Molecular basis and genetic testing strategies for diagnosing 21-hydroxylase deficiency, including CAH-X syndrome

Ja Hye Kim[1],
Gu-Hwan Kim[2],
Han-Wook Yoo[3],
Jin-Ho Choi[1]

[1]Department of Pediatrics, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea
[2]Medical Genetics Center, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea
[3]Department of Pediatrics, Bundang CHA Medical Center, CHA University, Seongnam, Korea

Congenital adrenal hyperplasia (CAH) is a group of autosomally recessive disorders that result from impaired synthesis of glucocorticoid and mineralocorticoid. Most cases (~95%) are caused by mutations in the *CYP21A2* gene, which encodes steroid 21-hydroxylase. CAH patients manifest a wide phenotypic spectrum according to their degree of residual enzyme activity. *CYP21A2* and its pseudogene (*CYP21A1P*) are located 30 kb apart in the 6q21.3 region and share approximately 98% of their sequences in the coding region. Both genes are aligned in tandem with the *C4*, *SKT19*, and *TNX* genes, forming 2 segments of the RCCX modules that are arranged as *STK19-C4A-CYP21A1P-TNXA-STK19B-C4B-CYP21A2-TNXB*. The high sequence homology between the active gene and pseudogene leads to frequent microconversions and large rearrangements through intergenic recombination. The *TNXB* gene encodes an extracellular matrix glycoprotein, tenascin-X (TNX), and defects in *TNXB* cause Ehlers-Danlos syndrome. Deletions affecting both *CYP21A2* and *TNXB* result in a contiguous gene deletion syndrome known as CAH-X syndrome. Because of the high homology between *CYP21A2* and *CYP21A1P*, genetic testing for CAH should include an evaluation of copy number variations, as well as Sanger sequencing. Although it poses challenges for genetic testing, a large number of mutations and their associated phenotypes have been identified, which has helped to establish genotype-phenotype correlations. The genotype is helpful for guiding early treatment, predicting the clinical phenotype and prognosis, and providing genetic counseling. In particular, it can help ensure proper management of the potential complications of CAH-X syndrome, such as musculoskeletal and cardiac defects. This review focuses on the molecular pathophysiology and genetic diagnosis of 21-hydroxylase deficiency and highlights genetic testing strategies for CAH-X syndrome.

Keywords: 21-hydroxylase deficiency, Congenital adrenal hyperplasia, Copy number variation, *CYP21A2*

Address for correspondence:
Jin-Ho Choi
Department of Pediatrics, Asan Medical Center, University of Ulsan College of Medicine, 88, Olympic-ro 43-Gil, Songpa-Gu, Seoul 05505, Korea
Email: jhc@amc.seoul.kr
https://orcid.org/0000-0003-1196-7826

Highlights

· The *CYP21A2* gene is located in the major histocompatibility complex class III region on chromosome 6p21.3, approximately 30 kb apart from its pseudogene, *CYP21A1P*.
· Molecular diagnosis of 21-hydroxylase deficiency (21-OHD) is challenging because of high sequence homology between *CYP21A2* and *CYP21A1P*, with approximately 98% homology at the exon level and 96% homology at the intron level.
· The phenotype of 21-OHD varies according to the degree of residual enzyme activity.
· Genetic diagnosis is helpful in guiding appropriate management and predicting the phenotype and prognosis of patients with 21-OHD.

Introduction

Congenital adrenal hyperplasia (CAH) is a heterogeneous group of enzyme deficiencies in the steroidogenic pathway of the adrenal cortex that result in impaired synthesis of cortisol and aldosterone by the adrenal glands.[1] Impaired cortisol synthesis causes prolonged elevation of adrenocorticotropic hormone, which leads to overstimulation of the adrenal cortex and subsequent adrenal gland hyperplasia. More than 95% of patients with CAH have steroid 21-hydroxylase deficiency (21-OHD), which is the lack of an enzyme encoded by the *CYP21A2* gene.[2] This condition is classified into classic and nonclassic forms based on the residual enzyme activity. The classic form is defined by severely reduced or absent enzyme activity, and it is subdivided into the salt-wasting (SW) form with little or no enzyme activity and the simple virilizing (SV) form with 1% to 5% of residual enzyme activity. Nonclassical CAH (NCCAH) is the mildest form of CAH; it does not cause virilization or a life-threatening adrenal crisis.

The *CYP21A2* gene spans 3.35 kb and consists of 10 exons.[3,4] It is located in the major histocompatibility complex class III region of chromosome 6p21.3, and tandem repeats of the active genes and its pseudogenes are aligned within the RCCX module. The tandem repeats often cause complex genomic recombination events between the *CYP21A2* and *CYP21A1P* genes, which are approximately 30 kb apart.

This review describes the molecular mechanism of 21-OHD, its genotype-phenotype correlations, the extended phenotypes of CAH-X syndrome, and genetic testing strategies. Many studies have reported inconsistent nomenclature of mutations in *CYP21A2* that result from numbering the initiation codon as 0. In this review, mutations are described according to nomenclature guidelines from the Human Genome Variation Society (https://varnomen.hgvs.org/), in which nucleotide numbering begins with the translation initiation codon (ATG) as the first codon number, and subsequent nucleotides are numbered from the 5′ to 3′ direction.[5]

Genetics of the *CYP21A2* gene and molecular mechanisms of 21-OHD

1. Genetics of *CYP21A2* and organization of the RCCX module

The *CYP21A2* gene is located in a complex organization of genes known as the RCCX module, which is a tandem repeat structure containing a series of functional and nonfunctional genes: serine/threonine kinase 19 (*STK19*, formerly *RP*), complement 4 (*C4*), steroid 21-hydroxylase (*CYP21*), and tenascin-X (*TNX*) (Fig. 1A).[3,6] The RCCX module consists of 2 segments, the short and long modules.[7] The short module contains *STK19B*, *C4B*, *CYP21A2*, and part of *TNXB*. The long module consists of part of the *STK19* gene, the full length of the *C4A* gene, the *CYP21A1P* gene, and the *TNXA* gene. Because of the high sequence homology between the active gene and the pseudogene, microconversions and large rearrangements occur through intergenic recombination.[1]

The *STK19* gene, which encodes a nuclear serine/threonine kinase protein, was recently proposed to be a regulator of
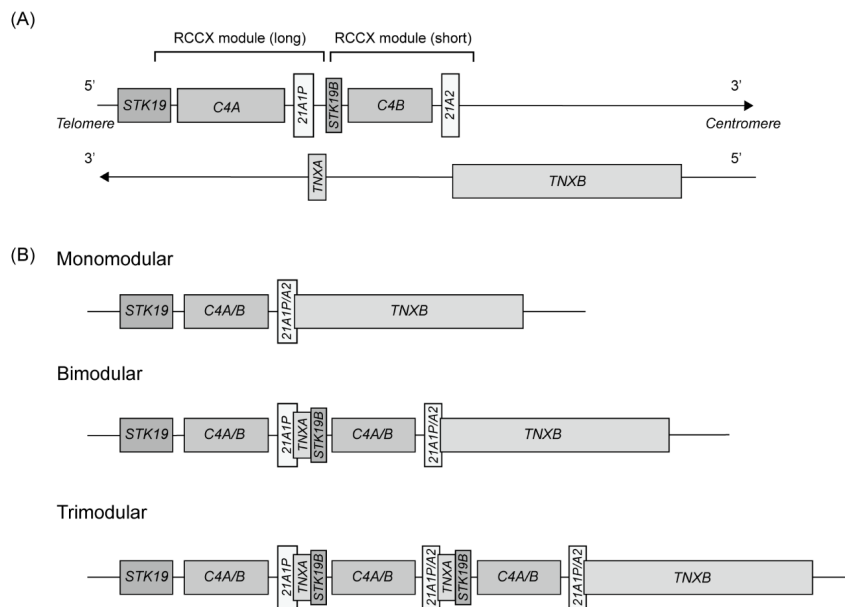

Fig. 1. Organization of the human RCCX module and copy number variations on chromosome 6p21. (A) Schematic representation of the RCCX module featuring 2 segments with the following gene arrangement: *STK19* (formerly RP1)-*C4A-CYP21A1P-TNXA-STK19B-C4B-CYP21A2-TNXB*. The arrow indicates transcriptional orientation. (B) Three different RCCX module arrangements: monomodular, bimodular, and trimodular structures.

NRAS activity.[8] STK19 phosphorylates NRAS, which facilitates interactions between NRAS and its downstream effectors, increasing the activation of the mitogen-activated protein kinase cascade. *C4A* (active gene) and *C4B* (pseudogene), each contain 41 exons and are highly homologous and essential elements of the humoral immune response. The size of the human *C4* gene varies from 16 kb to 22 kb due to the integration of the endogenous retroviral sequence (6.7 kb), human endogenous retrovirus-K (HERV-K), in intron 9 of the long C4 gene.[9] Approximately 65% of human *C4* gene contains HERV-K viral integration. Although no evidence links HERV-K to human disease, aberrant expression of these retroviral elements is associated with the development of autoimmune diseases.[10,11] Thus, C4 deficiency correlates strongly with autoimmune diseases such as systemic lupus erythematosus.[12] The transcription start site (5'-end) of the *CYP21A1P* and *CYP21A2* genes is located 2,466 bp downstream of both *C4A* and *C4B*.[13,14]

The *CYP21A2* gene encodes 21-hydroxylase, which plays a critical role in the synthesis of the 2 principal steroid hormones, aldosterone and cortisol. *CYP21A2* and its pseudogene *CYP21A1P* share a high degree of sequence homology, with 98% sequence identity in the coding region and 96% in the intronic region.[4,15] The pseudogene is inactivated by several pathogenic variants that prevent the synthesis of a functional protein.

The *TNXA* and TNXB genes have an opposite transcriptional direction compared with other cluster genes.[6,16] The 3'-ends of the *CYP21* genes overlap with the last exons of *TNXA* and *TNXB* (Fig. 1A).[6] The *TNXB* gene encodes an extracellular matrix glycoprotein, TNX, that consists of 44 exons spanning 68.2 kb. Defects in *TNXB* cause Ehlers-Danlos syndrome (EDS).[16] In contrast, *TNXA* is a truncated pseudogene spanning 4.5 kb and is homologous to *TNXB* from exons 32 to 44.

A germline copy number variation (CNV) is a fragment of DNA with variable copies that result from duplications or deletions during evolution. Multi-allelic CNVs are genomic segments that exhibit a variable number of copies among individuals. They can affect gene expression, protein function, and phenotypic traits, thereby contributing to genetic and phenotypic heterogeneity. The RCCX module is one of the most complex CNV loci in humans.[3] In Caucasians, RCCX CNV alleles typically have monomodular, bimodular, or trimodular segments, with a prevalence of approximately 17%, 69%, and 14%, respectively (Fig. 1B).[10] The genetic variations in RCCX CNVs make it difficult to accurately characterize molecular defects in *CYP21A2*.

2. Molecular mechanisms of genetic defects in *CYP21A2*

1) Point mutations by microconversion

As mentioned above, the molecular diagnosis of 21-OHD is challenging because of high sequence homology between *CYP21A2* and *CYP21A1P*. They differ by only 65 nucleotides in their coding and intronic regions.[4] Microconversion of sequence variants from *CYP21A1P* to *CYP21A2* is responsible for nonfunctional protein synthesis in 70%–75% of cases (Fig. 2).[2] Deleterious variants are transferred by small conversions from the pseudogene during meiosis. Microconversion events are caused by major recurrent variants: a splicing mutation (c.293-13A/C>G, also known as I2G), an 8-bp deletion in exon 3 (p.G111Vfs), one nucleotide insertion in exon 7 (p.L308Ffs), 1 nonsense mutation (p.Q319*), 3 missense mutations (p.P31L, p.I173N, and p.R357W), and 1 cluster conversion (p.I237N, p.V238E, and p.M240K) (Fig. 2).[17,18]

2) Large gene rearrangement (chimeric *CYP21A1P/CYP21A2* and *TNXA/TNXB*) by unequal meiotic crossover

In the remaining 25%–30% of cases, large gene rearrangements due to unequal crossover during meiosis result in large deletion, duplications, or other contiguous gene deletions.[2] Chimeric *CYP21A1P/CYP21A2* genes occur by homologous
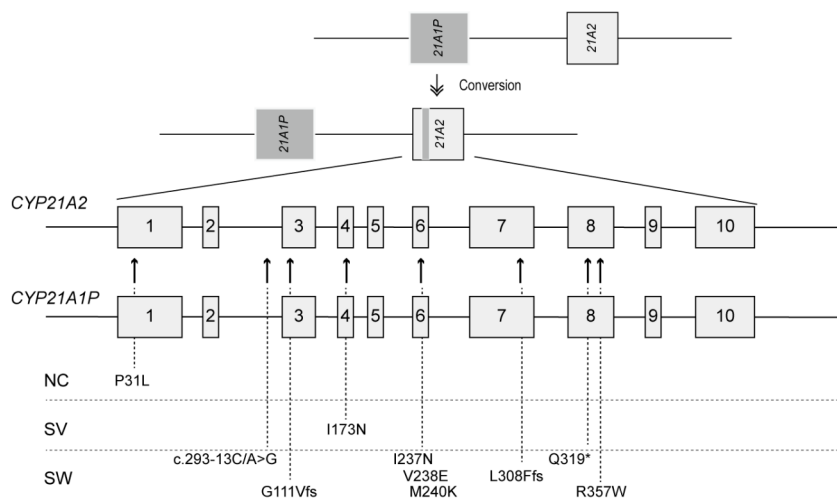
Fig. 2. Schematic representation of microconversion events and the location of common mutations within the *CYP21A2* gene transferred by gene conversion. NC, nonclassical; SV, simple virilizing; SW, salt-wasting.

recombination between these 2 genes. A 26- or 32-kb deletion (depending on the size of *C4B*) in the region involved in the 3'-end of *CYP21A1P, TNXA*, and *C4B* and the 5'-end of *CYP21A2* creates a single, nonfunctional chimeric gene (Fig 3A). To date, 9 different chimeric *CYP21A1P/CYP21A2* genes have been characterized based on the chimeric junction site (Fig. 3A).[19] These chimeric genes have been classified into classic and attenuated forms, depending on whether the junction site is upstream or downstream of the I2G mutation.[19] Seven chimeras (CH1, CH2, CH3, CH5, CH6, CH7, and CH8) carrying this splicing mutation (I2G) are associated with a severe SW phenotype.[18,19] In contrast, 2 chimeras (CH4 and CH9) harboring the promoter of *CYP21A1P* and the p.P31L mutation are associated with an attenuated phenotype.[19]

3) *De novo* spontaneous mutations in *CYP21A2*

Fewer than 5% of mutations in *CYP21A2* are not due to gene conversions.[2] Most of these variants are private and missense mutations. Founder effects have been observed in these pseudogene-independent mutations.[20-22]

3. Genotype−phenotype correlations for mutations in *CYP21A2*

Common major mutations correlate with phenotypes

that reflect varying degrees of residual enzyme activity (Fig. 4). Mutations in *CYP21A2* are categorized into 4 groups (0, A, B, and C) based on *in vitro* data according to the level of enzyme activity.[23] Group 0 is caused by null mutations, such as a large deletion, p.L308Ffs, p.Q319*, and p.R357W, that leave no enzyme activity. Group A consists of mutations that cause severely impaired enzyme activity, such as a homozygous or compound heterozygous I2G mutation. Group B reflects either homozygous or compound heterozygous for the p.I173N mutation with group 0 or A variants. Group C consists of milder homozygous or compound heterozygous mutations, such as p.V282L, p.P454S, and p.P31L, which retain 20%−60% of residual enzyme activity.[1]

Because most patients with 21-OHD harbor compound heterozygous mutations, the allele with the less severe mutation tends to have a greater influence on the clinical phenotype. Severe genotypes (group 0 and A) correlate well with the SW form, ranging from 91%−97%; however, this relationship is less strong for the intermediate severity groups (group B and C), which range from 45%−57%.[24] For example, patients carrying the p.P31L mutation predominantly have NCCAH, but that mutation has also been reported in the SV type.[1,24,25] This genotype-phenotype discordance could be explained by a genetic modifier. Combined variants of p.P31L and 4 single nucleotide polymorphisms in the promoter region (c.-126C>T, c.-113G>A, c.-110T>C, and c.-103A>G) have been shown to
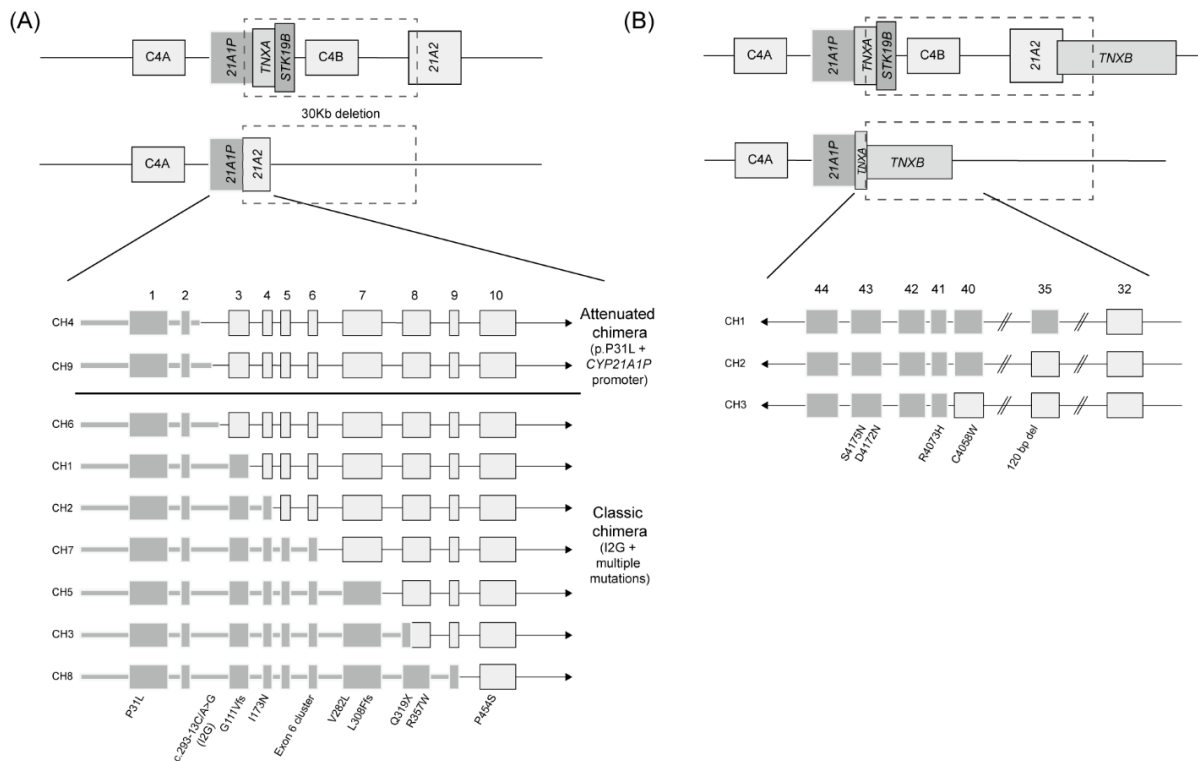

Fig. 3. Gene rearrangement due to unequal crossover during meiosis. (A) *CYP21A2* gene rearrangement caused by a 30-kb deletion, leading to a *CYP21A1P/CYP21A2* chimera. The figure illustrates the 9 types of chimera according to the different junction sites. Seven chimeras harboring the I2G mutation are associated with the classic salt-wasting phenotype, and 2 chimeras carrying a p.P31L mutation exhibit an attenuated phenotype. (B) Formation of a chimeric TNXA/TNXB gene that deletes the *CYP21A2* gene. Three different chimeric genes are demonstrated, depending on the junction site.

reduce transcriptional activity by 5-fold and enzyme activity by less than 6%.[26] This reduction leads to a more severe phenotype than would be expected on the basis of the genotype alone. Because little research on genetic modifiers in CAH has been done, further studies are needed to fully understand their role. Therefore, genotype-phenotype correlations are not always absolute, and clinical management should be based on clinical features and hormonal data rather than the genotype.[1,25,27]

4. Mutation spectrum of *CYP21A2* according to ethnicity

The worldwide incidence of classic CAH, based on neonatal screening, ranges from 1:14,000 to 1:18,000 births.[28] The prevalence of the major mutations varies by ethnicity and geographic region (Table 1). Deletion/conversion, I2G, and p.I173N mutations are the most common in most populations.[1] However, the p.V282L mutation is quite common in Ashkenazi Jews with 21-OHD, and it has rarely been observed in Asians.[27] The I2G and p.V282L mutations are more prevalent in the Middle Eastern population than elsewhere, and the p.Q319* mutation is the most common in the Tunisian population.[29,30] In Korean patients, large deletions are the most common mutation, followed by the I2G and p.I173N.[31] This ethnic diversity and specificity of *CYP21A2* mutations could facilitate targeted genetic screening and improve the efficiency of 21-OHD diagnosis.

CAH-X syndrome

Deletions affecting both the *CYP21A2* and *TNXB* genes result in a contiguous gene deletion syndrome known as CAH-X syndrome, which is characterized by a hypermobile form of EDS.[14,32] The tenascins are widely-expressed extracellular matrix proteins, and TNX is expressed in the connective tissues, such as skeletal muscle, heart, and blood vessels.[16,32] The role of TNX has been confirmed by *Tnx*-knockout mice, which have reduced collagen content and connective tissue architecture, suggesting that TNX plays an essential role in regulating collagen deposition in connective tissues.[33] The incidence of CAH-X syndrome is estimated to be 7%–15% of patients with 21-OHD.[34-36] However, the exact prevalence of the condition remains unknown because it is often underdiagnosed due to the variability of the phenotype and lack of awareness among physicians.

The *TNXA* gene is a partially duplicated gene segment that shares a high degree of sequence homology with *TNXB* from intron 32 to exon 45. *TNXA* is truncated at the 5' end and harbors a 120-bp deletion causing a frameshift and premature termination of translation.[37] A chimeric *TNXA/TNXB* gene is caused by unequal gene crossover in the RCCX module,

Fig. 4. Classification of common mutations in *CYP21A2* based on their residual enzyme activity.

Table 1. Allele frequencies in various regions

Variable	USA	Brazil	UK	Germany	China	Japan	South Korea
Alleles (n)	3,005	856	306	310	460	98	144
Allele frequencies (%)							
Deletions/duplications/rearrangements	20	9	31.9	20.3	25.2	5.1	31.3
I2G	22.9	21.1	12.6	30.3	35	39.8	28.5
p.I173N	8.2	7.5	12.6	19.7	14.3	14.3	15.3
p.Q319*	3.5	6.1	3.3	4.8	4.6	3.1	2.8
p.R357W	3.6	5.4	3.6	4.5	5.9	15.3	9
p.G111Vfs*21	2.1	1.8	0.5	1.6	4.3	4.1	0
E6 cluster	2.1	1.2	1.1	1	1.3	5.1	1.4
p.P31L	2.6	0.6	0.8	2.6	0.2	4.1	0
p.V282L	23.9	26.6	12.6	2.9	0.2	0	0
Others	11.1	20.7	21	12.3	9	9.1	11.7
Reference No.	27	21,52	53	54	55	56	31

resulting in the complete deletion of the *STK19B-C4B-CYP21A2* genes with a fraction of the *TNXA* and *TNXB* genes (Fig. 3B).[38] Three different *TNXA/TNXB* chimeras have been reported based on the location of the junction site.[39] CAH-X CH1 is caused by a 120 bp deletion in exon 35, which can be detected by a multiplex-ligation dependent probe amplification (MLPA) analysis. Because exons 32–34 and exons 36–44 are highly homologous, CNV detection in an MLPA analysis is difficult. CAH-X CH2 is characterized by an intact exon 35 and the p.C4058W mutation in exon 40. CAH-X CH3 has a cluster of 3 pseudogene-derived variants: p.R4073H in exon 41, p.D4172N in exon 43, and p.S4175N in exon 43.[38]

To date, the clinical features of more than 50 patients with CAH-X syndrome have been described.[34-36,38,40] Overall, skin laxity and musculoskeletal features, such as joint hypermobility and subluxation, are common. Chronic arthralgia is more prominent in adult patients with CAH-X syndrome than in children. Cardiac abnormalities such as septal defects, cardiac chamber enlargement, or great vessel enlargement occur in about 25% of patients with this condition. These patients also have gastrointestinal abnormalities, including gastroesophageal reflux, hernia, and rectal prolapse. Other findings include early-onset osteoarthritis, scoliosis, pectus excavatum, osteoporosis, bifid or elongated uvula, easy bruising, and delayed wound healing. Patients with biallelic deletion of *TNXB* have more severe clinical manifestations than those with monoallelic deletion.[35]

Molecular genetic testing in patients with 21-OHD

Molecular genetic testing for 21-OHD is a complex process due to the genomic structure of *CYP21A2* and the surrounding region. Various disease-causing mutations include point mutations, large gene rearrangements, multiple mutations in *cis*, and CNVs of variable sizes. Large gene rearrangements can be detected by Southern blot analyses, real-time quantitative polymerase chain reaction (PCR), and MLPA analyses. Currently, Sanger sequencing combined with an MLPA analysis is commonly used and can detect most mutations.

1. Southern blot analysis and quantitative PCR

Southern blot analyses were originally used to detect large gene deletions or duplications. This method uses restriction endonucleases that recognize short DNA sequences and cleave double-stranded DNA. The most commonly used restriction endonuclease is *TaqI*, which recognizes the palindromic DNA sequence of 5'-T^CGA-3', where the caret symbol (^) indicates the cleavage site. This enzyme produces 3.7- and 2.5-kb fragments of intact *CYP21A2* and the partial *TNXB* gene, respectively, and 3.2- and 2.4-kb fragments of *CYP21A1P* and the partial *TNXA* gene, respectively (Fig. 5A).[41] The 3.2- and 3.7-kb fragments are indicative of the *CYP21A1P* and *CYP21A2* genes by their size. Southern blot analyses show that the 3.2-kb fragment without the 3.7-kb fragment represents deletion of the *CYP21A2* gene (Fig. 5A). However, this method is time-consuming and requires large amounts of DNA.

Real-time, quantitative PCR is a rapid and sensitive method for detecting gene deletions and duplications. This method uses *CYP21A2*-specific and internal control gene primer sets and measures the threshold cycle (Ct) values of both the target and control genes. To assess the *CYP21A2* copy number, the difference in Ct for *CYP21A2* and the internal control gene is calculated and plotted against the logarithm of the *CYP21A2*/internal control ratio (%). However, some discrepancies with Southern blot analyses can occur due to limited discrimination between 2 and 3 RCCX copies per genome.[42]

2. PCR–based sequence analysis

Sanger sequencing allows the detection of point mutations and small insertions/deletions. PCR with oligonucleotide-specific primers designed for differential amplification of *CYP21A2* (3.48 kb) and *CYP21A1P* (3.49 kb) (Fig. 5B) has been widely used.[31,43] The alleles containing the fusion gene can be detected using a primer set containing *CYP21A1P*-specific forward and *CYP21A2*-specific reverse primers. A *CYP21A2*-specific forward primer and a *CYP21A1P*-specific reverse primer can be used to detect alleles containing a rearranged gene. PCR products using *CYP21A2*-specific primer sets are analyzed by automatic direct sequencing to detect point mutations.

When designing primers for PCR-based sequencing, it is crucial to ensure that they are specific for the target gene or region of interest to prevent amplification of the pseudogene. Nonamplification by PCR can result from allelic dropout, which must be considered, as for any other gene.[44] Several single nucleotide polymorphisms in the promoters of both genes can interfere with primer binding, resulting in allelic dropout.

Another well-established technique is long-range PCR for an 8.5-kb fragment using CYP779f/Tena32F primers (Fig. 5C) that encompass the entire *CYP21A2* gene and part of *TNXB*.[45] CYP779f (5'-AGGTGGGCTGTTTTCCTTTCA-3') is a common forward primer for the 5'-untranslated region (UTR) of the *CYP21A2* and *CYP21A1P* genes. The CYP779f primer anneals from c.-799 to c.-779 based on the nucleotide sequences first described by Higashi et al. in 1986 (c.-802 to c.-782 based on NM_000500.9).[3,46] The reverse primer (Tena32F) (5'-CTGTGCCTGGCTATAGCAAGC-3') anneals specifically to exon 32 of *TNXB*. This method can be used to detect *TNXA/TNXB* chimeras in patients with CAH-X syndrome. The amplified PCR products are then subjected to Sanger sequencing.

3. Multiplex–ligation dependent probe amplification analyses

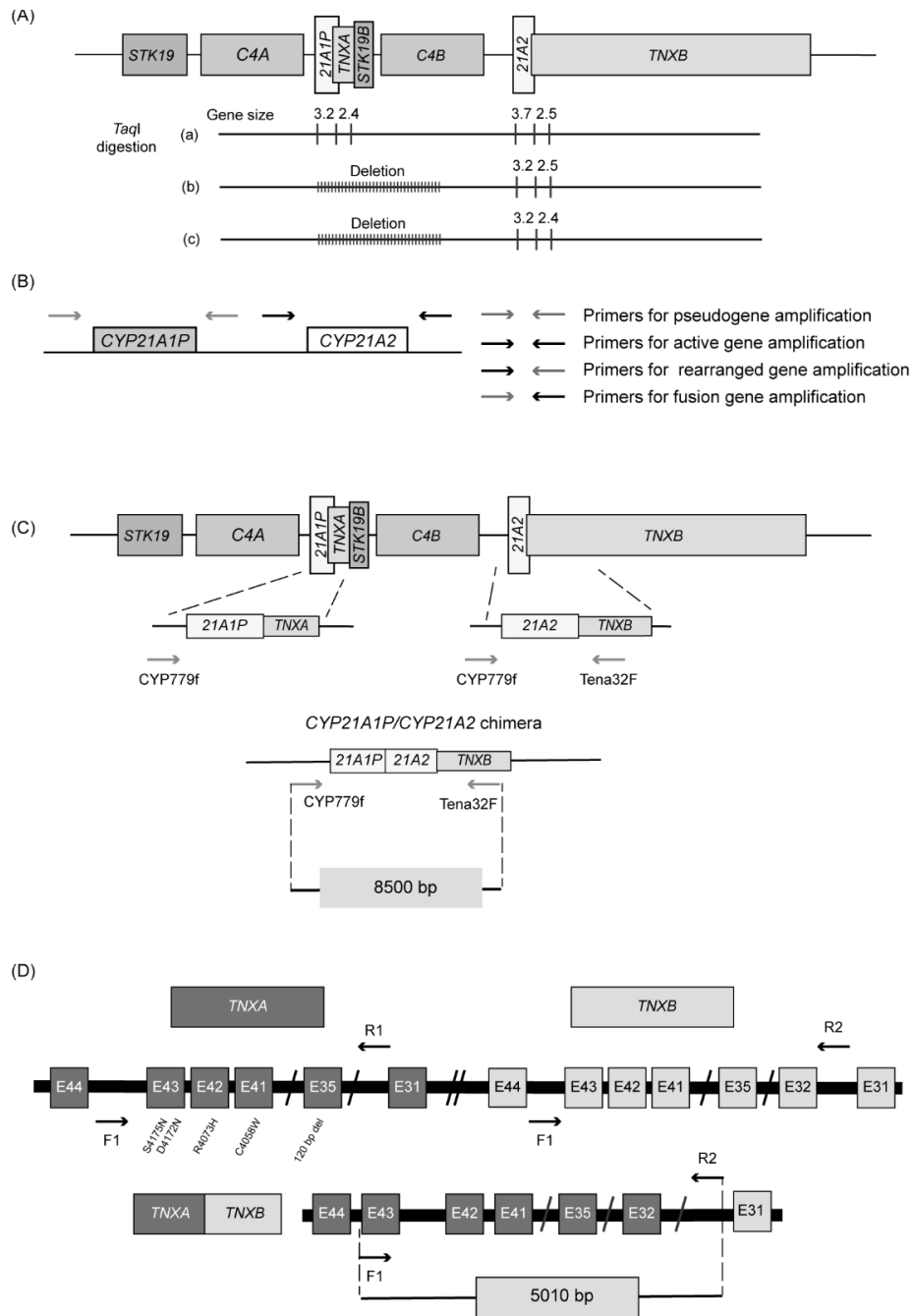The MLPA analysis is a robust method that can detect gene

Fig. 5. Schematic representation of a molecular genetic analysis of *CYP21A2*. (A) Southern blot analysis digested with *Taq*I restriction endonuclease. This method distinguishes the *CYP21A2* (3.7 kb) and *CYP21A1P* (3.2 kb) genes by fragment size. (B) Sequence-specific PCR amplification of the *CYP21A1P* and *CYP21A2* genes. Two set of primers create 4 possible pairwise reactions to obtain amplicons corresponding to *CYP21A1P*, *CYP21A2*, the fusion gene, and rearrangement products.[31] (C) Sequencing for a *CYP21A1P/CYP21A2* chimera using paired primers, CYP779f/Tena32F. These primers amplify the 8.5-kb fragment. The Tena32F primer is located in a nonduplicated area of *TNXB* in exon 32, and the CYP779f primer is complementary to the 5'-end of the *CYP21A1P* and *CYP21A2* genes. (D) Sequencing for a *TNXA/TNXB* chimera using mixed paired primers. The forward primer (F1) contains complementary sequences for intron 43 of *TNXA* and *TNXB*, and the reverse primer (R2) is specific for intron 31 of *TNXB*.

deletions, rearrangements, and fusion genes, and it requires only small amounts of DNA. A commercially available *CYP21A2*-MLPA kit (SALSA MLPA Probemix P050 CAH, MRC Holland, Amsterdam, The Netherlands) is widely used.[44] It contains 4 probes for the *CYP21A1P* gene and 8 probes for the *CYP21A2* gene, enabling it to detect gene deletions, large gene conversions,

a single nucleotide variant in the 5'-UTR (c.-113G>A), an 8 bp deletion in exon 3, p.I173N in exon 4, E6 cluster mutations in exon 6, and p.L308Ffs in exon 7. In addition, the probe mix contains 6 probes for the *TNXB* gene and can detect a 120 bp deletion at the boundary of exon 35 and intron 35 (CAH-X CH1). This method is limited to detecting variants that lie outside the target probe. Despite the challenges associated with genetic testing for 21-OHD, accurate genetic testing (up to 98%) can be achieved using PCR-based sequencing and MLPA.[44]

4. Molecular genetic diagnosis of CAH−X syndrome

PCR-based detection of chimeric genes and MLPA analyses have been carried out to confirm the diagnosis of CAH-X syndrome.[35] PCR-based detection of chimeric genes has been performed using a mixed-primer strategy: a common forward primer for intron 43 of *TNXA* and *TNXB* and a reverse primer for intron 31 of *TNXB* (Fig. 5D). This produces a 5,010 bp fragment of *TNXB*, from part of intron 31 to part of intron 43 (Fig. 5D). This PCR product is used as a template for the subsequent sequencing analysis to assess *TNXA/TNXB* chimeras. CAH-X CH1 alleles are assessed by Sanger sequencing of exon 35 in *TNXB*, whereas CH2 and CH3 alleles are identified by direct sequencing of exons 40, 41, and 43. An MLPA analysis can detect a 120-bp deletion in *TNXB* exon 35 that is derived from the *TNXA* pseudogene. However, that is a laborious process and is not routinely performed. Therefore, the diagnosis of CAH-X syndrome relies primarily on clinical assessment using the Beighton 9-point scale.[47]

5. Future directions

Recently, next-generation sequencing techniques have been increasingly used in clinical diagnostics, replacing traditional sequencing technologies. Because pseudogenes have high sequence homology, short read sequencing is limited in its ability to accurately map the reads to the correct genomic location. In addition, the complex genomic structure of RCCX module produces structural variations that are difficult to characterize using short-read sequencing technologies.[48]

Therefore, long-read sequencing (LRS) technologies have been developed, such as single molecule real time (SMRT) sequencing (Pacific Biosciences, Menlo Park, CA, USA) and Oxford Nanopore sequencing (Oxford Nanopore Technologies, Oxford, UK).[49] These techniques generate reads long enough to investigate repetitive or homologous regions. Long-range PCR encompassing the full-length of the *CYP21A1P/TNXA* and *CYP21A2*/partial *TNXB* genes that is followed by SMRT sequencing can detect various mutations, including point mutations and small insertions/deletions.[50] Locus-specific PCR using CYP779f/Tena32F followed by SMRT sequencing using the Sequel II platform (Pacific Biosciences) successfully identified pathogenic variants.[51] However, those studies showed that current LRS by PCR of a partial region is still limited in its ability to find CNVs and complex rearrangements. In addition, specialized bioinformatic algorithms are required to accurately detect mutations in highly homologous regions. Advances in bioinformatic tools and reductions in the cost of LRS-based testing will mitigate current challenges in the future.

Conclusions

The *CYP21A2* gene is located on the RCCX module, which is a complex, multiallelic, and tandem CNV. This highly homologous gene structure can lead to genetic rearrangement during meiosis. Genetic testing is challenging due to the presence of pseudogenes and complex genomic structures. Recent extensive studies of the RCCX modules have led to a better understanding of the spectrum of TNX-related disorders. Identification of CAH-X syndrome is important to enable early intervention for musculoskeletal and cardiac abnormalities. The combination of clinical and biochemical markers with appropriate genetic testing promotes better outcomes in patients with 21-OHD.

Notes

Author contribution: Conceptualization: JHK, JHC; Data curation: JHC; Formal analysis: JHC; Funding acquisition: JHC; Methodology: JHC; Project administration: JHC; Visualization: JHC; Writing - original draft: JHK, JHC; Writing - review & editing: GHK, HWY, JHC

ORCID
Ja Hye Kim: 0000-0002-2477-5199
Gu-Hwan Kim: 0000-0002-7056-7446
Han-Wook Yoo: 0000-0001-8615-186X
Jin-Ho Choi: 0000-0003-1196-7826

References

1. Claahsen-van der Grinten HL, Speiser PW, Ahmed SF, Arlt W, Auchus RJ, Falhammar H, et al. Congenital adrenal hyperplasia-Current insights in pathophysiology, diagnostics, and management. Endocr Rev 2022;43:91-159.

2. 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.

3. Higashi Y, Yoshioka H, Yamane M, Gotoh O, Fujii-Kuriyama Y. Complete nucleotide sequence of two steroid 21-hydroxylase genes tandemly arranged in human chromosome: a pseudogene and a genuine gene. Proc Natl Acad Sci U S A 1986;83:2841-5.

4. White PC, New MI, Dupont B. Structure of human steroid 21-hydroxylase genes. Proc Natl Acad Sci U S A 1986;83:5111-5.

5. den Dunnen JT, Dalgleish R, Maglott DR, Hart RK, Greenblatt MS, McGowan-Jordan J, et al. HGVS recommendations for the description of sequence variants: 2016 update. Hum Mutat 2016;37:564-9.

6. Morel Y, Bristow J, Gitelman SE, Miller WL. Transcript encoded on the opposite strand of the human steroid 21-hydroxylase/complement component C4 gene locus. Proc Natl Acad Sci U S A 1989;86:6582-6.

7. Lee H-H. Chimeric CYP21P/CYP21 and TNXA/TNXB genes in the RCCX module. Mol Genet Metab 2005;84:4-8.

8. Yin C, Zhu B, Zhang T, Liu T, Chen S, Liu Y, et al. Pharmacological targeting of STK19 inhibits oncogenic NRAS-driven melanomagenesis. Cell 2019;176:1113-27. e16.

9. Tassabehji M, Strachan T, Anderson M, Campbell RD, Collier S, Lako M. Identification of a novel family of human endogenous retroviruses and characterization of one family member, HERV-K(C4), located in the complement C4 gene cluster. Nucleic Acids Res 1994;22:5211-7.

10. Blanchong CA, Zhou B, Rupert KL, Chung EK, Jones KN, Sotos JF, et al. Deficiencies of human complement component C4a and C4b and heterozygosity in length variants of RP-C4-CYP21-TNX (Rccx) modules in Caucasians: the load of Rccx genetic diversity on major histocompatibility complex–associated disease. J Exp Med 2000;191:2183-96.

11. Grandi N, Cadeddu M, Pisano MP, Esposito F, Blomberg J, Tramontano E. Identification of a novel HERV-K(HML10): comprehensive characterization and comparative analysis in non-human primates provide insights about HML10 proviruses structure and diffusion. Mob DNA 2017;8:15.

12. Carroll MC. Complement and humoral immunity. Vaccine 2008;26:I28-33.

13. Wijesuriya SD, Zhang G, Dardis A, Miller WL. Transcriptional regulatory elements of the human gene for cytochrome P450c21 (steroid 21-hydroxylase) lie within intron 35 of the linked C4B gene. J Biol Chem 1999;274:38097-106.

14. Miller WL, Merke DP. Tenascin-X, Congenital adrenal hHyperplasia, and the CAH-X syndrome. Horm Res Paediatr 2018;89:352-61.

15. White PC, Grossberger D, Onufer BJ, Chaplin DD, New MI, Dupont B, et al. Two genes encoding steroid 21-hydroxylase are located near the genes encoding the fourth component of complement in man. Proc Natl Acad Sci U S A 1985;82:1089-93.

16. Bristow J, Tee MK, Gitelman SE, Mellon SH, Miller WL. Tenascin-X: a novel extracellular matrix protein encoded by the human XB gene overlapping P450c21B. J Cell Biol 1993;122:265-78.

17. Choi JH, Kim GH, Yoo HW. Recent advances in biochemical and molecular analysis of congenital adrenal hyperplasia due to 21-hydroxylase deficiency. Ann Pediatr Endocrinol Metab 2016;21:1-6.

18. Concolino P, Costella A. Congenital adrenal hyperplasia (CAH) due to 21-hydroxylase deficiency: a comprehensive focus on 233 pathogenic variants of CYP21A2 Gene. Mol Diagn Ther 2018;22:261-80.

19. Chen W, Xu Z, Sullivan A, Finkielstain GP, Van Ryzin C, Merke DP, et al. Junction site analysis of chimeric CYP21A1P/CYP21A2 genes in 21-hydroxylase deficiency. Clin Chem 2012;58:421-30.

20. Billerbeck AE, Mendonca BB, Pinto EM, Madureira G, Arnhold IJ, Bachega TA. Three novel mutations in CYP21 gene in Brazilian patients with the classical form of 21-hydroxylase deficiency due to a founder effect. J Clin Endocrinol Metab 2002;87:4314-7.

21. de Carvalho DF, Miranda MC, Gomes LG, Madureira G, Marcondes JAM, Billerbeck AEC, et al. Molecular CYP21A2 diagnosis in 480 Brazilian patients with congenital adrenal hyperplasia before newborn screening introduction. Eur J Endocrinol 2016;175:107-16.

22. Loidi L, Quinteiro C, Parajes S, Barreiro J, Lestón DG, Cabezas-Agrícola JM, et al. High variability in CYP21A2 mutated alleles in Spanish 21-hydroxylase deficiency patients, six novel mutations and a founder effect. Clin Endocrinol (Oxf) 2006;64:330-6.

23. Speiser PW, Dupont J, Zhu D, Serrat J, Buegeleisen M, Tusie-Luna MT, et al. Disease expression and molecular genotype in congenital adrenal hyperplasia due to 21-hydroxylase deficiency. J Clin Invest 1992;90:584-95.

24. Riedl S, Röhl FW, Bonfig W, Brämswig J, Richter-Unruh A, Fricke-Otto S, et al. Genotype/phenotype correlations in 538 congenital adrenal hyperplasia patients from Germany and Austria: discordances in milder genotypes and in screened versus prescreening patients. Endocr Connect 2019;8:86-94.

25. Kocova M, Anastasovska V, Falhammar H. Clinical outcomes and characteristics of P30L mutations in congenital adrenal hyperplasia due to 21-hydroxylase deficiency. Endocrine 2020;69:262-77.

26. Araujo RS, Billerbeck AE, Madureira G, Mendonca BB, Bachega TA. Substitutions in the CYP21A2 promoter explain the simple-virilizing form of 21-hydroxylase deficiency in patients harbouring a P30L mutation. Clin Endocrinol (Oxf) 2005;62:132-6.

27. New MI, Abraham M, Gonzalez B, Dumic M, Razzaghy-Azar M, Chitayat D, et al. Genotype–phenotype correlation in 1,507 families with congenital adrenal hyperplasia owing to 21-hydroxylase deficiency. Proc Natl Acad Sci U S A 2013;110:2611-6.

28. Speiser PW, Arlt W, Auchus RJ, Baskin LS, Conway GS, Merke DP, et al. Congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency: an endocrine society clinical practice guideline. J Clin Endocrinol Metab 2018;103:4043-88.

29. Romdhane L, Kefi R, Azaiez H, Halim NB, Dellagi K, Abdelhak S. Founder mutations in Tunisia: implications for diagnosis in North Africa and Middle East. Orphanet J Rare Dis 2012;7:52.

30. Kharrat M, Tardy Vr, M'Rad R, Maazoul F, Jemaa LB,

Refaï M, et al. Molecular genetic analysis of Tunisian patients with a classic form of 21-hydroxylase deficiency: identification of four novel mutations and high prevalence of Q318X mutation. J Clin Endocrinol Metab 2004;89:368-74.

31. Choi JH, Jin HY, Lee BH, Ko JM, Lee JJ, Kim GH, et al. Clinical phenotype and mutation spectrum of the CYP21A2 gene in patients with steroid 21-hydroxylase deficiency. Exp Clin Endocrinol Diabetes 2012;120:23-7.

32. Burch GH, Gong Y, Liu W, Dettman RW, Curry CJ, Smith L, et al. Tenascin–X deficiency is associated with Ehlers–Danlos syndrome. Nat Genet 1997;17:104-8.

33. Mao JR, Taylor G, Dean WB, Wagner DR, Afzal V, Lotz JC, et al. Tenascin-X deficiency mimics Ehlers-Danlos syndrome in mice through alteration of collagen deposition. Nat Genet 2002;30:421-5.

34. Morissette R, Chen W, Perritt AF, Dreiling JL, Arai AE, Sachdev V, et al. Broadening the spectrum of Ehlers Danlos syndrome in patients with congenital adrenal hyperplasia. J Clin Endocrinol Metab 2015;100:E1143-52.

35. Merke DP, Chen W, Morissette R, Xu Z, Ryzin CV, Sachdev V, et al. Tenascin-X haploinsufficiency associated with Ehlers-Danlos syndrome in patients with congenital adrenal hyperplasia. J Clin Endocrinol Metab 2013;98:E379-87.

36. Lao Q, Brookner B, Merke DP. High-throughput screening for CYP21A1P-TNXA/TNXB chimeric genes responsible for Ehlers-Danlos syndrome in patients with congenital adrenal hyperplasia. J Mol Diagn 2019;21:924-31.

37. Gitelman SE, Bristow J, Miller WL. Mechanism and consequences of the duplication of the human C4/P450c21/gene X locus. Mol Cell Biol 1992;12:2124-34.

38. Marino R, Garrido NP, Ramirez P, Notaristéfano G, Moresco A, Touzon MS, et al. Ehlers-Danlos syndrome: Molecular and clinical characterization of TNXA/TNXB chimeras in congenital adrenal hyperplasia. J Clin Endocrinol Metab 2021;106:e2789-802.

39. Carrozza C, Foca L, Paolis ED, Concolino P. Genes and pseudogenes: complexity of the RCCX locus and disease. Front Endocrinol (Lausanne) 2021;12:709758.

40. Lao Q, Mallappa A, Rueda Faucz F, Joyal E, Veeraraghavan P, Chen W, et al. A TNXB splice donor site variant as a cause of hypermobility type Ehlers–Danlos syndrome in patients with congenital adrenal hyperplasia. Mol Genet Genomic Med 2021;9:e1556.

41. Lee HH, Lee YJ, Chao MC. Comparing the Southern blot method and polymerase chain reaction product analysis for chimeric RCCX detection in CYP21A2 deficiency. Anal Biochem 2010;399:293-8.

42. Olney RC, Mougey EB, Wang J, Shulman DI, Sylvester JE. Using real-time, quantitative PCR for rapid genotyping of the steroid 21-hydroxylase gene in a North Florida population. J Clin Endocrinol Metab 2002;87:735-41.

43. Ravichandran L, Varghese D, R P, S AH, Korula S, Thomas N, et al. Allele-specific and multiplex PCR based tools for cost-effective and comprehensive genetic testing in Congenital Adrenal Hyperplasia. MethodsX 2022;9:101748.

44. Baumgartner-Parzer S, Witsch-Baumgartner M, Hoeppner W. EMQN best practice guidelines for molecular genetic testing and reporting of 21-hydroxylase deficiency. Eur J Hum Genet 2020;28:1341-67.

45. Krone N, Arlt W. Genetics of congenital adrenal hyperplasia. Best Pract Res Clin Endocrinol Metab 2009;23:181-92.

46. Lee HH, Chang SF, Tsai FJ, Tsai LP, Lin CY. Mutation of IVS2-12A/C>G in combination with 707-714delGAGACTAC in the CYP21 gene is caused by deletion of the C4-CYP21 repeat module with steroid 21-hydroxylase deficiency. J Clin Endocrinol Metab 2003;88:2726-9.

47. Beighton P, De Paepe A, Steinmann B, Tsipouras P, Wenstrup RJ. Ehlers-Danlos syndromes: revised nosology, Villefranche, 1997. Ehlers-Danlos National Foundation (USA) and Ehlers-Danlos Support Group (UK). Am J Med Genet 1998;77:31-7.

48. De Coster W, Van Broeckhoven C. Newest methods for detecting structural variations. Trends Biotechnol 2019;37:973-82.

49. Logsdon GA, Vollger MR, Eichler EE. Long-read human genome sequencing and its applications. Nat Rev Genet 2020;21:597-614.

50. Liu Y, Chen M, Liu J, Mao A, Teng Y, Yan H, et al. Comprehensive analysis of congenital adrenal hyperplasia using long-read sequencing. Clin Chem 2022;68:927-39.

51. Tantirukdham N, Sahakitrungruang T, Chaisiwamongkol R, Pongpanich M, Srichomthong C, Assawapitaksakul A, et al. Long-read amplicon sequencing of the CYP21A2 in 48 Thai patients with steroid 21-hydroxylase deficiency. J Clin Endocrinol Metab 2022;107:1939-47.

52. Marino R, Ramirez P, Galeano J, Perez Garrido N, Rocco C, Ciaccio M, et al. Steroid 21-hydroxylase gene mutational spectrum in 454 Argentinean patients: genotype–phenotype correlation in a large cohort of patients with congenital adrenal hyperplasia. Clin Endocrinol (Oxf) 2011;75:427-35.

53. Finkielstain GP, Chen W, Mehta SP, Fujimura FK, Hanna RM, Van Ryzin C, et al. Comprehensive genetic analysis of 182 unrelated families with congenital adrenal hyperplasia due to 21-hydroxylase deficiency. J Clin Endocrinol Metab 2011;96:E161-72.

54. Krone N, Braun A, Roscher AA, Knorr D, Schwarz HP. Predicting phenotype in steroid 21-hydroxylase deficiency? Comprehensive genotyping in 155 unrelated, well defined patients from Southern Germany. J Clin Endocrinol Metab 2000;85:1059-65.

55. Wang R, Yu Y, Ye J, Han L, Qiu W, Zhang H, et al. 21-hydroxylase deficiency-induced congenital adrenal hyperplasia in 230 Chinese patients: Genotype-phenotype correlation and identification of nine novel mutations. Steroids 2016;108:47-55.

56. Koyama S, Toyoura T, Saisho S, Shimozawa K, Yata J. Genetic analysis of Japanese patients with 21-hydroxylase deficiency: identification of a patient with a new mutation of a homozygous deletion of adenine at codon 246 and patients without demonstrable mutations within the structural gene for CYP21. J Clin Endocrinol Metab 2002;87:2668-73.