Clin. Lab. 2017;63:XXX-XXX

SHORT COMMUNICATION

Whole Exome Sequencing Identifies a c.C2566T Mutation in the *Androgen Receptor* in a Chinese Family

Zengge Wang [1, 3, *], Yulin Zhou [1, *], Ruijuan Yang [1, *], Zhongmin Xia [1, *], Huan Zeng [1], Liya Du [1], Jun Ren [2], Qiwei Guo [1]

[*] *These authors contribute equally to this work*
[1] *United Diagnostic and Research Center for Clinical Genetics, School of Public Health of Xiamen University & Xiamen Maternal and Child Health Hospital, Xiamen, 361102 Fujian, China*
[2] *Department of Dermatology, Zhongshan Hospital Xiamen University, Xiamen, 361003 Fujian, China*
[3] *Medical College of Xiamen University, Xiamen, 361102 Fujian, China*

SUMMARY

Background: **Whole exome sequencing (WES) is one of the most valuable tools for the detection of Mendelian diseases in clinical laboratory. We performed WES for a family of 46,XY disorders of gender development and compared the applicability of public databases for the subsequent phenotype studies of WES-identified mutations.**
Methods: **DNA samples from the two patients were analyzed by WES. The mutated protein was studied using the HomoloGene database, Polyphen2, and SIFT. The phenotype of the mutation was studied using ClinVar, the androgen receptor gene mutations database, AR database at Leiden Open Variation Database, and PubMed.**
Results: **A c.C2566T (p.R856C) mutation in the androgen receptor gene was detected for the patients. The *in silico* studies indicated that the p.R856C mutation is deleterious to the function of the androgen receptor. Unlike those of other databases, the variations listed in the androgen receptor gene mutations database were classified as complete androgen insensitivity-, partial androgen insensitivity-, or mild androgen insensitivity-relevant according to their clinical phenotype. In addition, the publications of the collected mutations in the androgen receptor gene mutations database are complete and easily accessible, which facilitates in depth studies of clinically identified mutations.**
Conclusions: **We identified a c.C2566T (p.R856C) mutation of the *AR* gene in cases of familial complete androgen insensitivity by WES, and provided genetic counseling to related family members. This is the first study reporting this mutation in Chinese patients. We also compared the applicability of several public databases for phenotype studies of clinically identified Androgen Receptor mutations and suggest that the androgen receptor gene mutations database best satisfies clinical demands.**
(Clin. Lab. 2017;63:xx-xx. DOI: 10.7754/Clin.Lab.2017.170316)

Correspondence:
Qiwei Guo
United Diagnostic and Research Center for Clinical Genetics
School of Public Health of Xiamen University & Xiamen Maternal and Child Health Hospital
Xiamen, 361102 Fujian
China
Phone: +86 05927805028
Email: guoqiwei@gmail.com

Jun Ren
Department of Dermatology
Zhongshan Hospital Xiamen University
Xiamen, 361004 Fujian
China
Phone: +86 05927805029
Email: renjun666@126.com

Short Communication accepted April 3, 2017

KEY WORDS

androgen insensitivity syndrome, androgen receptor gene mutations database (ARDB), c.C2566T, p.R856C, whole exome sequencing

INTRODUCTION

Androgen insensitivity syndrome (AIS, OMIM #300068) is the most common, identifiable cause of 46,XY disorders of gender development (DSD); its prevalence is 1 in 20,000 - 64,000 male births [1-3]. According to its clinical features, AIS can be classified into three subgroups: complete (CAIS), partial (PAIS), and mild (MAIS) [4].

Genetically, AIS is an X-linked recessive disorder caused by mutations in the androgen receptor (*AR*) gene, resulting in complete or partial androgen unresponsiveness. *AR* is a single-copy gene on the Xq12 locus; it is more than 90 kb in length and contains eight exons that encode a protein of 920 amino acids [5,6]. AR functions as a steroid hormone-activated transcription factor that contains four major functional domains: the N-terminal domain (NTD), DNA-binding domain (DBD), C-terminal ligand-binding domain (LBD), and a hinge region connecting the LBD with the DBD [7]. Two alternatively spliced variants encoding distinct isoforms of AR have been described. Approximately 500 germline mutations in the *AR* gene that completely or partially inactivate the AR protein and cause AIS have been reported [6]. A diagnosis of AIS in an individual with 46,XY DSD requires the exclusion of other etiologies that display clinical features similar to AIS [8]. Therefore, molecular analysis of the *AR* gene, specifically sequence analysis, is crucial for diagnosis of AIS. Identification of a specific *AR* mutation by sequencing not only allows precise diagnosis and/or prognosis, but also facilitates genetic counseling of related family members [8].

Based on our current knowledge of genetics, whole exome sequencing (WES) is one of the most valuable tools for clinical analysis of rare Mendelian diseases owing to its relatively low cost, high detection throughput for genetic variants and sample size, high yields for clinical variants (more than 25%), and the ability to interpret the results and allow a clinical diagnosis [9]. In this study, we identified a c.C2566T mutation of the *AR* gene in a Chinese family with familial CAIS by WES, and examined other family members for this mutation. In addition, we compared the applicability of public databases for phenotype studies of clinically identified *AR* mutations.

MATERIALS AND METHODS

Patients and clinical evaluation

The proband (family member III-3) was a 16-year-old Chinese woman who visited the Department of Gynecology for primary amenorrhea. Physical examination revealed an apparently normal female phenotype with normal breast development, scarce pubic and axillary hair, normal external genitalia, but no observable cervix. Ultrasonography revealed the absence of a uterus and presence of normal-sized bilateral gonads (size: left, 20 x 12 mm; right, 27 x 12 mm). As shown in Table 1, serum hormone test results showed elevated testosterone (T), dihydrotestosterone (DHT), and T/DHT levels. No other hormone abnormalities were detected. The karyotype was male 46,XY. According to the clinical features, the patient was diagnosed with CAIS. However, due to the elevated T/DHT ratio, the patient was sent to our center for further confirmation.

The proband's 20-year-old sister had similar clinical characteristics. When she was 17-years old, she visited the Department of Gynecology for primary amenorrhea. Ultrasonography revealed the absence of a uterus and presence of normal-sized bilateral gonads (size: left, 12 x 9 mm; right, 13 x 11 mm). The serum hormone test results are shown in Table 1. Her karyotype was also male 46,XY. Since the DHT test was not available at the hospital, she was diagnosed with DSD rather than AIS. The intra-abdominal gonads were subsequently removed by laparoscopic surgery, and histological examination revealed that they consisted of testicular tissue.

Sequencing analysis

We collected peripheral blood samples from three generations of the proband's family. Genomic DNA was extracted from 200 μL of each blood sample, using the Super/HF16 plus DNA Extraction System (MagCore, Xiamen, China) according to the manufacturer's protocol. DNA samples from the two DSD sisters were analyzed by commercial whole exome sequencing (WES; Novogene, Beijing, China). Briefly, the exome sequences from 1.0 μg of genomic DNA were enriched using a liquid capture system (Agilent SureSelect Human All Exon V5; Agilent Technologies, Santa Clara, CA, USA) according to the manufacturer's protocol and sequenced on a Hiseq4000 (Illumina) for paired-end 150-bp reads. Finally, a mean exome coverage of 100 x was obtained, allowing for examination of the selected region and sufficient depth to accurately match 99% of the targeted exome.

A suspected clinically relevant variation was detected in the WES. We then examined the genome of related family members for this variation, using a commercial Sanger sequencing service (Mingbo, Xiamen, China). The sequences of the amplification primers were: forward 5′-CCAAGTAGATGGTTCCCTGT-3′ and reverse 5′-AACTGCTCTGTTCTCAACTCC-3′. The forward primer was also used as the sequencing primer.

In silico analysis

The corresponding amino acid change was studied using the HomoloGene database (http://www.ncbi.nlm.nih.gov/homologene) [10], Polyphen2 (http://genetics.bwh.harvard.edu/pph2/) [11], and SIFT (http://sift.jcvi.org/) [12].

The phenotype-genotype correlation of the suspected clinically-relevant variation was studied using ClinVar (http://www.ncbi.nlm.nih.gov/clinvar) [13], the androgen receptor gene mutations database (ARDB, http://androgendb.mcgill.ca/) [6], AR database at Leiden

Table 1. Results of serum hormone tests.

	III-1	III-3	Reference range
FSH (mIU/mL)	5.19	12.13	Mid-ovulation: 3.85 - 8.78 Peak of mid-period: 4.54 - 22.54 Mid-luteal phase: 1.79 - 5.12
LH (mIU/mL)	40.02	35.93	Mid-ovulation: 2.12 - 10.89 Peak of mid-period: 19.18 - 103.03 Mid-luteal phase: 49 - 291
T (ng/mL)	5.46	6.84	0.1 - 0.75
DHT (pg/mL)	Not available	242.67	23.5 - 116
T/DHT	Not available	28.38	< 25

Table 2. Clinical phenotype of the p.R856C mutation in public databases.

Database	ClinVar	ARDB	LOVD	PubMed
Phenotype-genotype correlation	Pathogenic: 117			
	Likely pathogenic: 1	CAIS: 307	Unclassified	
	Benign: 3	PAIS: 139		
	Uncertain significance: 4	MAIS: 43		
	Conflicting interpretations: 3			
	Total: 128	Total: 489		
No. of publications reporting a p.R856C mutation in different continents				
Europe		6	7	0
Asia		2	1	0
North America		4 [23 - 26]	4	1
South America		2	3	2
Total		14	15	3

Open Variation Database (LOVD, http://grenada.lumc.nl/LOVD2/MR/home.php?select_db=AR) [14], and PubMed (http://www.ncbi.nlm.nih.gov/pubmed) [10].

Informed consent

Written informed consent was obtained from the patients and their family members for conducting the genetic tests and publishing the research data. The study protocol was approved by the ethics committee of the Xiamen Maternal and Child Health Hospital.

RESULTS

AR mutation analysis of the family

A cytidine to thymidine change at position 2,566 of the coding sequence of the *AR* gene (c.C2566T), with a corresponding arginine to cystine change in the protein sequence (p.R856C), was detected by WES for the DNA samples from patients III-1 and III-3. It should be noted that the nucleotide and amino acid numbering is based on NCBI reference sequence NM_000044.2; using GenBank mRNA sequence M20132.1 as a reference, the corresponding nucleotide and amino acid changes were c.C2563T and p.R855C [6]. The mutation was confirmed by Sanger sequencing (Figure 1). Related family members were also examined by Sanger sequencing for this mutation (Figure 1). The sequencing results revealed that the mutation found in the two patients was inherited from their mother (II-2), and the mutation was a *de novo* variation because it was absent in the genome of the grandmother (I-2). The sister of the affected patients (III-2) did not inherit the mutant allele (Figure 2).

Functional prediction of the p.R856C protein

Analysis of the AR protein, using the HomoloGene database, showed that the LBD is highly conserved between species, especially Arg856 (Figure 3). The func-

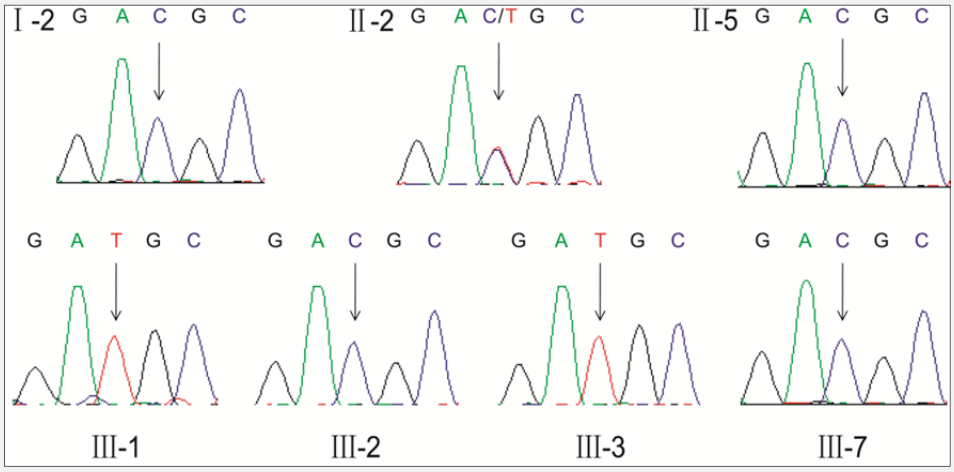

Figure 1. Sanger sequencing results of related family members.

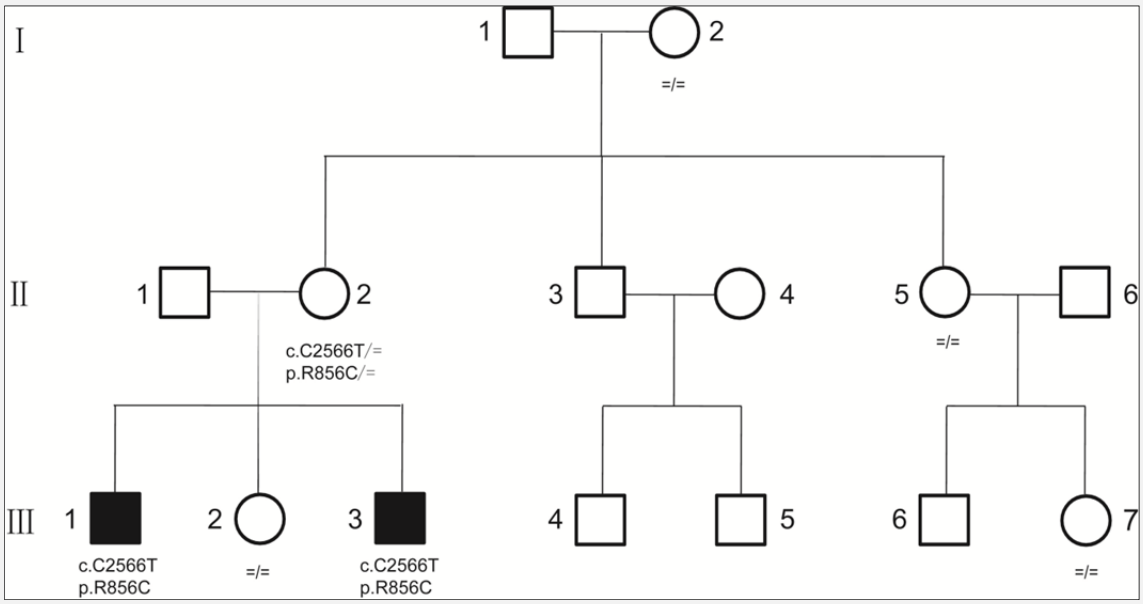

Figure 2. Pedigree analysis of the family affected with CAIS.

Black squares represent affected patients, and open circles and squares represent unaffected females and males, respectively.

tional effects of the p.R856C mutation were predicted considering sequence conservation, physicochemical differences, and the proximity of the mutation to the predicted functional domains, using PolyPhen, and the mutation was predicted to be probably damaging, with a score of 1.000. The functional importance of the

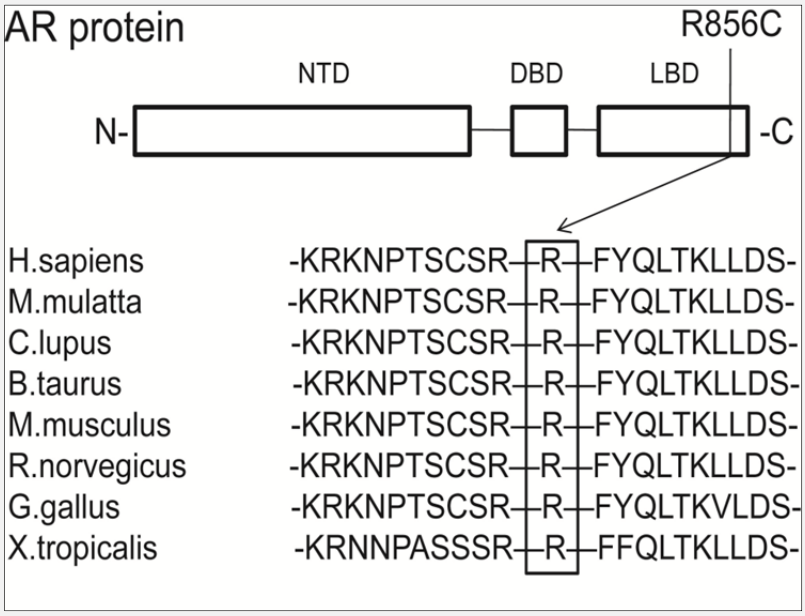

Figure 3. Conservation of AR Arg856 across species.

p.R856C mutation was predicted based on an alignment of orthologous and/or paralogous protein sequences, using SIFT, and the mutation was predicted to be "not tolerated". Therefore, these analyses indicated that the p.R856C mutation is deleterious to the function of the AR protein.

Comparison of the applicability of public databases for the study of clinical phenotype of the p.R856C mutation

We studied the clinical phenotype-genotype correlation of the p.R856C mutation, using four public databases: ClinVar, ARDB, LOVD, and PubMed. As shown in Table 2, a total of 128 clinically relevant variations were found in the ClinVar database, using "androgen receptor" as the keyword. According to the clinical phenotypes, the variations were classified as pathogenic, likely pathogenic, benign, uncertain significance, or conflicting interpretations; however, the p.R856C mutation was not found in this database. In contrast, more than 600 clinically relevant variations were found in ARDB, of which 489 are AIS-relevant. The AIS-relevant variations were classified as CAIS-, PAIS-, or MAIS-relevant according to their clinical phenotype. There were 613 unique variants found in the LOVD database; however, the variants were not classified according to clinical phenotype.

Excluding conference abstracts, there were 14 and 15 publications referring to the p.R856C mutation found in the ARDB and LOVD databases, respectively. The p.R856C mutation was reported in different continents and populations around the world. Most of the publications in the ARDB and LOVD databases could be retrieved from PubMed; however, only three publications were retrieved using "androgen receptor, R855C" as keywords when directly searching PubMed (Table 2).

DISCUSSION

According to the ARDB, approximately 500 AIS-relevant variations have been detected in the *AR* gene [6]. Arginine 856, which is conserved among species, is located in a hot spot for variations, and the R856C change has been reported at a higher frequency than other amino acid variations [1,6,15-29]. This mutation has been detected in different populations around the world, which implies that the R856C mutation does not display a population preference (Table 2). However, this is the first study reporting this mutation in Chinese patients; thus, it is a supplement to the existing data. Similar to a previous case [17], the R856C mutation in family member II-2 was a *de novo* variation; and a *de novo* mutation in this hot spot could be a potential explanation for the relatively high prevalence of R856C.

Protein analysis predicted that the R856C variation was deleterious to AR protein function, and previous reports showed it was associated with CAIS. Biochemical anal-

ysis and molecular dynamics simulations suggested that the R856C variation could induce critical misplacement of C-terminal helix 12, resulting in a loss-of-function of ligand binding [26]. The clinical features of the individuals with the R856C variation in our study, i.e., CAIS, are concordant with those reported in previous studies. In most hospitals in China, a diagnosis of AIS is based on clinical features, including medical imaging and hormonal data. However, a precise diagnosis of AIS in an individual with 46,XY DSD requires the exclusion of other known or unknown etiologies. It has been shown that deficiencies in testosterone biosynthesis and metabolism, e.g., 5α-reductase type 2 deficiency, can be misdiagnosed as AIS owing to the similar and/or untypical clinical features presenting in some cases, e.g., in affected young infants and pre-pubertal children [8]. Likewise, due to the uncertain cutoff value of T/DHT, an elevated T/DHT ratio was reported in some AIS cases (including our case) [30], which could lead to a false diagnosis of 5α-reductase type 2 deficiency. Furthermore, a considerable number of patients who do not carry *AR* mutations show AIS phenotypes, indicating that these patients could be afflicted with a disease phenotypically similar to, but genetically different from, AIS [6]. Therefore, sequence analysis of the *AR* gene as well as other known or unknown genes is very important. Identification of a specific mutation by sequencing not only allows a precise diagnosis and/or prognosis for the patient, but also facilitates genetic counseling of related family members.

Based on our current knowledge of genetics, WES is the most valuable tool for clinical analysis of rare Mendelian diseases owing to its relatively low cost, detection throughput for genetic variants and sample size, high yields for clinical variants (more than 25%), and the ability to interpret the results and allow a clinical diagnosis [9]. In this study, all eight exons of the *AR* gene were examined, and the c.C2566T mutation was accurately detected by WES in a relatively short period (14 days), at an affordable cost (approximately 600 US Dollars for each sample). In addition, WES provides a robust approach for systematic screening of the potential genetic changes responsible for AIS-like phenotypes, which could advance etiologic research of 46,XY DSD. Finally, based on the discovery by WEB, laparoscopic surgery will be performed to remove the intra-abdominal gonads of patient III-3. We also provided genetic tests and counseling to related family members and confirmed that the mutation was *de novo* in the patients' mother (II-2), and will not pass to the offspring of the patients' sister (III-2).

The ARDB is a useful database for studying the phenotype-genotype correlation of a clinically identified *AR* mutation because the pathogenicity of the collected mutations in this database has been confirmed and the mutations have been classified into different categories [6]. In terms of AIS, the phenotypes of the corresponding mutations are classified as CAIS, PAIS, or MAIS, which is more clinically informative than the classifications in the ClinVar database and AR database at LOVD, even though the AR database at LOVD shares identical resources with the ARDB [6]. In addition, the publications of the collected mutations in ARDB are complete and easily accessible, which facilitates in depth studies of clinically identified mutations.

CONCLUSION

Using WES, we identified a c.C2566T (p.R856C) mutation of the *AR* gene in cases of familial CAIS and provided genetic counseling to related family members. This is the first study reporting this mutation in Chinese patients. We also compared the applicability of public databases for phenotype-genotype correlation studies of clinically identified *AR* mutations and suggested that the ARDB best satisfies clinical demands.

Acknowledgement:
This work was supported by the Natural Science Foundation for Distinguished Young Scholars of Fujian Province (project no. 2015D012), the Natural Science Foundation of Fujian Province (project no. 2014D003), and the Medical Innovation Program of Fujian Province (2014-CXB-46).

Declaration of Interest:
The authors declare no conflict of interest.

References:

1. Ahmed SF, Cheng A, Dovey L, et al. Phenotypic features, androgen receptor binding, and mutational analysis in 278 clinical cases reported as androgen insensitivity syndrome. J Clin Endocrinol Metab 2000;85:658-65 (PMID: 10690872).

2. Morris JM. The syndrome of testicular feminization in male pseudohermaphrodites. Am J Obstet Gynecol 1953;65:1192-211 (PMID: 13057950).

3. Quigley CA, De Bellis A, Marschke KB, el-Awady MK, Wilson EM, French FS. Androgen receptor defects: historical, clinical, and molecular perspectives. Endocr Rev 1995;16:271-321 (PMID: 7671849).

4. Gottlieb B, Pinsky L, Beitel LK, Trifiro M. Androgen insensitivity. Am J Med Genet 1999;89:210-7 (PMID: 10727996).

5. Brown CJ, Goss SJ, Lubahn DB, et al. Androgen receptor locus on the human X chromosome: regional localization to Xq11-12 and description of a DNA polymorphism. Am J Hum Genet 1989; 44:264-9 (PMID: 2563196).

6. Gottlieb B, Beitel LK, Nadarajah A, Paliouras M, Trifiro M. The androgen receptor gene mutations database: 2012 update. Hum Mutat 2012;33:887-94 (PMID: 22334387).

7. Wang Q, Lu J, Yong EL. Ligand- and coactivator-mediated transactivation function (AF2) of the androgen receptor ligand-binding domain is inhibited by the cognate hinge region. J Biol Chem 2001;276:7493-9 (PMID: 11102454).

8. Akcay T, Fernandez-Cancio M, Turan S, Guran T, Audi L, Bereket A. AR and SRD5A2 gene mutations in a series of 51 Turkish 46,XY DSD children with a clinical diagnosis of androgen insensitivity. Andrology 2014;2:572-8 (PMID: 24737579).

9. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med 2013;369:1502-11 (PMID: 24088041).

10. NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 2016; 44:D7-19 (PMID: 26615191).

11. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet 2013;Chapter 7:Unit7 20 (PMID: 23315928).

12. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 2003;31:3812-4 (PMID: 12824425).

13. Landrum MJ, Lee JM, Benson M, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res 2016;44:D862-8 (PMID: 26582918).

14. Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. Hum Mutat 2011;32:557-63 (PMID: 21520333).

15. Brinkmann AO, Jenster G, Ris-Stalpers C, et al. Androgen receptor mutations. J Steroid Biochem Mol Biol 1995;53:443-8 (PMID: 7626493).

16. Hiort O, Sinnecker GH, Holterhus PM, Nitsche EM, Kruse K. The clinical and molecular spectrum of androgen insensitivity syndromes. Am J Med Genet 1996;63:218-22 (PMID: 8723113).

17. Ledig S, Jakubiczka S, Neulen J, et al. Novel and recurrent mutations in patients with androgen insensitivity syndromes. Horm Res 2005;63:263-9 (PMID: 15925895).

18. Malmgren H, Gustavsson J, Tuvemo T, Dahl N. Rapid detection of a mutation hot-spot in the human androgen receptor. Clin Genet 1996;50:202-5 (PMID: 9001799).

19. Sultan C, Lumbroso S, Poujol N, Belon C, Boudon C, Lobaccaro JM. Mutations of androgen receptor gene in androgen insensitivity syndromes. J Steroid Biochem Mol Biol 1993;46:519-30 (PMID: 8240973).

20. Cheikhelard A, Morel Y, Thibaud E, et al. Long-term followup and comparison between genotype and phenotype in 29 cases of complete androgen insensitivity syndrome. J Urol 2008;180: 1496-501 (PMID: 18710728).

21. Komori S, Sakata K, Tanaka H, Shima H, Koyama K. DNA analysis of the androgen receptor gene in two cases with complete androgen insensitivity syndrome. J Obstet Gynaecol Res 1997;23: 277-81 (PMID: 9255042).

22. Vasu VR, Saranya B, Jayashankar M, Munirajan AK, Santhiya ST. A novel splice site and two known mutations of androgen receptor gene in sex-reversed XY phenotype. Genet Test Mol Biomarkers 2012;16:749-55 (PMID: 22799610).

23. Avila DM, Wilson CM, Nandi N, Griffin JE, McPhaul MJ. Immunoreactive AR and genetic alterations in subjects with androgen resistance and undetectable AR levels in genital skin fibroblast ligand-binding assays. J Clin Endocrinol Metab 2002;87: 182-8 (PMID: 11788645).

24. De Bellis A, Quigley CA, Cariello NF, et al. Single base mutations in the human androgen receptor gene causing complete androgen insensitivity: rapid detection by a modified denaturing gradient gel electrophoresis technique. Mol Endocrinol 1992;6: 1909-20 (PMID: 1480178).

25. McPhaul MJ, Marcelli M, Zoppi S, Wilson CM, Griffin JE, Wilson JD. Mutations in the ligand-binding domain of the androgen receptor gene cluster in two regions of the gene. J Clin Invest 1992;90:2097-101 (PMID: 1430233).

26. Elhaji YA, Wu JH, Gottlieb B, et al. An examination of how different mutations at arginine 855 of the androgen receptor result in different androgen insensitivity phenotypes. Mol Endocrinol 2004;18:1876-86 (PMID: 15118070).

27. Melo KF, Mendonca BB, Billerbeck AE, et al. Clinical, hormonal, behavioral, and genetic characteristics of androgen insensitivity syndrome in a Brazilian cohort: five novel mutations in the androgen receptor gene. J Clin Endocrinol Metab 2003;88:3241-50 (PMID: 12843171).

28. Murono K, Mendonca BB, Arnhold IJ, Rigon AC, Migeon CJ, Brown TR. Human androgen insensitivity due to point mutations encoding amino acid substitutions in the androgen receptor steroid-binding domain. Hum Mutat 1995;6:152-62 (PMID: 7581 399).

29. Melo KF, Mendonca BB, Billerbeck AE, Costa EM, Latronico AC, Arnhold IJ. [Androgen insensitivity syndrome: clinical, hormonal and molecular analysis of 33 cases]. Arq Bras Endocrinol Metabol 2005;49:87-97 (PMID: 16544039).

30. Audi L, Fernandez-Cancio M, Carrascosa A, et al. Novel (60%) and recurrent (40%) androgen receptor gene mutations in a series of 59 patients with a 46,XY disorder of sex development. J Clin Endocrinol Metab 2010;95:1876-88 (PMID: 20150575).