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

Thu Hien Nguyen[1,2,3], Lien Nguyen Thi Kim[3], Thanh Ngan Nguyen Thi[3], Huong Giang Tran Thi[3], Hong Nhung Nguyen[4], Nguyen Thi Phuong Mai[5], Huy Hoang Nguyen[3,*]

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

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* gen, disorders of gender development.

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 able to 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 for chromosomal analysis, testicular determining factor and family history. The patient had a 46, XY and had a testicular differentiation factor. The uterus and ovaries were not detected by 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 (Agi- lent 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, NROBI, 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 mutation was predicted by some tool as Functional Analysis through Hidden Markov Models (FATHMM; http://fathmm.biocompu-te.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 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-HĐĐĐ, 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-targed reads accounting for 78.6% of the raw data were identified (Table 1).

Table 1. Pre-alignment statistics

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

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 2. Statistic group variants

SNP	Patient A-KTH
# of SNP	98.236
Synonymous Variant	11.743
Missense Variant	10.540
Stop Gained	95
Stop Lost	36
# of INDEL	12.875
Frameshift Variant	273
Inframe Insertion	164
Inframe Deletion	227

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.
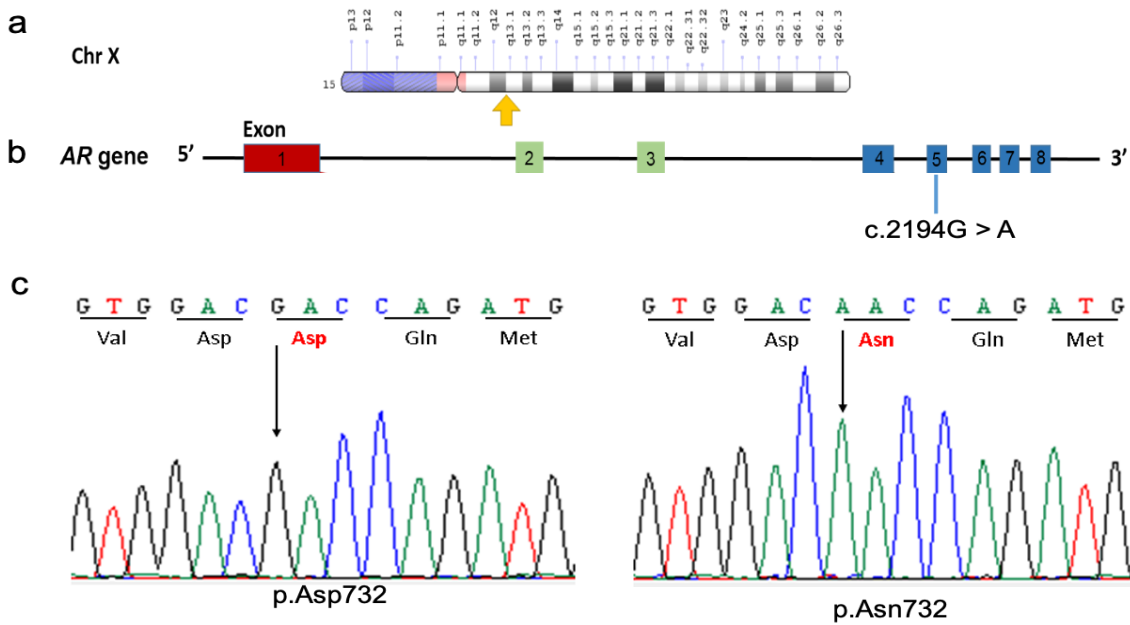

Figure 1. Mutation in the *AR* gene in patient A-KTH. a. Chromosome X; b. Location of mutation c.2194 G>A (p.Asp732Asn) in the *AR* gene; c. A normal allele G at the position 2194 of cDNA (left panel) in a healthy person and a missense mutation c.2194 G>A at the homozygous state in the patient A-KTH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

Tool	Predictions	Prediction Score
Mutation Taster	Disease causing	0.99
Pmut	Disease	0.91
PolyPhen 2	Probably damaging	0.96
Panther	Possibly damaging	0.5
Fathmm	Damaging	-7.99
Pon-P2	Pathogenic	0.868
PhD-SNP	Disease	
PROVEAN	Deleterious	-3,357
Snap2	Effect	20

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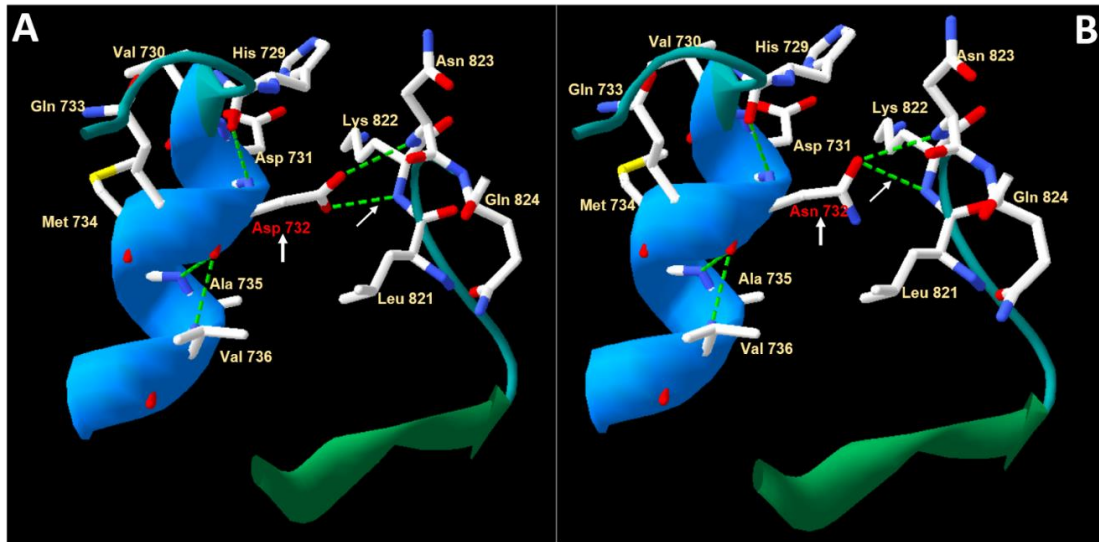
Figure 2. 3D structural simulation model. A. wild-type protein (p.Asp732); B. mutant protein (p.Asn732). Helix α-helix in blue, folded β-sheet in green, and coil structure in green. The red end of the amino acid has the carboxyl group and the blue end has the amino group. The dashed green line shows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

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

mutations which located in the LBD of the AR, see also McGill AR database (www.mcgill.ca/androgendb). The structure of LBD includes 11 α-helices and four short β strands. In which, two-stranded β-sheets are arranged in a typical three-layer antiparallel helical sandwich fold. The LBD, encoded by the distal part of exon 4 and exons 5–8, binding of ligand to the AR LBD causes a significant positional change in helix 12 (Tan et al., 2015).

Somatic missense mutations in the LBD usually lead to reduce specificity of AR to other hormones such as progesterone, estrogens and adrenal steroids, and ligand affinity and coregulator recruitment also have been affected (Eisermann et al., 2013). Therefore, the mutations that have been occurring in AR LBD have a significant effect on AR function (Rosa et al., 2002). An issue of current concern is the rapid emergence of drug resistance caused by mutations in the AR ligand binding region (LBD), which presents a major challenge for the drug industry to used AR antagonists. Therefore, the study of mutations in LBD has an important role to discover new effective anti-AR treatments (Pang et al., 2021).

At the time of medical examination, the patient was 15 years old and had no signs of menstruation, the patient had a completely feminized appearance, 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.

The mutation c.2194G > A, p.Asp732Asn has been reported in previous studies in patients with CAIS, which has been demonstrated to be deleterious (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 novel* 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

patients with mutations in the previously published *AR* gene, especially those with the p.Asp732Asn 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.

REFERENCES

Batista R. L., Costa E. M. F., Rodrigues A. de S., Gomes N. L., Faria J. A., Nishi M. Y., Arnhold I. J. P., Domenice S., Mendonca B. B de, 2018. Androgen insensitivity syndrome: a review. *Arch. Endocrinol. Metab.*, 62: 227–235. https://doi.org/10.20945/2359-3997000000031

Davison S. L., Bell R., 2006. Androgen physiology. *Semin. Reprod. Med.*, 24: 71–77. https://doi.org/10.1055/s-2006-939565

Eisermann K., Wang D., Jing Y., Pascal L. E., Wang Z., 2013. Androgen receptor gene mutation, rearrangement, polymorphism. *Transl. Androl. Urol.*, 2: 137. https://doi.org/10.3978/j.issn.2223-4683.2013.09.15

Hannema S. E., Scott I. S., Hodapp J., Martin H., Coleman N., Schwabe J. W., Hughes I. A., 2004. Residual activity of mutant androgen receptors explains wolffian duct development in the complete androgen insensitivity syndrome. *J. Clin. Endocrinol. Metab.*, 89: 5815–5822. https://doi.org/10.1210/jc.2004-0709

Li H., Durbin R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinforma. Oxf. Engl.*, 25: 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

McKenna A., Hanna M., Banks E., Sivachenko A., Cibulskis K., Kernytsky A., Garimella K., Altshuler D., Gabriel S., Daly M., DePristo M. A., 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.*, 20: 1297–1303. https://doi.org/10.1101/gr.107524.11

Modi P. K., Faiena I., Kim I. Y., 2016. Chapter 3 - Androgen Receptor, in: Mydlo J. H., Godec C. J. (Eds.), Prostate Cancer (Second Edition). *Academic Press, San Diego*, pp. 21–28. https://doi.org/10.1016/B978-0-12-800077-9.00003-7

Nadal M., Prekovic S., Gallastegui N., Helsen C., Abella M., Zielinska K., Gay M., Vilaseca M., Taulès M., Houtsmuller A. B., van Royen M. E., Claessens F., Fuentes-Prior P., Estébanez-Perpiñá E., 2017. Structure of the homodimeric androgen receptor ligand-binding domain. *Nat. Commun.*, 8: 14388. https://doi.org/10.1038/ncomms14388

Nguyen H. H., Khanh Nguyen N., Dung Vu C., Thu Huong Nguyen T., Nguyen N.-L., 2020. Late-Onset Ornithine Transcarbamylase Deficiency and Variable Phenotypes in Vietnamese Females With OTC Mutations. *Front. Pediatr.*, 8: 321. https://doi.org/10.3389/fped.2020.00321

Nguyen T. H., Nguyen N. L., Vu C. D., Ngoc C. T. B., Nguyen N. K., Nguyen H. H., 2021. Identification of three novel mutations in PCNT in vietnamese patients with microcephalic osteodysplastic primordial dwarfism type II. *Genes Genomics*, 43: 115–121. https://doi.org/10.1007/s13258-020-01032-5

Pang J. P., Shen C., Zhou W. F., Wang Y. X., Shan L. H., Chai X., Shao Y., Hu X. P., Zhu F., Zhu D. Y., Xiao L., Xu L., Xu X. H., Li D., Hou T. J., 2021. Discovery of novel antagonists targeting the DNA binding domain of androgen receptor by integrated docking-based virtual screening and bioassays. *Acta Pharmacol. Sin.*,. https://doi.org/10.1038/s41401-021-00632-5

Rosa S., Biason-Lauber A., Mongan N. P., Navratil F., Schoenle E. J., 2002. Complete Androgen Insensitivity Syndrome Caused by a Novel Mutation in the Ligand-Binding Domain of the Androgen Receptor: Functional Characterization. *J. Clin. Endocrinol. Metab.*, 87: 4378–4382. https://doi.org/10.1210/jc.2002-020139

Tan M. E., Li J., Xu H. E., Melcher K., Yong E., 2015. Androgen receptor: structure, role in prostate cancer and drug discovery. *Acta Pharmacol. Sin*., 36: 3–23. https://doi.org/10.1038/aps.2014.18

Vija L., Ferlicot S., Paun D., Bry-Gauillard H., Berdan G., Abd-Alsamad I., Lombès M., Young J., 2014. Testicular histological and immunohistochemical aspects in a post-pubertal patient with 5 alpha-reductase type 2 deficiency: case report and review of the literature in a perspective of evaluation of potential fertility of these patients. *BMC Endocr. Disord.*, 14: 43. https://doi.org/10.1186/1472-6823-14-43

Wang Z., Zhou Y., Yang R., Xia M., Zeng H., Du L., Ren J., Guo Q., 2017. Whole Exome Sequencing Identifies a c.C2566T Mutation in the Androgen Receptor in a Chinese Family. *Clin. Labi*, 63. https://doi.org/10.7754/Clin.Lab.2017.170316

Wu X., Zhou Q., Mao J., Lu S., Nie M., 2010. Mutational analysis of androgen receptor gene in four Chinese patients with male pseudohermaphroditism. *Fertil. Steril.*, 93: 2076.e1–4. https://doi.org/10.1016/j.fertnstert.2009.11.025

Yuan S. M., Zhang Y. N., Du J., Li W., Tu C. F., Meng L. L., Lin G., Lu G. X., Tan Y. Q., 2018. Phenotypic and molecular characteristics of androgen insensitivity syndrome patients. *Asian J. Androl.*, 20: 473–478. https://doi.org/10.4103/aja.aja_17_18