

Application note

Targeted HiFi sequencing for congenital adrenal hyperplasia

Introduction

Congenital adrenal hyperplasia (CAH) refers to a group of rare inherited autosomal recessive disorders that affects the adrenal glands. CAH is characterized by an inability to produce typical levels of critical hormones, like cortisol, aldosterone, and androgens. CAH is most commonly due to a lack of the 21-hydroxylase enzyme, which is responsible for converting cholesterol to cortisol within the adrenal glands. Without this enzyme, the adrenal glands may produce too little cortisol and/or aldosterone and too much androgen, resulting in 21-hydroxylase deficiency (21OHD) (National Organization for Rare Disorders, 2018).

More than 1,300 genetic variants have been reported with about 230 affecting human health. More than 95% of the pathogenic variants causing 21OHD are caused by intergenic recombination. Approximately 75% of the deleterious variants are introduced by small gene conversions from the pseudogene during meiosis. In the remaining 20–25% of the cases, CAH is due to gross misalignment from unequal crossing over during meiosis, which can lead to gene deletions, duplications, and other mutations involving *CYP21A2* and other contiguous genes (Pignatelli et al., 2020).

In this Application Note, we focus on HiFi sequencing of full-length PCR amplicons spanning the 3.3 kb *CYP21A2* and its pseudogene, *CYP21A1P*. PacBio HiFi reads are long (up to 25 kb) and accurate (99.9%) making them well suited to analyze the complexity of CAH-related pathogenic variants, where gene conversions and chimeric genes are the predominant mutation types.

Workflow overview

Here, we present general instructions for co-amplification of full-length *CYP21A2* and *CYP21A1P* genes from human genomic DNA (gDNA) samples using barcoded primers in a single round of PCR, as shown in Figure 1. The recommended input gDNA per sample is 50 ng. Typical amplicon sizes for *CYP21A2* and *CYP21A1P* are 10.1 kb and 8.9 kb, respectively, enabling a better SV and breakpoint detection (Tantirukdham et al., 2022).

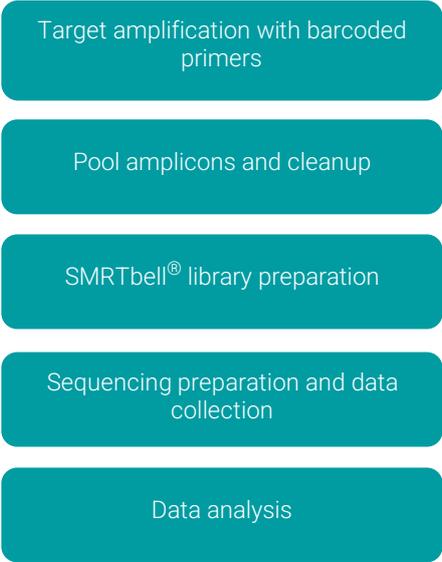


Figure 1. General steps of HiFi targeted sequencing

Multiplexing approach

Though this targeted sequencing approach can achieve a higher multiplexing level, this document provides primer designs for barcoding up to 24 samples. In order to increase the samples that can be multiplexed in a SMRT® Cell, additional primers may be designed with unique barcode sequences and validated with the workflow.

PCR

Table 1 lists sequences for barcoded Forward and Reverse primers for *CYP21A2* and *CYP21A1P* genes. The 8 barcoded forward primers are common for both *CYP21A2* and *CYP21A1P* genes. Three barcoded reverse primers are specific for *CYP21A2* and 3 are specific for *CYP21A1P*. The 8 common Forward primers and 3 specific reverse primers can be used in different combinations to barcode up to 24 samples. Both *CYP21A2* and *CYP21A1P* genes are co-amplified in a single PCR reaction (Figure 2). The CYP-LU2 forward primers and CYP-RD2 reverse primers are used at a final PCR reaction concentration of 0.20 µM. The CYP-LU2 forward primers and CYP-LD1 reverse primers are used at a final concentration of 0.050 µM.

Each oligo contains a 5' buffer sequence (GCATC), a 16-base barcode (in bold), and forward or reverse primer sequences.

SMRTbell library construction and sequencing

It is recommended to QC the PCR products for pooling prior to SMRTbell library construction. The instruction for constructing SMRTbell libraries is [found here](#). Finally, the SMRTbell library is prepared for sequencing on PacBio long-read systems.

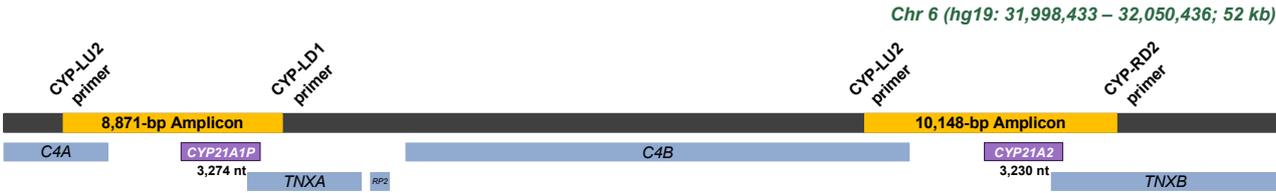


Figure 2. Amplicon positions for co-amplification of the gene and pseudogene. CYP-LU2 is a common forward primer.

Barcoded forward primer for CYP21A2 and CYP21A1P	
>CYP-LU2_BC1001_Forward	/5Phos/GCATCCACATATCAGAGT CGCG GCTTTGAGAGTTCTCTGAGTAGG
>CYP-LU2_BC1002_Forward	/5Phos/GCATC ACACAC AGACTGT GAGG GCTTTGAGAGTTCTCTGAGTAGG
>CYP-LU2_BC1003_Forward	/5Phos/GCATC ACACATCTCGT GAGAGGCTTTGAGAGTTCTCTGAGTAGG
>CYP-LU2_BC1004_Forward	/5Phos/GCATCC ACGCACACACGCG GCTTTGAGAGTTCTCTGAGTAGG
>CYP-LU2_BC1005_Forward	/5Phos/GCATCC ACTCGACTCTCGCGT GCTTTGAGAGTTCTCTGAGTAGG
>CYP-LU2_BC1006_Forward	/5Phos/GCATCCATATATATCAGCTGTGCTTTGAGAGTTCTCTGAGTAGG
>CYP-LU2_BC1007_Forward	/5Phos/GCATCTCTGTATCTCTATGTGGCTTTGAGAGTTCTCTGAGTAGG
>CYP-LU2_BC1008_Forward	/5Phos/GCATC ACAGT CGAGCGCT GCGG GCTTTGAGAGTTCTCTGAGTAGG
Barcoded reverse primers for CYP21A1P	
>CYP-LD1_BC1017_Reverse	/5Phos/GCATCATATAGCGCGCTGTGCTGAACGTGTTTTGGGACATG
>CYP-LD1_BC1018_Reverse	/5Phos/GCATCC ACAGT GAGC ACGTG ACTGAACGTGTTTTGGGACATG
>CYP-LD1_BC1019_Reverse	/5Phos/GCATCATCTGATAGAGTGTGTCTGAACGTGTTTTGGGACATG
Barcoded reverse primers from CYP21A2	
>CYP-RD2_BC1041_Reverse	/5Phos/GCATC GCGTGCACGCGG GAGAGAGTGAGCCTCTCACAGG
>CYP-RD2_BC1042_Reverse	/5Phos/GCATC GCGCTCGT CGAGCGAGGAGTGAGCCTCTCACAGG
>CYP-RD2_BC1043_Reverse	/5Phos/GCATCTATGTAGAGCTCTATAGAGTGAGCCTCTCACAGG

Table 1. Example of CYP21A2 and CYP21A1P primers for 24-plex of samples. The sequences in bold represent the barcodes.

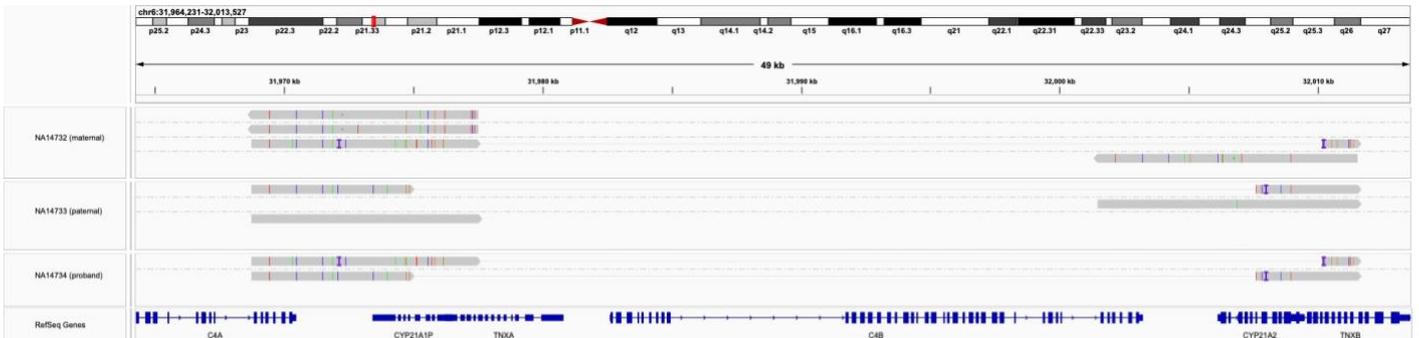


Figure 3. IGV view of a trio, tested with an amplicon based HiFi workflow. NA14732 (maternal) NA14733 (paternal) and NA14734 (proband). The mother has 2 typical pseudogene alleles, one typical CYP21A2, and one hybrid-fusion allele that is non-functional. Link lines between two parts of the fusion/hybrid consensus represent a single amplicon spanning both genes. The father has one each normal gene/pseudogene allele, and one hybrid fusion allele, but different from the mother’s fusion allele. The proband inherited both fusions, both of which are non-functional.

Recommended analysis tools

HiFi reads may be produced and demultiplexed on the sequencing instrument or in SMRT® Analysis off-instrument. HiFi reads for each sample can then be analyzed using pbaa, which generates phased consensus sequences by clustering HiFi reads and enables haplotype phasing for accurate diploidy calls using downstream community tools.

Some recommended tools that have been tested and optimized for HiFi reads are:

- Small variant calling: Google DeepVariant, bcftools, and GATK.
- Structural variant calling: pbsv, others.
- Gene/pseudogene rearrangement: PB-Motif (Stephens et al., 2021).

Resources and References

More details: please visit

<https://www.pacb.com/products-and-services/analytical-software/targeted-sequencing/>

Available materials

1. Targeted sequencing for amplicons – Best Practices <https://www.pacb.com/wp-content/uploads/Application-Brief-Targeted-sequencing-Best-Practices.pdf>
2. Preparing multiplexed amplicon libraries using SMRTbell prep kit 3.0 <https://www.pacb.com/wp-content/uploads/Procedure-checklist-Preparing-multiplexed-amplicon-libraries-using-SMRTbell-prep-kit-3.0.pdf>

More technical info: <https://www.pacb.com/wp-content/uploads/McLaughlin-ESHG-2021-Targeting-Clinically-Significant-Dark-Regions-of-the-Human-Genome-with-High-Accuracy-Long-Read-Sequencing.pdf>

Customer workflow example

In Tantirukdham et al., Dr. Shotelersuk and team explored 48 putative cases of CAH caused by 21-OHD to determine whether PacBio long-read technology could be leveraged to detect pathogenic variants and optimize clinical workflows.

Dr. Shotelersuk considered the possibility of simplifying his team’s workflow to get more data using a long-read sequencing approach. The team implemented a new workflow with two key goals in mind:

- Detect pathogenic variants in the *CYP21A2* gene
- Determine the workflow’s effectiveness in positive 21-OHD cases

Their 100% concordance rate in cases of confirmed affected subjects highlights the opportunity to consolidate conventional molecular techniques without compromising performance.

See our [gene testing blog post](#) to read more about this study.

References

- Liu, Y et al., (2022) [Comprehensive Analysis of Congenital Adrenal Hyperplasia Using Long-Read Sequencing](#). *Clinical Chemistry*, 68(7), 927–939.
- Congenital adrenal hyperplasia*. NORD (National Organization for Rare Disorders). (2018, June 26). Retrieved September 21, 2022, from <https://rarediseases.org/rare-diseases/congenital-adrenal-hyperplasia/>
- Pignatelli, D., et al. (2019) [The Complexities in Genotyping of Congenital Adrenal Hyperplasia: 21-Hydroxylase Deficiency](#). *Frontiers in Endocrinology*, 10, 432.
- Stephens, Z., et al. (2021) [PB-Motif-A Method for Identifying Gene/Pseudogene Rearrangements With Long Reads: An Application to CYP21A2 Genotyping](#). *Frontiers in Genetics*, 1359.
- Tantirukdham, N., et al. (2022) [Long-read amplicon sequencing of the CYP21A2 in 48 Thai patients with steroid 21-hydroxylase deficiency](#). *The Journal of Clinical Endocrinology & Metabolism*.

Research use only. Not for use in diagnostic procedures. © 2022 Pacific Biosciences of California, Inc. (“PacBio”). All rights reserved. Information in this document is subject to change without notice. PacBio assumes no responsibility for any errors or omissions in this document. Certain notices, terms, conditions and/or use restrictions may pertain to your use of PacBio products and/or third-party products. Refer to the applicable PacBio terms and conditions of sale and to the applicable license terms at [pacb.com/license](https://www.pacb.com/license). Pacific Biosciences, the PacBio logo, PacBio, Circulomics, Omniome, SMRT, SMRTbell, Iso-Seq, Sequel, Nanobind, and SBB are trademarks of PacBio. All other trademarks are the sole property of their respective owners.