Indian Journal of Pediatrics (October 2025) 92(10):1056–1061
https://doi.org/10.1007/s12098-024-05249-0

ORIGINAL ARTICLE

Advocating Targeted Sequential Screening over Whole Exome Sequencing in 21-Hydroxylase Deficiency

Lavanya Ravichandran[1] · Shriti Paul[1] · A. Rekha[2] · Deny Varghese[1] · R Parthiban[1] · H.S. Asha[1] · Sarah Mathai[3] · Anna Simon[3] · Sumita Danda[2] · Nihal Thomas[1] · Aaron Chapla[1]

Received: 7 June 2024 / Accepted: 14 August 2024 / Published online: 20 September 2024

Abstract

Objectives Whole exome sequencing (WES) has emerged as the preferred method for diagnosing a range of Mendelian disorders. Nonetheless, the applicability of WES in genetic diagnosis of 21-hydroxylase deficiency (21-OHD) remains uncertain due to the intricacies involved in molecular analysis of the *CYP21A2* gene.

Methods In this case series, authors report the outcomes of couples or families who underwent WES followed by focused sequential strategy (FSS) targeting *CYP21A2* gene hotspot mutations and targeted sequencing of genes associated with Congenital Adrenal Hyperplasia (CAH).

Results This analysis revealed that WES, when compared to FSS, resulted in six false-negative findings out of seven subjects and one false-positive result. These results were corroborated through validation using Multiplex Ligation-Dependent Probe Amplification (MLPA) and Sanger sequencing.

Conclusions One major limitation of exome sequencing lies in target enrichment, which often achieves less than 95% coverage of the regions of interest, potentially leading to false negatives. This challenge is particularly pronounced when deciphering the complex genetics of 21-OHD, characterized by intricate pseudogene-derived rearrangements and gene conversions. Additionally, next-generation sequencing (NGS) analysis of the *CYP21A2* gene is not straightforward due to reads aligning to pseudogene regions, necessitating stringent computational pipelines with defined targets. However, simple genotyping assays have shown a high positive yield of pseudogene-derived mutations in over 80% of cases, while targeted NGS can be valuable in subjects with initially negative results. Therefore, WES is not recommended as the primary testing method for 21-OHD and may be better suited for rare forms of CAH once *CYP21A2* mutations have been ruled out.

Keywords 21-hydroxylase deficiency · Congenital Adrenal Hyperplasia · *CYP21A2* · Whole exome sequencing

Introduction

Mutations in the *CYP21A2* gene are the leading cause of congenital adrenal hyperplasia (CAH). The prevalence of these mutations is as high as 90–95% of all patients with CAH and result in partial to complete loss of 21-hydroxylase activity [1]. The majority of pathogenic alleles implicated in 21-hydroxylase deficiency (21-OHD) arise from asymmetric meiotic recombination events between the highly similar functional gene (*CYP21A2*) and its pseudogene counterpart (*CYP21A1P*) [2]. These deleterious mutations encompass large deletions and gene conversions, occurring in approximately 25–30% of cases, as well as point mutations resulting from micro conversions, which are observed in 75% of cases [3].

With several complex mutations transferred from the pseudogene, genetic testing in 21-OHD requires specific strategies that warrant locus specific amplification of *CYP21A2* gene, precise identification of chimeras/gene conversions and genotyping of common point mutations. It also requires expertise in interpreting the results with extensive knowledge on RCCX module and the chimeras [4].

✉ Aaron Chapla
aaronchapla@gmail.com

[1] Department of Endocrinology, Diabetes and Metabolism, Christian Medical College, Vellore, India

[2] Department of Medical Genetics, Christian Medical College, Vellore, India

[3] Department of Pediatric Endocrinology, Christian Medical College, Vellore, India

Multiple techniques with a focused sequential approach are often required to identify the actual extent of these mutations [5]. Initial gene enrichment is carried out with locus specific long range PCR [6, 7]. Genotyping assays like allele-specific PCR and restriction fragment length polymorphism (RFLP) have been used to identify hotspot point mutations [5, 8, 9]. The results of long range PCR and the hotspot mutations can pin point the chimeras whereas multiplex ligation-dependent probe amplification (MLPA) is utilized to identify copy number variations (CNVs) in this gene [10]. In the recent years, a few studies on the utility of next-generation sequencing (NGS) in CAH have been reported. However, the NGS data analysis for the *CYP21A2* gene is challenging and requires stringent computational pipelines to analyze these mutations [5, 11].

As WES becomes more accessible, clinicians and geneticists are increasingly favoring it for routine genetic diagnosis across various disorders. In this report, the authors delineate the challenges and pitfalls encountered when employing WES for 21-OHD, drawing from the case findings.

Material and Methods

This case series includes seven subjects from four families who underwent WES outsourced to an external laboratory. They were also tested for mutations in CAH related genes using an in-house developed focused sequential strategy (FSS) at Molecular Endocrinology Laboratory, CMC, Vellore [5, 10]. FSS strategy includes long range and allele specific PCRs to identify large 30 kb deletion and hotspot mutations in *CYP21A2* gene followed by targeted NGS for seven genes in CAH -- *CYP21A2*, *CYP11B1*, *CYP17A1*, *CYP19A1*, *POR*, *StAR* and *HSD3B2* genes. MLPA was carried out to identify other large deletions and CNVs.

Results

MEL001: A 37-y-old married female, with no consanguineous relation to her 43-y-old husband, presented with a history of four recurrent pregnancy losses and one neonatal death with atypical genitalia. The husband was documented to have oligospermia from a semen analysis conducted in 2016, along with a family history of abortions.

Both partners underwent karyotyping, which returned normal results. To further investigate potential genetic factors, WES was performed to assess carrier status for inborn errors of metabolism (IEM) and mutations associated with CAH. The WES results revealed that the wife did not carry any clinically significant variants related to the observed phenotype. However, the husband was identified as a carrier for the *CYP21A2*:Ala392Thr variant associated with 21-OHD.

Given the recurrent miscarriages experienced by the wife, the husband's positive carrier status for 21-OH deficiency, and the suspicion of CAH in the previous child with atypical genitalia, both partners underwent testing with FSS. MLPA analysis confirmed the absence of large deletions in both partners (Figs. 1 and 2). Subsequent targeted NGS of the wife's sample revealed heterozygosity for three variants in the *CYP21A2* gene (NM_000500.9): c.-113G>A, c.-126C>T, and c.-110T>C (Fig. 2). The sequencing depth for these variants was 99, 113, and 99 reads, respectively. Concurrently, MLPA results demonstrated a heterozygous loss of probe in the promoter region, consistent with the promoter variant identified through NGS (Fig. 1).

The *CYP21A2*:c.-113G>A variant is situated upstream of the gene in the untranscribed region. Its allele frequency in the gnomAD Genomes database is 0.00239. This variant commonly co-segregates with -126C>T, -110T>C, and -103A>G variants in the promoter region. In this case, the promoter variant co-segregated with -126C>T and -110T>C variants. These variants are pseudogene derived and have been reported in non-classical CAH [12]. Functional analysis have demonstrated that c.-113G>A can impair transcriptional activity by 80% and diminish the nuclear binding capacity [13].

Conversely, targeted NGS analysis revealed that the husband did not carry the *CYP21A2* (NM_000500.9): c.1147G>A(p.Ala392Thr) variant previously identified through exome sequencing. The sequencing depth for this locus in targeted NGS was 119 reads (Fig. 2), and this finding was further confirmed by Sanger sequencing (Fig. 2). Additionally, he tested negative for other pathogenic or likely pathogenic variants in the panel of CAH-associated genes assessed.

MEL 002: The proband, a 37-y-old female diagnosed with CAH, presented with symptoms including vomiting, loose stools, hyperpigmentation and clitoromegaly at birth. She was diagnosed with salt-losing CAH and underwent clitoroplasty in 1989. Further, her younger sister, a 30-y-old female, exhibited a similar phenotype at birth and underwent clitoroplasty in 1999.

Both affected siblings had a karyotype of 46XX. WES performed in 2014 yielded negative results for both siblings. Although the original coverage details were unavailable, the documentation indicated very low coverage in certain target regions despite repeated sequencing. In 2021, utilizing an in-house FSS, both siblings were identified as compound heterozygous for a 30 kb deletion with chimera 1 on one allele and the Ile172Asn hotspot mutation on the other allele.

MEL 003: A non-consanguineous couple experienced the loss of two children during the early neonatal period, likely due to CAH. The first male child exhibited clinical signs

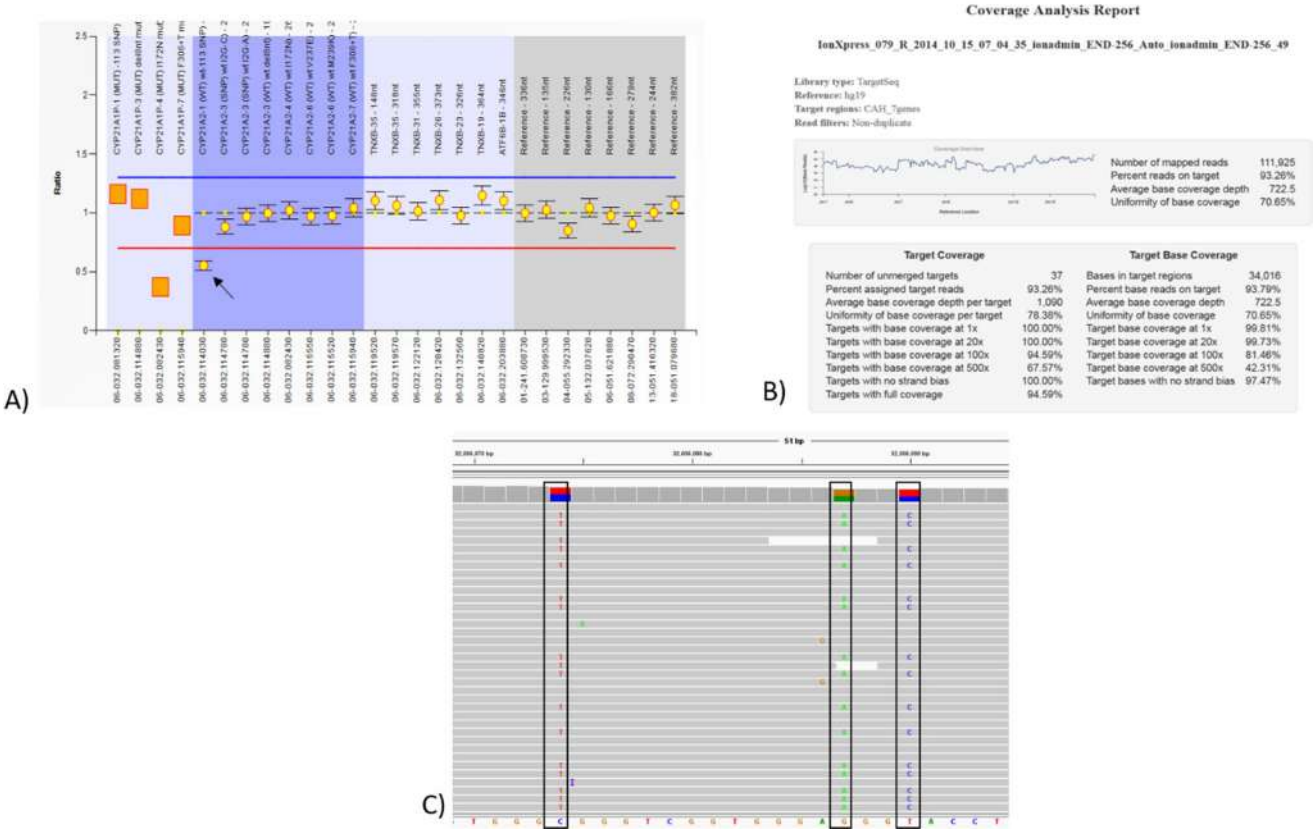

Fig. 1 **A**) MLPA results in wife showing no large deletion/duplication. It also shows heterozygous loss of probe '*CYP21A2* exon 1 (-113 SNP)' indicated by a *black arrow* that corresponds to the promoter mutation c.-113G>A. **B**) MEL-001: Coverage analysis report for wife's sample- generated from Torrent suit software for seven genes in CAH panel in 37 amplicons with a mean coverage of 722.5X and 99.73% of the target having a minimum coverage of 20X reads. **C**) Integrative Genomics Viewer (IGV.2.16.2) snapshot of *CYP21A2* gene reads aligned to hg19 reference genome showing heterozygous positive promoter mutation *CYP21A2*(NM_000500.9):c.-113G>A co-segregating with c.-126C>T, and c.-110T>C

of hypospadias, underdeveloped scrotum, recurrent vomiting, electrolyte imbalance, and passed away at four months of age. Their second pregnancy ended in spontaneous abortion. The third male baby presented with clinical indications of failure to thrive, developmental delay, seizures, vomiting, tachypnea, pneumonia, electrolyte imbalance, hyponatremia, and hypoglycemia, and passed away at six months of age.

WES results revealed no CAH mutations in the couple. However, utilizing the FSS, the wife was identified as a carrier for the *CYP21A2*:Gln319Ter (Q319X) hotspot mutation, while the husband was found to be a carrier for the *CYP21A2*:Gly111ValfsTer21 (8 bp deletion hotspot mutation).

MEL 004: A five-month-old boy, born to parents in a third-degree consanguineous marriage, was diagnosed with primary adrenal insufficiency and hyperpigmentation. The family had a history of spontaneous abortion in two previous conceptions, and an elder sibling who is phenotypically normal.

Sequencing coverage details for MEL004 were unavailable. WES results returned negative for this subject; however, utilizing a targeted approach, he tested positive for the P30L hotspot mutation. Table 1 provides the coverage details and the genetic test results from WES vs. targeted approach in the four families.

Discussion

With a notably high carrier frequency of the *CYP21A2* hotspot mutations [14] and a high prevalence of CAH in India [15], there is an urgent need for cost-effective strategies to address the disease burden effectively. Despite significant advancements in sequencing technology, many clinicians and scientists in the endocrine specialty remain unfamiliar with the utility or optimal application of hotspot screening, targeted NGS, and WES strategies in CAH testing.

In this case series, the authors present the genetic test results of individuals clinically suspected of having CAH, as well as at-risk couples seeking genetic evaluation for carrier status of CAH-related genes. Initially, whole exome sequencing failed to detect mutations in six out of seven subjects tested and erroneously identified one subject as a carrier for a *CYP21A2* mutation.

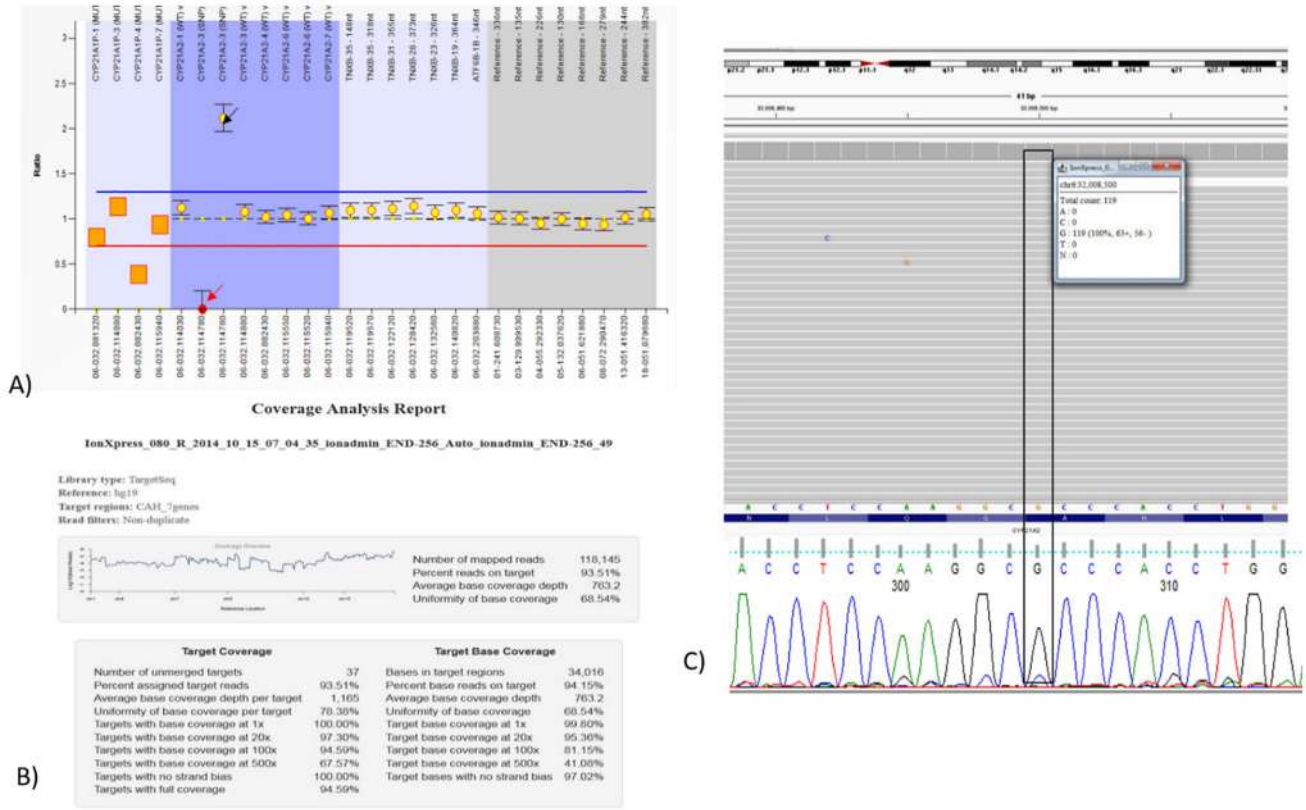

Fig. 2 (**A**) MLPA results in the husband showing normal probe ratio of *CYP21A2* gene. The intron 2 splice mutation I2G has two wild-type alleles (A/C) and one mutant allele (G) on this loci. MLPA utilizes two wildtype probes to identify the presence or absence of I2G mutation. The subject was homozygous for wildtype A allele reflected with a probe ratio 2 (*black arrow*) and a probe ratio 0 for wildtype C allele (*red arrow*). (**B**) MEL-001: Coverage analysis report for husband's sample- generated from Torrent suit software for seven genes in CAH panel in 37 amplicons with a mean coverage of 763.2X and 95.36% of the target having a minimum coverage of 20X reads indicating gaps in the sequencing regions. These gaps were patched up using amplicon-based sequencing in the subsequent runs. (**C**) Snapshot from Integrative Genomics Viewer (IGV.2.16.2) showing *CYP21A2* gene reads aligned to hg19 reference genome showing that the husband is negative for *CYP21A2*(NM_000500.9):c.1147G>A(p.Ala392Thr). Below is the sanger chromatogram corresponding to negative results for *CYP21A2*:c.1147G>A

The NGS platforms currently used in clinical practice for targeted next-generation sequencing (tNGS) and WES are all short read sequencers (SRS). With these platforms, the target genes are fragmented, sequenced, and aligned to the human reference genome to identify disease-causing variants [16]. Therefore, achieving optimal coverage of the sequencing targets with sufficient depth is a primary requirement in NGS analysis for accurate variant calling. A minimum of 20X coverage in the sequenced reads of the complete target that provides 99% sensitivity is considered for variant analysis [17, 18]. The major limitation in exome sequencing is with achieving optimal coverage of the target region. On an average, a WES targets around 30 Mb of the human genome ($3 * 10^7$ base pairs) with > 5,00,000 exons sequenced [19, 20]. Larger the target size, higher is the chance of having sequencing gaps. Several studies have reported significantly lower coverage in WES while compared to target panels [21], that may result in missing of clinically significant variants and false negative results [17].

In this study, whole exome sequencing yielded a mean depth of 102.41X and 92.98% of the target having 20x (average of four samples in whom coverage details were available). A 20x coverage of 93% indicates that there could be sequencing gaps up to 7% and this could have been a contributing factor to the missed variants. Further, insufficient depth at the loci where variants are identified might also result in false positive results. In addition to challenges in achieving optimal coverage and sufficient depth in short read sequencers, NGS analysis of *CYP21A2* gene by itself is complicated as the highly homologous reads align both to the functional gene and pseudogene coordinates in the reference genome. Indigenous bioinformatic pipelines, allowing alignment of reads only to the functional gene [11] and defining the target with specific bedfiles [5], is required to achieve optimal read alignment. This could be another reason explaining false positive and negative results with WES in these subjects.

Table 1 CAH genetic screening results from WES vs. targeted approach in four families

Family	Family member	WES results		Results from targeted approach with sanger validation	
		Coverage details	Variants identified	Genotyping for common hotspot mutations with ASPCR	Targeted NGS
MEL 001 Carrier testing	Wife	Mean depth of 100.85X and 94.92% at 20X.	Negative	Negative	Mean depth of 722.5X and 99.73% at 20X ***CYP21A2*: c.-113G>A**
	Husband	Mean depth of 115.45X and 95.34% at 20X	*CYP21A2*:Ala392Thr	Negative	Mean depth of 763.2X and 95.36% at 20X **Negative (sequencing gaps covered)**
MEL 002 Genetic diagnosis in CAH affected siblings	Sibling 1	NA	Negative	30 kb deletion with chimera 1 / Ile172Asn	NR
	Sibling 2	NA	Negative	30 kb deletion with chimera 1 / Ile172Asn	NR
MEL 003 Carrier testing	Husband	Mean depth of 107.3X with 90.3% at 20X	Negative	*CYP21A2*:Gly111ValfsTer21	NR
	Wife	Mean depth of 86.02X with 91.35% at 20X	Negative	*CYP21A2*:Gln319Ter	NR
MEL 004 Genetic diagnosis for CAH	Proband	NA	Negative	*CYP21A2*:Pro30Leu	NR

ASPCR Allele Specific PCR; *CAH* Congenital adrenal hyperplasia; *NA* Not available; *NGS* Next-generation sequencing; *NR* Not required; *WES* Whole exome sequencing

In disorders like 21-OH deficiency, 90% of the cases are positive for the pseudogene derived hotspot mutations. These mutations can be detected using simple genotyping assays like long range and allele specific PCR. Five out of seven subjects in this case series were positive with Allele Specific PCR (ASPCR). Only two samples (MEL-001) required a targeted NGS. Targeted sequencing offers much deeper coverage (200X-800X) as compared to WES. which is highly reliable for variant analysis in clinical samples [22]. Moreover, PCR based primary enrichment of gene regions utilized in tNGS enables rapid filling of gaps with sanger or amplicon based NGS sequencing. In case of family MEL-001, utilizing tNGS, the authors achieved an average coverage of 722.5X with 99.73% of the target with 20x reads in the wife's sample (Fig. 1). The husband's sample had an average coverage of 763.2X with 95.36% of reads at 20x reads (Fig. 2). The gaps in the sequenced region of the husband (*CYP11B1* exons 6–9) were identified and filled up using amplicon based sequencing covering the complete coding and splice site regions of the target genes. The tNGS also provided sufficient depth in the loci of identified variants.

In a recent study by Gunsel et al., WES was utilized in determining carrier frequency of several autosomal recessive disorders [23]. The study reported highest carrier frequency of *CYP21A2* mutations (14.7%) among 100 Cypriot individuals screened. However, it is noteworthy that the reported mutations comprised only four missense and one splice variants. There were no large deletions or small indels identified in the study while large 30 kb deletion, 8 bp deletion in exon 3 ((Gly111ValfsTer21), single base insertion in exon 7 (p.F306insT) are some of the most common pseudogene derived mutations in *CYP21A2* gene reported across various populations including previous studies in Cypriot groups [24–26]. It is crucial to highlight that large deletions and gene conversion by itself account for 15–20% of the cases with 21-OH deficiency [27, 28]. The absence of these findings in the 100 subjects screened suggest a potential limitation of WES in detecting such deletions. This increases the possibility that WES might have overlooked these deletions, emphasizing the need for complementary methods to enhance the comprehensive assessment of genetic variations associated with 21-OH deficiency.

With these limitations outlined above, cost-effective FSS could serve as the preferred initial approach for CAH diagnosis and carrier screening. When large rearrangements and common point mutations in the *CYP21A2* gene have been ruled out using long-range PCR and genotyping assays, tNGS is recommended as the second line of genetic testing. This recommendation is contingent upon the availability of well-established computational pipelines and a comprehensive understanding of complex mutations in 21-OHD. WES is deemed more appropriate for investigating rare forms of CAH once the possibility of 21-OHD has been eliminated using targeted approaches.

Acknowledgements This work was carried out as a part of Ph.D. thesis of LR. The authors acknowledge Regional Centre for Biotechnology,

Faridabad to which the Ph.D programme is affiliated. They also thank DST-INSPIRE for the Ph.D. fellowship.

Authors' Contributions LR: Original draft preparation; LR, SP, RA, DV, PR: Methodology; AHS, SM, AS, SD: Clinical management and data collection; NT, AC: Supervision; AC: Study design and conceptualization; LR, RA, AC: Data analysis and interpretation; AHS, SM, AS, SD, NT: Critical revision and manuscript editing. NT and AC will act as guarantors for this manuscript.

Funding The study has been supported by Indian Council of Medical Research (ICMR) as a part of Adhoc project (F.No.54/08/2022-HUM/BUS).

Declarations

Conflict of Interest None.

References

1. El-Maouche D, Arlt W, Merke DP. Congenital adrenal hyperplasia. Lancet. 2017;390:2194–210.
2. Carrozza C, Foca L, De Paolis E, Concolino P. Genes and pseudogenes: complexity of the RCCX locus and disease. Front Endocrinol. 2021;12. https://doi.org/10.3389/fendo.2021.709758.
3. Kim JH, Kim G-H, Yoo H-W, Choi J-H. Molecular basis and genetic testing strategies for diagnosing 21-hydroxylase deficiency, including CAH-X syndrome. Ann Pediatr Endocrinol Metab. 2023;28:77–86.
4. Xu Z, Chen W, Merke DP, McDonnell NB. Comprehensive mutation analysis of the CYP21A2 gene: an efficient multistep approach to the molecular diagnosis of congenital adrenal hyperplasia. J Mol Diagn. 2013;15:745–53.
5. Ravichandran L, Korula S, Asha HS, et al. Allele-specific PCR and next-generation sequencing based genetic screening for congenital adrenal hyperplasia in India. Eur J Med Genet. 2021;64:104369.
6. Greene CN, Cordovado SK, Turner DP, Keong LM, Shulman D, Mueller PW. Novel method to characterize CYP21A2 in Florida patients with congenital adrenal hyperplasia and commercially available cell lines. Mol Genet Metab Rep. 2014;1:312–23.
7. Lee H-H. Mutational analysis of CYP21A2 gene and CYP21A1P pseudogene: long-range PCR on genomic DNA. Methods Mol Biol. 2014;1167:275–87.
8. Mathur R, Menon PSN, Kabra M, Goyal RK, Verma IC. Molecular characterization of mutations in Indian children with congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency. J Pediatr Endocrinol Metab. 2001;14:27–35.
9. Marumudi E, Sharma A, Kulshreshtha B, Khadgawat R, Khurana ML, Ammini AC. Molecular genetic analysis of CYP21A2 gene in patients with congenital adrenal hyperplasia. Indian J Endocrinol Metab. 2012;16:384–8.
10. Ravichandran L, Varghese D, R P, et al. Allele-specific and multiplex PCR based tools for cost-effective and comprehensive genetic testing in congenital adrenal hyperplasia. MethodsX. 2022;9:101748.
11. Gangodkar P, Khadilkar V, Raghupathy P, et al. Clinical application of a novel next generation sequencing assay for CYP21A2 gene in 310 cases of 21- hydroxylase congenital adrenal hyperplasia from India. Endocrine. 2021;71:189–98.
12. Xu C, Jia W, Cheng X, et al. Genotype-phenotype correlation study and mutational and hormonal analysis in a Chinese cohort with 21-hydroxylase deficiency. Mol Genet Genomic Med. 2019;7:e671.
13. Araújo RS, Mendonca BB, Barbosa AS, et al. Microconversion between CYP21A2 and CYP21A1P promoter regions causes the nonclassical form of 21-hydroxylase deficiency. J Clin Endocrinol Metab. 2007;92:4028–34.
14. Ravichandran L, Paul S, A R, et al. High carrier frequency of CYP21A2gene mutations in Southern India - underscoring the need for genetic testing in congenital adrenal hyperplasia. Endocrine. 2024;85:363–9.
15. ICMR Task Force on Inherited Metabolic Disorders. Newborn screening for congenital hypothyroidism and congenital adrenal hyperplasia. Indian J Pediatr. 2018;85:935–40.
16. Mardis ER. DNA sequencing technologies: 2006–2016. Nat Protoc. 2017;12:213–8.
17. Kong SW, Lee I-H, Liu X, Hirschhorn JN, Mandl KD. Measuring coverage and accuracy of whole exome sequencing in clinical context. Genet Med. 2018;20:1617–26.
18. Meynert AM, Ansari M, FitzPatrick DR, Taylor MS. Variant detection sensitivity and biases in whole genome and exome sequencing. BMC Bioinformatics. 2014;15:247.
19. Rabbani B, Tekin M, Mahdieh N. The promise of whole-exome sequencing in medical genetics. J Hum Genet. 2014;59:5–15.
20. Seaby EG, Pengelly RJ, Ennis S. Exome sequencing explained: a practical guide to its clinical application. Brief Funct Genomics. 2016;15:374–84.
21. Yu H, Yu H, Zhang R, et al. Targeted gene panel provides advantages over whole-exome sequencing for diagnosing obesity and diabetes mellitus. J Mol Cell Biol. 2023;15:mjad040.
22. Platt CD, Zaman F, Bainter W, et al. Efficacy and economics of targeted panel versus whole-exome sequencing in 878 patients with suspected primary immunodeficiency. J Allergy Clin Immunol. 2021;147:723–6.
23. Gunsel AS, Ergoren MC, Kemal H, et al. Determination of carrier frequency of actionable pathogenic variants in autosomal recessive genetic diseases in the Turkish Cypriot population. Genes (Basel). 2023;14:1967.
24. Skordis N, Kyriakou A, Tardy V, et al. Molecular defects of the CYP21A2 gene in Greek-Cypriot patients with congenital adrenal hyperplasia. Horm Res Paediatr. 2010;75:180–6.
25. Phedonos A, Shammas C, Skordis N, Kyriakides T, Neocleous V, Phylactou L. High carrier frequency of 21-hydroxylase deficiency in Cyprus. Clin Genet. 2013;84:585–8.
26. Neocleous V, Fanis P, Phylactou LA, Skordis N. Genotype is associated to the degree of virilization in patients with classic congenital adrenal hyperplasia. Front Endocrinol (Lausanne). 2018;9:733.
27. Hannah-Shmouni F, Chen W, Merke DP. Genetics of congenital adrenal hyperplasia. Best Pract Res Clin Endocrinol Metab. 2009;23:181–92.
28. Pignatelli D, Carvalho BL, Palmeiro A, Barros A, Guerreiro SG, Macut D. The complexities in genotyping of congenital adrenal hyperplasia: 21-hydroxylase deficiency. Front Endocrinol (Lausanne). 2019;10:432.