A novel pathogenic variant in an Iranian Ataxia telangiectasia family revealed by next-generation sequencing followed by in silico analysis

Mohammad Amin Tabatabaiefar a,b, Paria Alipour c, Azam Pourahmadiyan c, Najmeh Fattahi d, Laleh Shariati e, Neda Golchin f, Javad Mohammad-Jasli g,h

a Department of Genetics and Molecular Biology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran
b Pediatric Inherited Diseases Research Center, Research Institute for Primordial Prevention of Non-communicable Disease, Isfahan University of Medical Sciences, Isfahan, Iran
c Cellular and Molecular Research Center, Basic Health Sciences Institute, Shahrekord University of Medical Sciences, Shahrekord, Iran
d Clinical Biochemistry Research Center, Cardiovascular Research Institute, Isfahan University of Medical Sciences, Isfahan, Iran
e Cardiovascular Research Center, Cardiovascular Research Institute, Isfahan University of Medical Sciences, Isfahan, Iran
f Ahvaz Noor Genetics Laboratory, Ahvaz, Iran
g Department of Medical Genetics, School of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran
h Department of Genetics and Molecular Biology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

A B S T R A C T

Ataxia telangiectasia (A-T) is a neurodegenerative autosomal recessive disorder with the main characteristics of progressive cerebellar degeneration, sensitivity to ionizing radiation, immunodeficiency, telangiectasia, premature aging, recurrent sinopulmonary infections, and increased risk of malignancy, especially of lymphoid origin. A-T is caused by a mutation in the ATM gene, which plays an important role in the activation of cell-cycle checkpoints and initiation of DNA repair in response to DNA damage. Targeted next-generation sequencing (NGS) was performed on an Iranian 5-year-old boy presented with truncal and limb ataxia, telangiectasia of the eye, Hodgkin lymphoma, hyper pigmentation, total alopecia, hepatomegaly, and dysarthria. Sanger sequencing was used to confirm the candidate pathogenic variants. Computational docking was done using the HEX software to examine how this change affects the interactions of ATM with the upstream and downstream proteins. Three different variants were identified comprising two homozygous SNPs and one novel homozygous frameshift variant (c.80468047delTA, p.Thr2682ThrfsX5), which creates a stop codon in exon 57 leaving the protein truncated at its C-terminal portion. Therefore, the activation mozygous SNPs and one novel homozygous frameshift variant were confirmed as pathogenic based on the American College of Medical Genetics and Genomics guideline. This study expands the spectrum of ATM pathogenic variants in Iran and demonstrates the utility of targeted NGS in genetic diagnostics.

© 2017 Published by Elsevier B.V.

1. Introduction

Ataxia telangiectasia (A-T)(MIM_208900), as an autosomal recessive cerebellar ataxia [1], has been described with different characteristics such as progressive cerebellar degeneration, sensitivity to ionizing radiation, immunodeficiency, telangiectasia, premature aging, recurrent sinopulmonary infections, and increased risk of malignancy, especially of lymphoid origin, poor growth, gonadal atrophy, delayed pubertal development, and insulin-resistant diabetes [2]. The worldwide frequency of A-T is estimated to be between 1 in 40,000 and 1 in 100,000 live births [3]. A-T occurs as a result of mutation within ATM (MIM# 607585). The gene, previously mapped to the chromosomal region 11q22-23 [4], was identified by Savitsky et al. [5]. ATM is a large gene with 66 exons, which spans over 150 kb of genomic DNA [6,7], and encodes a protein with 3050 residues. It is a member of phosphoinositide 3-kinase-related gene, previously mapped to the chromosomal region 11q22-23 [4]. The American College of Medical Genetics and Genomics guideline. This study expands the spectrum of ATM pathogenic variants in Iran and demonstrates the utility of targeted NGS in genetic diagnostics.

© 2017 Published by Elsevier B.V.
FANCD2 in order and activate both cell cycle checkpoints and start DNA repair [9,10].

Out of all ATM pathogenic variants which are almost 800 so far (http://www.hgmd.cf.ac.uk/ac/index.php), truncating variants are the most common variants which are likely to create highly unstable protein fragments [11,12]. Pathogenic variants have been spread uniformly over the coding region of the ATM, and no mutational hotspot has been identified in this gene [13].

Next-generation sequencing (NGS) is a fast and cost-effective sequencing technology through which millions of DNA sequence reads are generated in a single run. It has revolutionized the way genome variations are identified for genetic disease [14,15].

There are few reports of pathogenic variants in the ATM gene from the Iranian population [16]. In this study, we applied NGS to the molecular diagnosis of A-T in an Iranian patient and report a novel pathogenic variant.

2. Materials and methods

2.1. Subjects

A 5-year-old boy with the clinical presentation of A-T, born to an Iranian consanguineous family of Persian descent, was referred to Ahvaz Noor Medical Genetics Laboratory in Ahvaz, Southwest Iran. Genetic evaluation showed the truncal and limb ataxia oculocutaneous telangiectasia, Hodgkin's lymphoma, hyperpigmentation, total alopecia, hepatomegaly and dysarthria. Noticeably, the elder son of this family was also an A-T patient and showed the same symptoms, at the age of 3.5 years old, as the proband but more severely affected. He died at the age of 8 due to the respiratory failure caused by the metastasis of Hodgkin's lymphoma to the lungs. The primary clinical diagnosis for the proband was mainly based upon the symptoms of progressive cerebellar ataxia, dysarthria, and telangiectasia of the eye which was the same as his elder brother. The proband showed elevated levels of serum α-fetoprotein (=150.30 ng/ml), alkaline phosphatase (+1505), creatinine (=226), and also showed altered immunoglobulin profiles (lgG = 7.42 g/l, lgA < 0.6 g/l, lgM = 2.44 g/l, lgE < 0.1 g/l), but no cerebellar atrophy was detected. The parents were clinically unaffected without any neurological impairment, and had no familial history other than their children. The family signed a written consent form before undergoing the genetic test. The study was approved by the Review Board of Isfahan University of Medical Sciences.

2.2. DNA extraction and mutation screening via NGS

Based on PAXgene Blood DNA Kit (QIAGEN, Germany) protocol, genomic DNA from blood samples was extracted and assessed in terms of quality and quantity using agarose gel and a Nanodrop 2000 instrument (Thermo Fisher Scientific Inc., USA), respectively. The NGS test was performed using a custom-designed Nimblegen chip capturing the fragments, including exons, splice sites, and the adjacent intron sequences within the ATM gene. First, for sample enrichment, fragmentation and polishing of genomic DNA were done. Then, linkers were ligated to the fragmented samples. When array hybridization was completed, the fragments of interest were washed from the array using a target fragment elution wash. Following enrichment of the ATM gene, sequencing was performed using HiSeq2000 instrument (Illumina, San Diego, CA) with the 90 bp pair-end reads. The variant analysis was performed via mapping the sequence read to the reference genomic DNA (UCSC hg19) using the Burrows-Wheeler alignment software. GATK software was used to identify SNPs and indels [17]. Then, the identified variants (frequency > 1%) and synonymous substitutions were filtered out applying public databases, including 1000 Genome Project (http://www.1000genomes.org), HapMap samples, and dbSNP.

2.3. The 3-D structure of the ATM protein and computational docking

SWISS-MODEL was used to predict the 3D structure of the ATM protein, and its truncated form based upon the homology modeling method. To predict the effect of the ATM frameshift variant on its function, ATM interaction through its C-terminal with some proteins such as TIP60 and p53 [9,10] was examined via docking stimulation using the Hex software. The PDB file of the two proteins is needed as an input. Then, the energy of two-protein interaction was calculated known as Etot. Etot is the total calculated interaction energy of the system [18].

2.4. Co-segregation study and pathogenicity examination

The co-segregation of the variant with the A-T phenotype was done using Sanger sequencing for the proband and other members within the family in order to analyze the variant reported via NGS. A set of primer was designed via Primer 3 software (F: 5'-ACACCGGGCTAAAGTTGTA-3' and R: 5'-CAAAATCCCCAATAAGCCAGAAA-3') to amplify the region of small deletion in exon 57 of all family members. Then, the amplification was subject to bidirectional Sanger sequencing using an ABI 3130 automated sequencer. The sequences were compared with the ATM gene reference sequence. Based on the criteria introduced by the American College of Medical Genetics and Genomics (ACMG) guideline [19], the pathogenicity of the frameshift variant in ATM gene was examined.

3. Results

3.1. Molecular finding

Targeted NGS for exons and exon-intron boundaries of the ATM gene was performed for an affected member of the A-T family. The test platform was examined over the 95% of the ATM gene with sensitivity over 99%. Point mutation, micro-insertion, deletion and duplication (<20 bp) were simultaneously detected. As a result, three different variants were detected (Table 1); two homozygous substitution variants (p.Asp1853Asn and p.Asn1983Ser), and a small frameshift deletion (c.8046_8047delTA, p.Thr2682ThrsfsX5). Based on data provided in the NCBI Clinivar database, both homozygous substitution variants are benign, while the p.Asp1853Asn variant in combination with IVS38-8>T>C in cis position is possibly associated with bilateral breast cancer [20]; it also predicts late fibrosis after radiotherapy in cancer patients [21].

According to the ACMG guideline, this variant (c.8046_8047delTA, p.Thr2682ThrsfsX5) is predicted to be pathogenic based on some lines of evidence; Very strong (PVS1): it is a null variant in the ATM gene which loss of function is a known mechanism of disease. Supporting (PP1): the co-segregation study using the Sanger sequencing method reveals that the variant is co-segregating with the disease (Fig. 1). Moderate (PM2): this variant is absent from the controls of the Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium database. The novelty of this variant was also examined using international mutation and polymorphism databases such as HGMD Database, Swissvar, OMIM, ensemble, ClinVar. After elucidating the pathogenicity of the variant, during her third pregnancy, the mother

<table>
<thead>
<tr>
<th>Variant Name</th>
<th>RS-ID</th>
<th>Frequency in dbSNP</th>
<th>Hapmap</th>
<th>1000-genome</th>
<th>BGI</th>
</tr>
</thead>
<tbody>
<tr>
<td>p.Asp1853Asn</td>
<td>rs1801516</td>
<td>0.006</td>
<td>0</td>
<td>0.0128</td>
<td>0</td>
</tr>
<tr>
<td>p.Asn1983Ser</td>
<td>rs659243</td>
<td>1</td>
<td>0.992</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>p.Thr2682ThrsfsX5</td>
<td>Novel</td>
<td>–</td>
<td>–</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

a dbSNP: SNP frequency in the dbSNP database.

b Hapmap: SNP frequency in Asian population in the Hapmap database.

c 1000-genome: SNP frequency in all the samples from the 1000 Genome Project.

d BGI’s: SNP frequency in 901’s local ‘200 people genome database.'
requested prenatal diagnosis. Being heterozygous rather than homozygous for the ATM mutation, the fetus was predicted to be unaffected by A-T which was confirmed after birth (the rhombic symbol in Fig. 1).

3.2. 3D structure of ATM and its truncated form

The 3D structure modeling of the ATM protein using SWISS-MODEL shows that this protein has missed a number of its residues including PIKK and FACT domains. The acetylation of the FACT domain in the C-terminal portion is a primary step for the activation of the ATM protein. Thus, deletion of the C-terminal portion could remarkably change its structure, and disrupt its activation during double-strand break (Fig. 2).

3.3. Truncated ATM could not get activated in DNA damage response

Truncated ATM fails to interact with Tip60, as an upstream protein, and p53, as a downstream protein. The novel frameshift variant was examined to determine how it could affect the ATM function through its interactions with other proteins. The EtoT docking scores of the truncated and the normal ATM proteins are significantly different; the interaction of the truncated protein with the target proteins requires much more energy (Fig. 3), leading to disrupted interactions with upstream and downstream proteins such as Tip60 and p53, respectively. Therefore, the dimer-monomer transition does not occur as the result of disrupted interaction with Tip60. Alternatively, ATM cannot phosphorylate its target proteins without its kinase domain, which forms when the C- and N-terminal portions of ATM get together. The net result is the failure of the truncated ATM protein to activate the proteins involved in cell cycle checkpoints and DNA repair processes.

4. Discussion

This study applies the targeted NGS to A-T disorder and reports a new frameshift variant (c.8046_8047delTA, p.Thr2682ThrfsX5) in the ATM gene from a 5-year-old proband presented with A-T clinical features born to a consanguineous family. This new frameshift variant results in early termination in exon 57 to produce a truncated protein. This variant makes the protein lose its C-terminal portion, including PIKK and FACT domains, which are necessary for its activation in response to radiation. The variant was categorized as pathogenic based on the ACMG guideline [19].

Activation of downstream proteins in response to double-strand break depends on ATM activation via dimer-monomer transition following its auto-phosphorylation [22]. Tip60 is an upstream protein in this pathway, which activates ATM, and makes its interaction domain accessible for its target proteins. Dimer-monomer transition occurs through acetylation of ATM in the FATC domain [9,23]. Thus, the interaction disruption could inevitably affect the function of the protein. The mentioned frameshift deletion variant prevents the ATM protein from interacting with Tip60, thereby inhibiting its auto-phosphorylation. Even if the protein can auto-phosphorylate, deletion of the PIKK domain within the ATM protein disables the phosphorylation of target proteins in response to radiation since the process requires the assembly of both C-terminal and N-terminal portions [22].

ATM is a large gene with 66 small exons with no mutational hotspot [13]. In order to save time and energy, we applied an accurate and cost-effective diagnostic method which examines all exons and exon-intron boundaries of ATM genes in a proband with the clinical presentation of A-T. A novel frameshift variant was identified using NGS method located in exon 57 resulting in protein truncation. So far, 12 pathogenic variants related to A-T have been reported in this exon (9 missense variants: Tyr2677Cys, Thr2682Ser, Pro2699Leu, Lys2700Asn, Ile2702Arg, Cys2704Arg, Asp2708Glu, Val2716Ala; one frameshift variant; Ile2702fs; two nonsense variants: Lys2700Ter, Gln2714Ter) [24–32].

Frameshift and nonsense variant are among the majority types of ATM variants [11,28,33,34] constituting about 70% of A-T case [35]. This suggests that the A-T classical phenotype is due to null alleles resulting in total loss of protein function. Truncating variants, which lead to complete absence of ATM kinase activity lead to severe phenotypes [36], while typical [28] or milder [37,38] phenotypes are due to missense variants in the ATM gene.

Cerebellar ataxia is one of the manifestations in A-T as well as some other neurodegenerative disorders. On the other hand, in these neurodegenerative disorders, loss of function of DNA-repair components is common. Therefore, it suggests that a normal cerebellar development needs an intact DNA damage response system [39].

Based on evidence provided in the ACMG guideline, the reports of truncating variants downstream of the extreme 3 prime of an identified truncating variant are confirmatory for its pathogenicity [19]. This novel variant causes the protein to gain early stop codon in exon 57 resulting in A-T classical phenotype. This suggests that it is a loss-of-function variant. Also, there are other null variants, related to the A-T phenotype, downstream of this variant, which result in truncated proteins (c.8266A>T, c.8283-8284delTC, c.8307G>A, c.8373C>A, c.8287C>T, c.8793T>A, c.8833delCT, c.8879G>A, c.9139C>T, c.9026T>G).

In a study in China, 12 novel variants were reported from 8 unrelated families with no presence of a homozygous and founder effect variant. This provides a clue for the diversity of Chinese population and the of ATM variants [40]. In contrast, many founder mutations have been detected in Norwegian, Costa Rican, Polish, Italian and Amish/Mennonite population through haplotype analysis [13]. The similarity of 55%
of ATM variants in 11 Norwegian families with the same haplotype is evidence of founder effect [41]. With the high rate of consanguinity in Iran (38.6%) [42], founder ATM variants are predicted to exist more frequently, and this could facilitate the diagnostics by developing a rapid and cost-effective method to detect the pathogenic variant.

In conclusion, our study expands the spectrum of ATM pathogenic variants in Iran and confirms the utility of targeted NGS sequencing in genetic diagnosis.

Acknowledgement

This work was supported by Ahvaz Noor Genetics Laboratory and Isfahan University of Medical Sciences (grant no. 195122).

References

[2] A. Nissenkorn, Y. Levy-Shraga, Y. Banet-Levi, A. Lahad, I. Sarouk, D. Modan-Moses, Evidence of founder effect [41]. With the high rate of consanguinity in Iran (38.6%) [42], founder ATM variants are predicted to exist more frequently, and this could facilitate the diagnostics by developing a rapid and cost-effective method to detect the pathogenic variant. And this could facilitate the diagnostics by developing a rapid and cost-effective method to detect the pathogenic variant.


This work was supported by Ahvaz Noor Genetics Laboratory and Isfahan University of Medical Sciences (grant no. 195122).


