**BJMG** 26 (1), 2023 • 51-56
DOI: 10.2478/bjmg-2023-0012
sciendo

CASE REPORT

# ANDROGEN INSENSITIVITY SYNDROME DUE TO NON-CODING VARIATION IN THE ANDROGEN RECEPTOR GENE: REVIEW OF THE LITERATURE AND CASE REPORT OF A PATIENT WITH MOSAIC C.-547C>T VARIANT

Noveski P[1], Plaseski T[2], Dimitrovska M[2], Plaseska-Karanfilska D[1]

[*]**Corresponding Authors:** Dijana Plaseska-Karanfilska, MD, PhD, Research Centre for Genetic Engineering and Biotechnology 'Georgi D. Efremov', Macedonian Academy of Sciences and Arts, Krste Misirkov 2, 1000 Skopje, Republic of Macedonia, Tel. +389 2 3235 410, E-mail: dijana@manu.edu.mk

## ABSTRACT

Sexual development (SD) is a complex process with strict spatiotemporal regulation of gene expression. Despite advancements in molecular diagnostics, disorders of sexual development (DSD) have a diagnostic rate of ~50%. Androgen insensitivity syndrome (AIS) represents the most common form of 46,XY DSD, with a spectrum of defects in androgen action. Considering the importance of very strict regulation of the SD, it is reasonable to assume that the genetic cause for proportion of the DSD lies in the non-coding part of the genome that regulates proper gene functioning. Here we present a patient with partial AIS (PAIS) due to a mosaic *de novo* c.-547C>T pathogenic variant in the 5'UTR of androgen receptor (*AR*) gene. The same mutation was previously described as inherited, in two unrelated patients with complete AIS (CAIS). Thus, our case further confirms the previous findings that variable gene expressivity could be attributed to mosaicism. Mutations in 5'UTR could create new upstream open reading frames (uORFs) or could disrupt the existing one. A recent systematic genome-wide study identified *AR* as a member of a subset of genes where modifications of uORFs represents an important disease mechanism. Only a small number of studies are reporting non-coding mutations in the *AR* gene and our case emphasizes the importance of molecular testing of the entire *AR* locus in AIS patients. The introduction of new methods for comprehensive molecular testing in routine genetic diagnosis, accompanied with new tools for in sillico analysis could improve the genetic diagnosis of AIS, and DSD in general.

**Keywords:** Androgen insensitivity syndrome (AIS); 5'UTR variant; upstream open reading frames (uORFs); non-coding variation; mosaicism

## INTRODUCTION

Disorders of Sexual Development (DSD) represent a group of conditions that affect the development of the reproductive system, classified as 46,XY DSD, 46,XX DSD and sex chromosomal DSD. The etiology of DSDs can be complex with many different genetic and environmental factors contributing to their development. In congenital DSD, the severity and age of onset are highly variable and depend on the biological function of the affected gene, but the variability could also be associated with specific mutations in the affected gene, and even a variable phenotype could be observed in patients with the same mutation. Studies using genetic testing with sequencing and deletion/duplication analysis of the *AR* gene identified causality in approximately 50% of the 46,XY DSD cases [1]. Next generation sequencing (NGS) analysis, using a targeted gene panel, showed a 60% diagnostic rate in patients with 46,XY DSD disorders of androgen synthesis and action, but only a 19% detection rate in the patients with 46,XX DSD [2].

Androgen insensitivity syndrome (AIS) represents a frequent form of 46,XY DSD where pathogenic mutations in the androgen receptor (*AR*) gene are responsible for a spectrum of defects in androgen action: complete androgen insensitivity syndrome (CAIS), partial androgen insensitivity syndrome (PAIS) and mild androgen insensitivity syndrome (MAIS) [3]. So far, pathogenic mutations in affected patients were identified primarily in the coding regions and conserved splice sites of the *AR,* and rarely, single exon deletions/duplications or whole gene deletions

[1] Research Centre for Genetic Engineering and Biotechnology 'Georgi D. Efremov', Macedonian Academy of Sciences and Arts, 1000 Skopje, Republic of North Macedonia

[2] University Clinic of Endocrinology, Diabetes and Metabolic Disorders, Clinical Centre "Mother Teresa", 1000 Skopje, Republic of North Macedonia

were also reported [3]. Genotype-phenotype correlation exists for certain *AR* mutations in CAIS patients [4], but also, different AIS phenotypes for identical mutations were observed [5]. This variable expressivity is associated with several mutations and is often attributed to the oligogenic factors [6, 7], sometimes, however, also to mosaicism [8]. Recently, the recurrent germline pathogenic variant c.-547C>T in the 5' untranslated region (5'-UTR) of the *AR* gene was described in two unrelated patients with CAIS [9]. Here, we report on a mosaic form of the same pathogenic variant in a patient with PAIS. The current understanding of the involvement of the 5'UTR variation in highly penetrant diseases is discussed and an overview of the known *AR* non-coding pathogenic variants is given.

## METHODS

**Patient presentation**

A twenty-four year old patient with a disorder of sexual development and 46,XY karyotype was referred to our laboratory for genetic testing. According to the previous medical records, the patient was born with ambiguous external genitalia, described as a hypoplastic penis (resembling hypertrophic clitoris), partially covered with bifid scrotum (resembling oedematous labia), and in whom gonadal structures were identified after ultrasound examination. Urethrocystography showed the presence of a male urethra. Blind ending structure resembling a vagina was also observed. No uterus was detected. A human chorionic gonadotropin (hCG) stimulation test showed a positive response, resulting in increased testosterone production. At the age of seven months, cytological gonadal punction was performed, and, according to the chromatin status, cells resembling the Sertoli cells were observed. Although the parents were advised to raise the child as a female and a feminizing genitoplasty was performed at 6 years of age, they have reared him as a male and he decided to undergo surgical gender reassignment at 24 years of age.

**Genetic analysis**

We have analysed in total 4 DNA samples: 3 DNA samples from the patient and one DNA sample from his mother. Initially, DNA from the patient's blood (peripheral white blood cells) was isolated using standard phenol/chloroform protocol. In the second patient's admission, DNA was isolated from the blood sample of patient's mother as well as from a new blood sample and a buccal swab from the patient. This time the extraction of the DNA was performed using MagCore Super automated nucleic acids extractor (RBC Bioscience Corp., Taiwan).

Since the patient's phenotype in combination with the results of the hCG test and 46XY karyotype was indicative for the presence of AIS, Sanger sequencing targeting the exons and exon/intron boundaries of the androgen receptor (*AR*) gene was performed. Exon 1 sequencing was performed with the inclusion of the 5'UTR region, which is a standard practice in our laboratory. We then performed a multiple ligation-dependent probe analysis (MLPA) for detection of the exon copy number changes in the *AR* gene, using the AR P074-A3 Androgen insensitivity syndrome kit (MRC-Holland, Amsterdam, The Netherlands). To confirm the biological relatedness between the patient and patient's mother, a comparison was performed of the allelic profiles of 15 polymorphic short tandem repeat (STR) loci using previously the published multiplex PCR of fluorescently labelled primers [10]. We have used also another multiplex PCR of fluorescently labelled primers [11] which previously have been used to detect mosaic loss of sex chromosome in blood cells [12]. Both multiplex PCR reactions also contain primers amplifying the *SRY* gene. The Sanger sequencing, MLPA and fluorescent multiplex PCR reactions were analyzed on the ABI PRISM 3500 Genetic Analyzer (Thermo Fisher Scientific, Waltham, MA, USA). For the MLPA data analysis, Coffalyser.Net software (https://www.mrcholland.com/technology/software/coffalyser-net) was used. Electropherograms from the Sanger sequencing were analyzed with Sequencing Analysis v5.4 (Thermo Fisher Scientific, Waltham, MA, USA). ImageJ software (https://imagej.nih.gov/ij/index.html) was used to compare area ratios of the normal and mutant allele's fluorescence peaks from the electropherograms of the Sanger sequencing. UTRannotator [13], a plugin to the Ensembl VEP analysis software [14], was used for *in silico* prediction of the possible impact of the 5'UTR variant.

## RESULTS

Sanger sequencing did not reveal any pathogenic variant in the protein coding sequence and splice regions of the *AR* gene. MLPA analysis did not show the presence of deletions/duplications of the *AR* exons. However, we have discovered "heterozygous" nucleotide change in the 5'UTR region of the *AR* gene, c.-547C>T, with an unequal ratio between normal and mutated allele (T allele being over represented compared to the C allele). This was an unusual finding since every nucleotide change in patient's X chromosome should have been presented as hemizygous, considering the patient's 46, XY karyotype. One reason for this unusual finding could be the presence of a small proportion of cell line harboring 47,XXY karyotype, with one of the X chromosomes carrying the normal C allele. However, karyotype and multiplex QF-PCR analysis did not reveal presence of extra X chromosome. Other pos-

sibilities for the unbalanced ratios of the detected alleles could be the presence of DNA contamination or a mosaic mutation. To resolve this, a second blood sample from the patient was obtained, together with the blood sample from the patient's mother. At the same time, a sample from the patient's buccal swab was also obtained. Sanger sequencing confirmed the previous finding of both mutant and normal alleles in the DNA samples isolated from the second blood sample and the buccal swab (Figure 1, A and B). The mutation was not present in the mother (Figure 1 C), which resolved the origin of the mutation as a *de novo* event. Biological relatedness between patient and his mother was confirmed with the STR markers. Comparison of the areas of the normal and mutant alleles indicated that the normal allele was present in approximately 20% of the cells. This was obtained as mean value from several different electropherograms (two from Figure 1, A and B and 5 others not shown). *In silico* analysis with the UTRannotator showed that the c.-547C>T creates translational initiation codon ATG, and that newly created uORF terminates after 186 nucleotides. Also, UTRannotator predicted that the sequence, flanking a newly created initiation codon, shows moderate match to the Kozak consensus sequence.

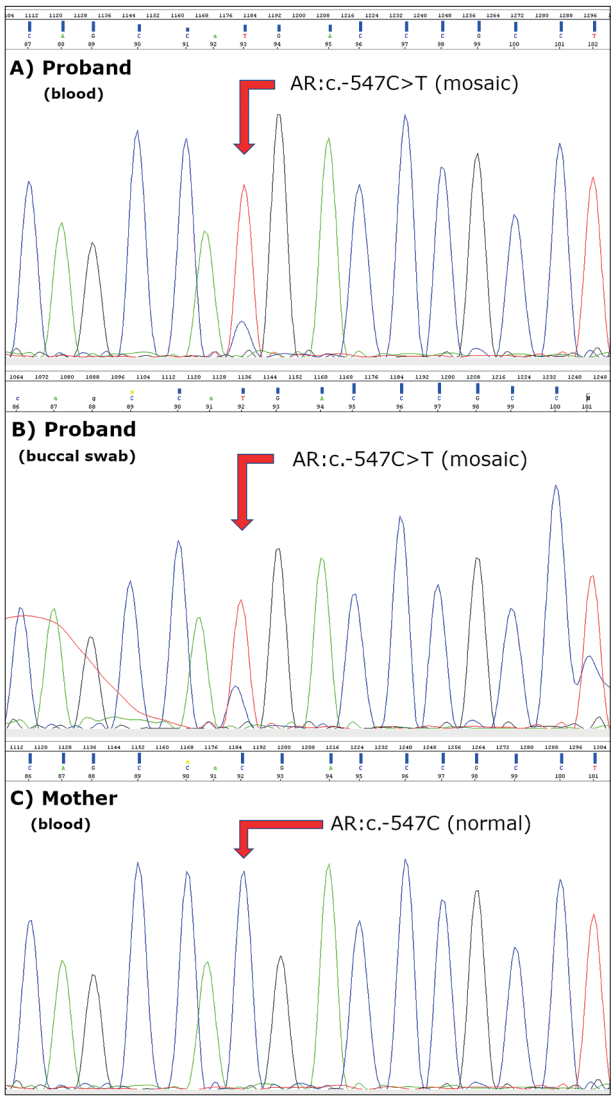


**Figure 1.** Electropherograms from Sanger sequencing for DNA isolated from: A) Patient's blood; B) Patient's buccal swab and C) Patient's mother blood.

## DISCUSSION

Sexual development is a complex process starting at the beginning of the embryonic development and lasting through puberty, involving a large number of genes with specific temporal and spatial regulation of their expression. For the genetic diagnosis of rare disorders, a single-gene testing is still useful in the cases with a clear etiology, although the use of a panel of associated genes, analyzed with short reads massively parallel sequencing, is already an established practice.

Despite this advancement in the genetic diagnosis, there is still a large proportion of undiagnosed DSD patients, with diagnostic rate being ~50% [15]. There is emerging evidence that 5′UTR variants could be a cause of penetrant human diseases. Approximately half of the human transcripts naturally contain upstream open reading frames (uORF) in their 5′UTR which modulate the production of the main protein by disturbing the protein translation [16, 17]. Variants that introduce or disrupt uORF are rare and are subjected to strong negative selection because of their ability to cause a disease [17-19]. A recent study based on a large cohort of human whole genome sequences identified a subset of genes, among which was *AR* gene, where high-impact uORF-perturbing variants would have a deleterious effect on gene function [17].

Previous functional analysis of the c.-547C>T pathogenic variant confirmed the mutation's deleterious effect on *AR* function and association with CAIS phenotype [9]. This was in line with the predicted functional consequence given by the UTRannotator tool. A variant was also not identified in the Genome Aggregation Database (gnomAD), a large-scale population database for variant frequencies [20]. The presence of PAIS in our patient, compared to CAIS in patients described by Hornig et al., is in line with the previous findings that variable expressivity in AIS could be attributed to mosaicism [8, 21]. Specifically, there is a phenotypic overlap consisting of the absence of a uterus, blind ending vagina and positive hCG test as well as a phenotypic difference consisting of the presence of ambiguous genitalia in our patient. This stands in contrast to the presence of female genitalia in the patients described

by Hornig et al. In individuals with PAIS and ambiguous genitalia, the tendency is to assign sex of rearing after an expert evaluation has been completed [22]. This was the procedure that was followed in our case, and we are unaware for the reasons behind parents' decision. Although preferable for studies of androgen action, we have not tested patient's gonadal tissue or genital skin fibroblasts for mosaicism, but the presence of the same rate of the mutation in two different cell types is suggestive that the mutation is present in the entire body, in approximately the same rate.

In general, regions outside the *AR* coding sequences and conserved splice sites have been less extensively studied in AIS patients. The non-coding *AR* pathogenic variants reported on so far are presented in Table 1. Recently, there was a report of a large family pedigree of PAIS patients with disease causing LINE-1 retrotransposon insertion in the 5′UTR region of the *AR* gene [23]. Another study identified deep intronic mutation in intron 6, creating an alternative splice acceptor site of the *AR* gene, in a family with PAIS patients [24]. The same mutation was found in an unrelated patient, also with PAIS [25]. Deep intronic mutations were observed in CAIS patients [24, 26, 27].

Studying the noncoding region of the *AR* gene is highly challenging due to its genomic size of 186,500 nucleotides, as compared of only 2763 nucleotides of the coding sequence. However, the probability of identifying pathogenic variants outside of the AR coding regions in AIS patients is high [25]. Further advancements in genetic diagnostics with the introduction of targeted sequencing of the whole genomic locus of *AR* gene, using targeted long-read sequencing which also preserves epigenetic information [28, 29], could better clarify the role of noncoding variation in AIS.

In conclusion, comprehensive genetic testing by targeting the entire *AR* gene locus, together with the development of appropriate annotation tools, would contribute to the identification of the missing heritability of AIS.

## DECLARATION OF INTEREST

The authors report no conflicts of interest.

## REFERENCES

1. Kyriakou A, Lucas-Herald AK, McGowan R, Tobias ES, and Ahmed SF. Disorders of sex development: advances in genetic diagnosis and challenges in management. Advances in Genomics and Genetics. 2015; 5: 165-177.
2. Eggers S, Sadedin S, van den Bergen JA, Robevska G, Ohnesorg T, Hewitt J, *et al.* Disorders of sex development: insights from targeted gene sequencing of a large international patient cohort. Genome Biol. 2016; 17(1): 243.
3. Batista RL, Costa EMF, Rodrigues AS, Gomes NL, Faria JA, Jr., Nishi MY, *et al.* Androgen insensitivity syndrome: a review. Arch Endocrinol Metab. 2018; 62(2): 227-235.
4. Boehmer AL, Brinkmann O, Brüggenwirth H, van Assendelft C, Otten BJ, Verleun-Mooijman MC, *et al.* Genotype versus phenotype in families with androgen insensitivity syndrome. J Clin Endocrinol Metab. 2001; 86(9): 4151-60.
5. Gottlieb B, Beitel LK, Nadarajah A, Paliouras M, and Trifiro M. The androgen receptor gene mutations database: 2012 update. Hum Mutat. 2012; 33(5): 887-94.

**Table 1.** Non-coding disease causing mutations in AR gene reported in AIS patients.

| AR Mutation | Mutation type and location within non-coding AR sequence | Phenotype (number of affected individuals and family history) | Inheritance | Reported by [ref] |
|---|---|---|---|---|
| c.-547C>T | SNV in 5'UTR | CAIS (n=2, unrelated cases) | Inherited | Hornig et al. [9] |
| c.-547C>T | SNV in 5'UTR | PAIS (n=1, no family history) | *De novo*, mosaic | This study |
| LINE-1 insertion at c.-268 | Insertion of >800 nucleotides in 5'UTR | PAIS (n=9; one family, 4 generations) | Inherited | Batista et al. [23] |
| c.2450–42G>A | Deep intronic SNV in intron 6 | PAIS (n=3; one family, 3 generations) | Inherited | Ono et al. [24] |
| c.2450–42G>A | Deep intronic SNV in intron 6 | PAIS (n=1, uncle with similar symptoms but without genetic test) | Not reported, presumable inherited | Kalinchenko et al. [25] |
| c.1769–11T>A | Deep intronic SNV in intron 2 | CAIS (n=3; one family, 3 generations, reported as PAIS but phenotype is more suggestive of CAIS, as described in [23]) | Inherited | Brüggenwirth et al. [26] |
| c.2450-118A>G | Deep intronic SNV in intron 6 | CAIS (n=2; one family, 2 generations) | Inherited | Känsäkoski et al. [27] |

SNV - single nucleotide variant

6. Camats N, Flück CE, and Audí L. Oligogenic Origin of Differences of Sex Development in Humans. Int J Mol Sci. 2020; 21(5).
7. Boehmer AL, Brinkmann AO, Nijman RM, Verleun-Mooijman MC, de Ruiter P, Niermeijer MF, *et al.* Phenotypic variation in a family with partial androgen insensitivity syndrome explained by differences in 5alpha dihydrotestosterone availability. J Clin Endocrinol Metab. 2001; 86(3): 1240-6.
8. Köhler B, Lumbroso S, Leger J, Audran F, Grau ES, Kurtz F, *et al.* Androgen insensitivity syndrome: somatic mosaicism of the androgen receptor in seven families and consequences for sex assignment and genetic counseling. J Clin Endocrinol Metab. 2005; 90(1): 106-11.
9. Hornig NC, de Beaufort C, Denzer F, Cools M, Wabitsch M, Ukat M, *et al.* A Recurrent Germline Mutation in the 5'UTR of the Androgen Receptor Causes Complete Androgen Insensitivity by Activating Aberrant uORF Translation. Plos one. 2016; 11(4): e0154158.
10. Noveski P, Terzic M, Vujovic M, Kuzmanovska M, Sukarova Stefanovska E, and Plaseska-Karanfilska D. Multilevel regression modeling for aneuploidy classification and physical separation of maternal cell contamination facilitates the QF-PCR based analysis of common fetal aneuploidies. Plos one. 2019; 14(8): e0221227.
11. Plaseski T, Noveski P, Trivodalieva S, Efremov GD, and Plaseska-Karanfilska D. Quantitative fluorescent-PCR detection of sex chromosome aneuploidies and AZF deletions/duplications. Genetic testing. 2008; 12(4): 595-605.
12. Noveski P, Madjunkova S, Sukarova Stefanovska E, Matevska Geshkovska N, Kuzmanovska M, Dimovski A, *et al.* Loss of Y chromosome in peripheral blood of colorectal and prostate cancer patients. Plos one. 2016; 11(1): e0146264.
13. Zhang X, Wakeling M, Ware J, and Whiffin N. Annotating high-impact 5'untranslated region variants with the UTRannotator. Bioinformatics. 2021; 37(8): 1171-1173.
14. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, *et al.* The Ensembl Variant Effect Predictor. Genome Biol. 2016; 17(1): 122.
15. Baetens D, Mendonça BB, Verdin H, Cools M, and De Baere E. Non-coding variation in disorders of sex development. Clin Genet. 2017; 91(2): 163-172.
16. Soukarieh O, Meguerditchian C, Proust C, Aïssi D, Eyries M, Goyenvalle A, *et al.* Common and Rare 5'UTR Variants Altering Upstream Open Reading Frames in Cardiovascular Genomics. Front Cardiovasc Med. 2022; 9: 841032.
17. Whiffin N, Karczewski KJ, Zhang X, Chothani S, Smith MJ, Evans DG, *et al.* Characterising the loss-of-function impact of 5' untranslated region variants in 15,708 individuals. Nat Commun. 2020; 11(1): 2523.
18. Calvo SE, Pagliarini DJ, and Mootha VK. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proc Natl Acad Sci U S A. 2009; 106(18): 7507-12.
19. Lee DSM, Park J, Kromer A, Baras A, Rader DJ, Ritchie MD, *et al.* Disrupting upstream translation in mRNAs is associated with human disease. Nat Commun. 2021; 12(1): 1515.
20. Chen S, Francioli LC, Goodrich JK, Collins RL, Kanai M, Wang Q, *et al.* A genome-wide mutational constraint map quantified from variation in 76,156 human genomes. bioRxiv. 2022: 2022.03.20.485034.
21. Holterhus PM, Brüggenwirth HT, Hiort O, Kleinkauf-Houcken A, Kruse K, Sinnecker GH, *et al.* Mosaicism due to a somatic mutation of the androgen receptor gene determines phenotype in androgen insensitivity syndrome. J Clin Endocrinol Metab. 1997; 82(11): 3584-9.
22. Gottlieb B and Trifiro MA. Androgen Insensitivity Syndrome [Updated 2017 May 11]. In: Adam MP, Mirzaa GM, Pagon RA, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2023. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1429/. 1993.
23. Batista RL, Yamaguchi K, Rodrigues ADS, Nishi MY, Goodier JL, Carvalho LR, *et al.* Mobile DNA in Endocrinology: LINE-1 Retrotransposon Causing Partial Androgen Insensitivity Syndrome. J Clin Endocrinol Metab. 2019; 104(12): 6385-6390.
24. Ono H, Saitsu H, Horikawa R, Nakashima S, Ohkubo Y, Yanagi K, *et al.* Partial androgen insensitivity syndrome caused by a deep intronic mutation creating an alternative splice acceptor site of the AR gene. Sci Rep. 2018; 8(1): 2287.
25. Kalinchenko NY, Petrov VM, Panova AV, and Tiulpakov AN. [A deep intronic mutation in AR gene causing androgen insensitivity syndrome: difficulties of diagnostics]. Probl Endokrinol (Mosk). 2021; 67(5): 48-52.
26. Brüggenwirth HT, Boehmer AL, Ramnarain S, Verleun-Mooijman MC, Satijn DP, Trapman J, *et al.* Molecular analysis of the androgen-receptor gene in a family with receptor-positive partial androgen insensitivity: an unusual type of intronic mutation. Am J Hum Genet. 1997; 61(5): 1067-77.

27. Känsäkoski J, Jääskeläinen J, Jääskeläinen T, Tommiska J, Saarinen L, Lehtonen R*, et al.* Complete androgen insensitivity syndrome caused by a deep intronic pseudoexon-activating mutation in the androgen receptor gene. Sci Rep. 2016; 6: 32819.
28. Miller DE, Lee L, Galey M, Kandhaya-Pillai R, Tischkowitz M, Amalnath D*, et al.* Targeted long-read sequencing identifies missing pathogenic variants in unsolved Werner syndrome cases. J Med Genet. 2022; 59(11): 1087-94.
29. Miller DE, Sulovari A, Wang T, Loucks H, Hoekzema K, Munson KM*, et al.* Targeted long-read sequencing identifies missing disease-causing variation. Am J Hum Genet. 2021; 108(8): 1436-1449.