Abstract citation ID: bvaf149.383

Adrenal (Excluding Mineralocorticoids)

SUN-413

Generic Heterogeneity of the RCCX Module Copy Number Variations Using Long-Read Sequencing in the Human Pangenome Data

Jun-Hong Park, MD[1], *Ja Hye Kim, MD and PhD*[1], *Naeun Kwak*[1], *Soyoung Park*[1], *Ji-Hee Yoon*[2], *Seo-Yeon Kim*[3], *Dohyung Kim*[1], *Soojin Hwang*[1], *Gu-Hwan Kim*[1], *Jun Kim*[4], *Han-Wook Yoo, MD,PhD*[5], *and Jin-Ho Choi, MD,PhD*[6]

[1]ASAN MEDICAL CENTER, Seoul, Korea, Republic of; [2]Kangbuk Samsung Hospital, Seoul, Korea, Republic of; [3]Chungnam National University, Deajeon, Korea, Republic of; [4]Chungnam National University, Daejeon, Korea, Republic of; [5]CHA BUNDANG MEDICAL CENTER, Seongnam-si, Gyeonggi-do, Korea, Republic of; [6]Asan Medical Center, Seoul, Korea, Republic of.

Disclosure: J. Park: None. **J. Kim:** None. **N. Kwak:** None. **S. Park:** None. **J. Yoon:** None. **S. Kim:** None. **D. Kim:** None. **S. Hwang:** None. **G. Kim:** None. **J. Kim:** None. **H. Yoo:** None. **J. Choi:** None.

Purpose: Germline copy number variation (CNV) is defined as a DNA segment with varying copy numbers resulting from duplications or deletions during evolution. The RCCX module represents one of the most complex CNV loci in humans, and its genetic variability poses challenges in accurately identifying molecular defects in the *CYP21A2* gene. This study aimed to investigate the genetic heterogeneity within the RCCX module using data from the Human Pangenome Reference Consortium (HPRC), which includes whole-genome sequencing data from 47 individuals using long-read sequencing (LRS) platform. Additionally, we sought to establish an enhanced reference framework to facilitate the comprehensive genetic analysis of patients with congenital adrenal hyperplasia. **Methods:** To identify gene structures within the RCCX locus in the HPRC, known gene sequences from the GRCh38 reference genome were used as queries in a BLAST analysis against the genome assemblies of 94 alleles. This analysis determined the positions and sequences of these genes within each assembly, revealing diverse structural forms, including mono-, bi-, tri-, *CYP21* genes, as well as *TNX* genes. The sequences of the *CYP21* genes were then compared to the known sequences from the GRCh38 reference genome using graph-based pangenome-level variant calling to identify specific variants. Active and pseudogene-specific markers (58 for *CYP21* genes and 11 for *TNX* genes) were used to classify the genes as active, pseudogene, or fusion. The presence of all markers indicated *CYP21A1P*, partial presence indicated a *CYP21* fusion, and complete absence indicated *CYP21A2*. **Results:** The analysis identified ten different allele types within the RCCX region. The most

prevalent configuration, observed in approximately 80% (75/94) of the samples, was bimodular, with 71% (67/94) containing one copy each of *CYP21A2* and *CYP21A1P*. In the bimodule configuration, 70% (70/94) of individuals had one copy each of *TNXA* and *TNXB*, while four individuals exhibited a *TNXA-TNXB* fusion alongside a *TNXB* gene. One individual had two copies of *TNXB* without *TNXA*. Monomodular alleles, accounting for 14% (13/94), consistently featured the active *CYP21A2* and *TNXB* genes. Trimodular alleles constituted 6% (6/94) with one allele containing two copies of *CYP21A1P* and one copy of *CYP21A2*, another having one copy of *CYP21A1P* and two copies of *CYP21A2,* and the remaining allele comprising two copies of *CYP21A1P* and one *CYP21A1P*/*CYP21A2* fusion gene. Regarding *TNX* genes, three individuals had two copies of *TNXA* and one copy of *TNXB*, while another three individuals had one copy of *TNXA*, a *TNXA-TNXB* fusion, and one copy of *TNXB*. **Conclusions:** Using the LRS and genome assembly analysis, we were able to identify not only the number of CNVs but also the specific genotypes. We demonstrated the extensive allelic diversity present in the general population.

Presentation: Sunday, July 13, 2025