is correlated with the number of repeats (Rocchi and Pennuto, 2013) [30]. Additionally, the number of the repeat is also correlated to the age of onset anticipation. For example, the largest CAG expansion with 51 repeats had age of onset in late thirty. However, muscle weakness is observed at an age between 50 to 60 with repeats less than 51 (Stephen T. warren, 1993) [34]. The CAG repeat over 38 to 62 leads to full mutation of the disease. Yet, the range of the disease is 9 to 36. The neurological degeneration caused by the fragmentation of AR gene located in the nuclei of the cell. Mild AIS symptoms are seen in SBMA such as feminized skin, less hair, azoospermia, oligospermia and gynecomastia (Jordan and Lieberman, 2008; Tanaka et al., 2012) [11,37].

As the mutation causes the testicular disfunction, it is strongly associated with the male infertility. Even though, genotypical the individual is male (XY), they have external female characteristics. Definitely, it is a gain of function mutation which is strongly associated with abnormal male testicular development, with or without muscular atrophy (Warren, 1993) [34].

**Spinocerebellar ataxia type 1 (SCA1)**

SCA1 inherited as an autosomal dominant disorder (Longshore and Tarleton, 1996) [15] which is characterized by progressive Neuro-degenerative disorder. Due to the loss of Neurons in spinocerebellar track, SCA1 exhibits wide range of neurological symptoms such as motor and muscle weakness, ophthalmoparesis and ataxia (Warren, 1993) [34]. Typically, the symptoms are observed at late forties (Kumaran et al., 2014) [14]. It is also a kind of poly glutamine disorder (Paulson, 2009; Warren, 1993) [34]. The CAG repeats of SCA1 are located on the ATXN1 gene (Platono et al. 2016) [29], exon 8 and chromosome 6. Approximately 2 in 100,000 suffer from SCA1 (Rengaraj et al., 2005) [29]. CAG repeats ranging between 6-39 are observed in normal individuals. The intermediate condition ranges between 36-38 whereas the full mutation occurs with a repeat number more than 39 to 44 (Vishwakarma et al., 2018) [38]. Anticipation has been most common in SCA1 indicates that the severity of disease increased as into the successive generations (Warren, 1993) [34]. SCA1 is observes in 30s and 40s (Kumaran et al., 2014) [14]. The juvenile form of the disease is more severe and individual may die within 16 years. Slurry speech, unbalancing in downstairs, lack of control are commonly observed. Problem in choking of food and drink may also observe in some cases. As the disease becomes more severe other signs related to ataxia are prominently observed such as dysmetria, dysdiadochokinesia, and hypotonia become apparent (Warren, 1993) [34].

**Conclusion**

The present review has enlisted several characteristics of triplet repeats with possible explanation for each type of repeats involve in disease.

**Characteristic of the repeat expansion**

**Microsatellite instability**: the repeats are located on the satellite region of the DNA. Once the methylation or methyl group is removed from the satellite repeats, it becomes unstable.

**The threshold of repeat length**: not all the triplet repeats can expand into the full mutation because the repeat has several threshold values which are necessary for the expansion. The repeat tends to expand when the base pairs exceed more than 200 in a repeat. As followed by the expansion mechanism or strand slippage, in each successive generation of cell division, the base pairs in repeat become increases. Once it crossed the threshold of 200 bp, it becomes more actively involve in expansion.

**Genetic anticipation**: Anticipation is a process in which the severity of disease increase as the age of the inset increases with increased repeat numbers. The mechanism of genetic anticipation was reported in 1918 for the disease myotonic dystrophy. The anticipation is commonly observed in all types of triplet repeat expansions except Friedreich ataxia.

**Severity of disease**: As the number of the repeats increase the severity of the disease increases gradually. It is also notable that in some of the cases in which the full mutation repeats are near the premutation or just started, the severity of disease is mild. Conclusively, we can say, the epigenetic alterations are the major causes of TRE, especially methylation. Each mechanism described are only explained by prokaryotic models, evidently it is not explained in humans. Further, the mechanism of anticipation, how repeats are expanded in each successive generations are still not explained clearly.

**References**