DETECTION OF c.G2194A MUTATION IN AR GENE OF A VIETNAMESE PATIENT WITH ANDROGEN INSENSITIVITY SYNDROME: A CASE REPORT

Thu Hien Nguyen1,2,3, Lien Nguyen Thi Kim3, Thanh Ngan Nguyen Thi3, Huong Giang Tran Thi3, Hong Nhung Nguyen1, Nguyen Thi Phuong Mai5, Huy Hoang Nguyen3,*

1Graduate University of Science and Technology, VAST, Vietnam
2Forensic Medicine, Military Hospital 103, Vietnam Military Medical University, Vietnam
3Institute of Genome Research, VAST, Vietnam
4Institute of Biotechnology, VAST, Vietnam
5Human Genetics Department, National Hospital of Pediatrics, Vietnam

Received 15 September 2021; accepted 24 May 2022

ABSTRACT

Androgen insensitivity syndrome (AIS) is a rare X-linked recessive androgen receptor (AR) disorder. However, the overlap in clinical manifestations between AIS and other disorders of sex development can cause clinical diagnostic difficulties. Applying the whole coding region sequencing method is an optimal method for the diagnosis of AIS. In this study, whole-exome sequencing was performed to screen mutations in the AR gene as well as genes related to disorders of sex development (DSD). Sanger sequencing was applied to validate the mutations in the patient. One missense mutation in the AR gene which was reported previously was identified in the patient. In this site, nucleotide G is changed to A at position 2194 on cDNA (c.G2194A), leading to a substitution of aspartic at position 732 aspartic to asparagine (p.Asp732Asn). However, this is the first published case in a Vietnamese with this mutation. Our study expands the mutation spectrum of the AR gene in Vietnamese patients and confirms the usefulness of whole-exome sequencing in the diagnosis of AIS. The results of the study are the basis for supporting doctors in prenatal diagnosis and giving reasonable advice to patients and families.

Keywords: Androgen insensitivity syndrome, AR gene, disorders of gender development.


*Corresponding author email: nhhoang@igr.ac.vn

©2022 Vietnam Academy of Science and Technology (VAST)
INTRODUCTION

Androgen insensitivity syndrome (AIS, OMIM #300068) is the most common of 46 XY disorders of gender development (DSD). The prevalence of AIS proven via molecular diagnosis is estimated to range from 1:20,000-64,000 in XY (chromosomally male) births. The severity of the disease depends on the structure and sensitivity of the abnormal androgen receptor, which can manifest under three phenotypes: complete androgen insensitivity syndrome (CAIS), partial androgen insensitivity syndrome (PAIS), and mild androgen insensitivity syndrome (MAIS) (Yuan et al., 2018). Mainly the body of patients with AIS is not able to respond in whole or in part to male sex hormones. Therefore, they may have mostly female external sex characteristics or signs of both male and female sexual development.

The AR gene encodes for the protein androgen, also known as the androgen receptor which is located on chromosome Xq11-12. This receptor belongs to the steroid hormone group of nuclear receptors with the estrogen receptor (ER), glucocorticoid receptor (GR), progesterone receptor (PR) and mineralocorticoid receptor (MR) (Modi et al., 2016). The AR receptor is a ligand-dependent transcription factor, including three functional domains as the other nuclear receptors: a large N-terminal domain (NTD) (1-555 residues), a DNA-binding domain (DBD) (556-623 residues), a hinge domain (624-665 residues) and a C-terminal ligand-binding domain (LBD) (666–919 residues). In some cell types, testosterone directly interacts with the androgen receptor, while in others, testosterone is metabolized by 5-alpha-reductase to dihydrotestosterone (DHT), a more potent factor that can activate androgen receptors (Davison & Bell, 2006). The binding of the AR to its native ligands 5α-dihydrotestosterone (DHT) and testosterone initiates male sexual development and differentiation (Tan et al., 2015).

Mutations in the AR gene cause androgen insensitivity syndrome, in which, about 30% of reported mutations in AIS were de novo (Batista et al., 2018). Mutations can be found throughout the AR gene, but the distribution of mutations varies in different forms of AIS. For CAIS patients, variants were usually detected on the NTD domain, while mutations in the LBD domain (exon 5, 6) were often present in PAIS patients. Mutations in MAIS were generally less phenotypically related and commonly found in the NTD domain (Yuan et al., 2018).

In this study, we detected a known mutation in a patient with disorders of sex development by the whole exome sequencing method. This finding elucidated the causes of clinical manifestations and helps doctors provide advice to patients and families.

MATERIALS AND METHODS

Patient

The patient was born in 2001 and was determined to have a female appearance after birth. In 2015, the patient visited at the National Children’s Hospital for examination due to delayed menstruation at puberty. The patient was 1m78 tall and completely female in appearance. The patient was assigned to a female gender due to ultrasound. In the family on the mother’s side, there are two younger sisters who had both female appearance and amenorrhea. The patient was clinically diagnosed with androgen insensitivity syndrome. The patient who participated in the study was agreed by his parents. In this study, the patient is denoted as A-KTH.

Whole exome sequencing and Sanger sequencing

The QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) was used to extract the DNA from the whole blood of the patient. The DNA library was prepared according to the instructions of the Agilent SureSelect Target Enrichment Kit (Agilent Technologies, CA, USA) and sequenced on Illumina sequencer (Illumina, CA, USA). The
bioinformatics tools used for data analysis were Burrows-Wheeler Aligner tool v0.7.12 (Li & Durbin, 2009), Picard tool v1.130 (http://broadinstitute.github.io/picard/), Genome Analysis Toolkit v3.4.0 (McKenna et al., 2010) and SnpEff v4.1 (http://snpeff.sourceforge.net/SnpEff.html) (Nguyen et al., 2021). Then, candidate variants in the pathogenic genes related to DSD (AR, SRY, NR5A1, SRD5A2, HSD17B3, HSD17B2, MAMLD1, NROB1, CYP11A1, CYP21A2, CYP17, CYP11B1, and CYP19A1) were screened.

Sanger sequencing was performed for validation of the variants of interest identified in exome analysis using BigDye v3.1 reagent (Applied Biosystems).

**In silico analyses and pathogenic interpretation**

The pathogenicity of candidate mutations was predicted by some tool as Functional Analysis through Hidden Markov Models (FATHMM; http://fathmm.biocompute.org.uk/inherited.html), Mutation Taster (http://www.mutationtaster.org/), PANTHER (http://www.pantherdb.org/tools/hmmScoreForm.jsp), Pathological Mutations on proteins (Pmut; http://mmb.irbbarcelona.org/PMut/analyses/new), functional Analysis through Hidden Hidden Markov Models (FATHMM; http://fathmm.biocompute.org.uk/inherited.html). Multiple alignments of amino acid sequences of AR of multiple species were performed using multiple sequence alignment in Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) in order to determine the conservation level of amino acids at the position of substitution. 3D structural simulation model of the protein is made by Swiss-PdbViewer 4.1.0 software.

The study had been approved by the ethical committee of IGR IRB No: 05-2020/NCHG-HDDD, October 2, 2020.

**RESULTS AND DISCUSSION**

Whole exome sequencing and Sanger analysis

The Illumina sequencer yielded a total of 77,848,976 read bases (bp) with 46.7% GC. The percentages of Q20 and Q30 were 97.5 and 90.1, respectively. After processing the data with bioinformatics software, 61,215,624 on-targeted reads accounting for 78.6% of the raw data were identified (Table 1).

<table>
<thead>
<tr>
<th>Sample</th>
<th>Total Reads</th>
<th>%GC</th>
<th>Q20 (%)</th>
<th>Q30 (%)</th>
<th>Initial Mappable Reads</th>
<th>Non-Redundant Reads</th>
<th>On-Target Reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-KTH</td>
<td>77,848,976</td>
<td>46.7</td>
<td>97.5</td>
<td>90.1</td>
<td>77,768,985 (99.9%)</td>
<td>75,238,862 (96.4%)</td>
<td>61,215,624 (78.6%)</td>
</tr>
</tbody>
</table>

Note: Initial Mappable Reads: Number of mapped reads to the human genome; Non-Redundant Reads: Number of de-duplicate reads from Picard tools; On-Target Reads: Number of reads mapped to target regions.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Patient A-KTH</th>
</tr>
</thead>
<tbody>
<tr>
<td># of SNP</td>
<td>98.236</td>
</tr>
<tr>
<td>Synonymous Variant</td>
<td>11.743</td>
</tr>
<tr>
<td>Missense Variant</td>
<td>10.540</td>
</tr>
<tr>
<td>Stop Gained</td>
<td>95</td>
</tr>
<tr>
<td>Stop Lost</td>
<td>36</td>
</tr>
<tr>
<td># of INDEL</td>
<td>12.875</td>
</tr>
<tr>
<td>Frameshift Variant</td>
<td>273</td>
</tr>
<tr>
<td>Inframe Insertion</td>
<td>164</td>
</tr>
<tr>
<td>Inframe Deletion</td>
<td>227</td>
</tr>
</tbody>
</table>

The variants were called and annotated by using Genome Analysis Toolkit v3.4.0 (McKenna et al., 2010) and SnpEff v4.1 (http://snpeff.sourceforge.net/SnpEff.html) (Table 2). Among them, only a missense mutation in the AR gene (c.2194G>A; p.Asp732Asn) was determined as disease-causing mutation.

Verification results by Sanger sequencing showed that missense mutations occurred in exon 5 of the AR gene (Fig. 1). This mutation
changes (guanine) G at nucleotide 2914 to adenine (A) in cDNA, leading to a replacement of Asp with Asn at the 732 positions in AR protein.

**Figure 1.** Mutation in the AR gene in patient A-KTH. a. Chromosome X; b. Location of mutation c.2194 G>A (p.Asp732Asn) in the AR gene; c. A normal allele G at the position 2194 of cDNA (left panel) in a healthy person and a missense mutation c.2194 G>A at the homozygous state in the patient A-KTH (right panel)

**In silico analyses and pathogenic predictions**

Mutation p.Asp732Asn was predicted to affect the function of the protein or has the potential to cause disease by 8 different bioinformatics tools (Table 3). In detail, the mutation p.Asp732Asn was predicted to be “disease causing” by the Mutation Taster; “Probably damaging” by the PolyPhen-2, PANTHER; “Disease” by Pmut, PhD-SNP analyses; “Deleterious” by the PROVEAN analyses; “Effect” by the SNAP2 analyses, “Damaging” by Fathmm analyses and “Pathogenic” by the Pon-2 analyses (Table 3). Each tool is evaluated based on the different criteria, therefore, the consistent results show the reliability of the prediction.

**Table 3. In Silico Analyses**

<table>
<thead>
<tr>
<th>Tool</th>
<th>Predictions</th>
<th>Prediction Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutation Taster</td>
<td>Disease causing</td>
<td>0.99</td>
</tr>
<tr>
<td>Pmut</td>
<td>Disease</td>
<td>0.91</td>
</tr>
<tr>
<td>PolyPhen 2</td>
<td>Probably damaging</td>
<td>0.96</td>
</tr>
<tr>
<td>Panther</td>
<td>Possibly damaging</td>
<td>0.5</td>
</tr>
<tr>
<td>Fathmm</td>
<td>Damaging</td>
<td>-7.99</td>
</tr>
<tr>
<td>Pon-P2</td>
<td>Pathogenic</td>
<td>0.868</td>
</tr>
<tr>
<td>PhD-SNP</td>
<td>Disease</td>
<td></td>
</tr>
<tr>
<td>PROVEAN</td>
<td>Deleterious</td>
<td>-3.357</td>
</tr>
<tr>
<td>Snap2</td>
<td>Effect</td>
<td>20</td>
</tr>
</tbody>
</table>
Predicting the impact of mutations on 3D structural simulation models

The mutant was built with a 3D structure based on the published structural model on the Protein Data Bank (PDB), code pdb3v49. By using Swiss-PdbViewer 4.1.0 software, the amino acid aspartic was converted to the amino acid asparagine (Fig. 2).

![Figure 2. 3D structural simulation model. A. wild-type protein (p.Asp732); B. mutant protein (p.Asn732). Helix α-helix in blue, folded β-sheet in green, and coil structure in green. The red end of the amino acid has the carboxyl group and the blue end has the amino group. The dashed green line shows hydrogen bonding.](image)

From the 3D structural model, the amino acid Asp732 is located in an α-helix chain of the AR protein. In addition to peptide bonds with adjacent amino acids, Asp732 is also linked with further amino acids by hydrogen bonds. Specifically, Asp732 binds to His729, Ala735 and Val736 on the same α-helix. Besides, Asp732 also binds to Lys822 and Asn823 on the coil structure. When the mutation occurs, the number of hydrogen bonds of Asn732 does not change. However, as Asn has only one carboxyl group, the hydrogen bond of Asn732 with Lys822 and Asn823, is concentrated on one end of Asn732. Furthermore, aspartic is acidic, negatively charged, and asparagine is a neutral amino acid, not charged. Due to different chemical properties, when mutations occur, it will affect the structure and function of the protein.

The androgen receptor (AR/NR3C4) belongs to the steroid receptor subfamily of nuclear receptors (NRs) which pivotal role in a number of endocrine-related diseases. According to previous studies, a large number of mutations in the AR gene have been associated with castration-resistant prostate cancer. However, mutations in the AR gene are known more as a cause of androgen insensitivity syndromes which are classified as complete (CAIS), partial (PAIS) or mild (MAIS) has depended on the clinical phenotype (Nadal et al., 2017). Due to the clinical similarity between AIS and several other DSD diseases, especially the ambiguous phenotypes of PAIS and MAIS, whole-genome sequencing is an effective option for diagnosis correct and explain the cause of the disease (Wang et al., 2017).

To date, over 1,000 mutations in AR have been reported in the AIS and prostate cancer-related, of these more than 900 mutations have been associated with AIS. Furthermore, a majority of these mutations are missense
The patient was 15 years old and had no signs of menstruation, the external genitalia was female, the patient had feminine behavior with mentally completely normal. The clinical results showed that the patient was a karyotype 46, XY with testis determining factor and diagnosed with CAIS. The p.Asp732Asn mutation was detected on exon 5 of the AR gene in the patient, this is a reported mutation. The predictive results of the in silico analysis showed a significant level of damage of this mutation to AR protein function. These are also highly reliable bases that have been used by many studies to predict the impact of new mutations (Nguyen et al., 2020; Wang et al., 2017).

On the other hand, when comparing the AR protein sequences of human with other proteins species, the results showed Asp732 residue is a highly conserved site. The longer a position has been preserved, the more likely that it will have a deleterious effect.

Interestingly, the mutant protein's activity was reduced when androgen concentration was low and at high androgen concentration, the functional defect of the mutant protein is ameliorated (Hannema et al., 2004; Vija et al., 2014; Wu et al., 2010). Functional studies have shown this mutation reduced 95% protein activity compared with the wild type (Hannema et al., 2004). Furthermore, the highly conserved loop between helices 3 and 4 of nuclear receptors, is directly involved in coactivator recruitment (Eisermann et al., 2013). Asp732 is located in the helix 4, therefore, the mutation might affect the recruitment ability of AR protein. Taking together, the p.Asp732Asn mutation is the cause AIS expression of the patient A-KTH.

Interestingly, the mutant protein's activity was reduced when androgen concentration was low and at high androgen concentration, the functional defect of the mutant protein is ameliorated (Hannema et al., 2004). This is very important and has great practical significance in research to find treatment options and improve the condition of the patient.

The p.Asp732Asn mutation in the patient A-KTH could be a de novo mutation or inherited from the mother. However, it is difficult to determine the genetic cause in this patient's family because the patient's parents did not agree to conduct further. At the time of diagnosis, the patient was 8 years old, when she was about to begin puberty. Therefore, this diagnostic result is very meaningful in helping doctors have appropriate interventions and genetic counselling for the patient’s family. From there, parents and families will have the right direction of nurturing and gender orientation for their children.

CONCLUSION

We identified a missense mutation in the AR gene in a Vietnamese patient by using whole-exome sequencing. The patient’s phenotype is consistent with the expression in
Detection of c.G2194A mutation in AR gene

patients with mutations in the previously published AR gene, especially those with the p.Aspg32Asn mutation. These results lead to an accurate diagnosis of the patient with AIS. Our findings demonstrate that whole-exome sequencing is a useful diagnostic tool for AIS.

Acknowledgements: The authors thank the family for their participation in this study. This research was funded by the Graduate University of Science and Technology, Vietnam Academy of Science and Technology, grant number GUST.STS.DT2020-SH01.

REFERENCES


