Integrating Copy Number Value (CNV) Analysis into High-Resolution HLA Typing for Improved Allele Assignment: From Ambiguity to Accuracy

Gaurav Tripathi[1], Farzan Modaresi[1], Russ Martin[1], Zahra Kashi[1]
[1]Kashi Clinical Laboratories

AIM

Copy Number Value (CNV) refers to the quantitative estimate of how many copies of a specific HLA gene (or allele) are present in a sample, typically inferred from sequencing read depth in NGS-based typing. In a normal diploid genome, the expected copy number is two (one per chromosome). CNV imbalance may indicate the presence of cryptic alleles, gene duplication events, or technical artifacts (e.g., PCR amplification bias or allele dropout) (Figure 1-2). This consideration is particularly relevant for HLA-DRB loci, where variable gene content and CNV can contribute to allelic imbalance, interpretive ambiguity, and potential misassignment. Such an imbalance was observed in a sample from a 28-year-old O+, Hispanic male NMDP donor, for whom an EDTA blood specimen was submitted for confirmatory HLA typing.

METHOD

DNA was extracted from the blood sample using an automated extractor per the manufacturer's protocol. After concentration and quality assessment, libraries were prepared and sequenced on the MiSeq i100 for high-resolution HLA typing.

RESULTS

The FASTQ file was uploaded to ScisCloud for HLA typing. The sample passed QC but flagged a mismatch in Class II HLA-DRB1, assigning DRB1*03:01 and DRB1*14:06. Review of the flagged exon-2 region showed that DRB1*14:06 was a misassignment due to extra targets (Figure 3).

Analysis in TypeStream Visual showed consistent Class I and II typing without DRB1 flags (Figure 4); however, DRB1 read depth was observed unusually high (>38,000). CNV analysis revealed a ratio of ~1:4 instead of the expected 1:1 or 1:2, indicating allelic imbalance and likely misclassification (Figure 5).

This pattern may suggest the possible presence of DR1, DR8, or DR10, despite the software assigning DRB1*03:01 and DRB1*14:06 based on high depth-read. Once the sample was re-analyzed the high read depth DRB1*14:06 was removed and DRB1*08:02 was called (Figure 6-7).

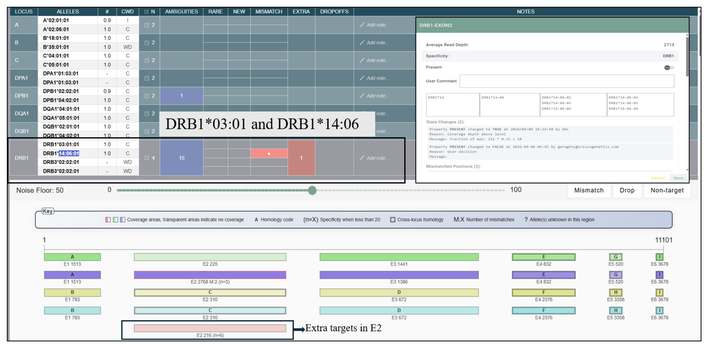


Figure 3: Analysis of sample for high resolution HLA Typing: Figure representing ScisCloud analysis of sample for high resolution HLA Typing results where DRB1*03:01 and DRB1*14:06 was assigned by software with an error saying Extra targets in Exon2 of DRB1

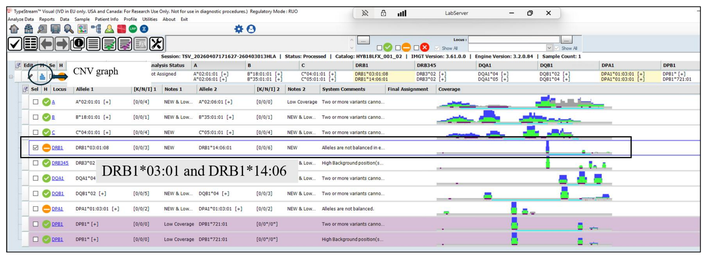


Figure 4: Analysis of sample for NGS based high resolution HLA Typing: Figure representing TypeStream Visual (TSV) analysis of sample for high resolution HLA Typing results where DRB1*03:01 and DRB1*14:06 was assigned by TSV

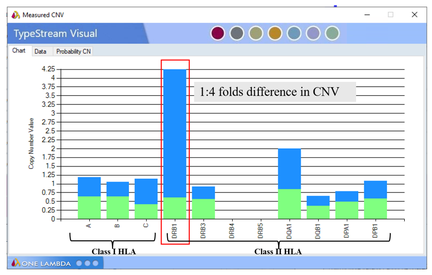


Figure 5: Copy Number Value difference in class I and class II HLA Typing where HLA DRB1 showed 1:4 fold difference in CNV indicator of allele misclassification

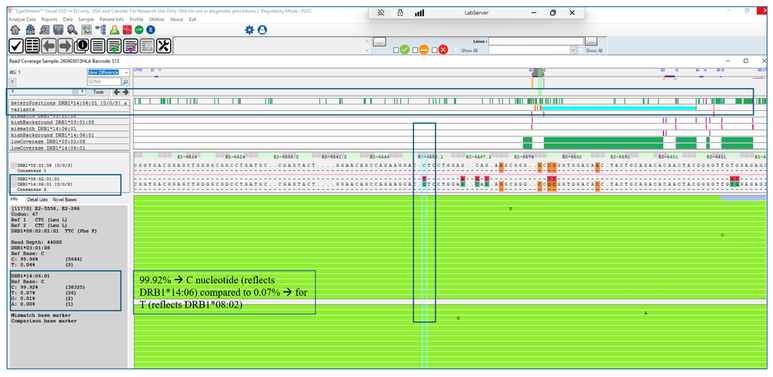


Figure 6: Allele comparator analysis: NGS HybriType TSV analysis ruled out DRB1*08:02:01:01 as many differences reflected in the allele comparator analysis, however, CNV analysis (figure 3) reflected errors in software analysis for extra attention in allele assignment.

WHAT IS COPY NUMBER VALUE IN HLA ANALYSIS?

- An inferred estimate of how many copies of each allele are present per locus – *not directly measured, it's calculated* (Figure 1)
- Calculated CNV is compared to the biologically expected number of gene copies for each locus (Figure 1)

How does it work?

TSV performs copy number detection for sequenced genes and reports measured and calculated copy numbers for each allele along with a confidence value

- Analyzes read counts by comparing within and across samples
- Heavily uses read counts and allele balance

What information does it provide?

- ✓ Homozygous vs Heterozygous
- ✓ Homozygous vs Hemizygous
- ✓ Nullizygous results
- ✓ Additional copies of genes – *very rare*

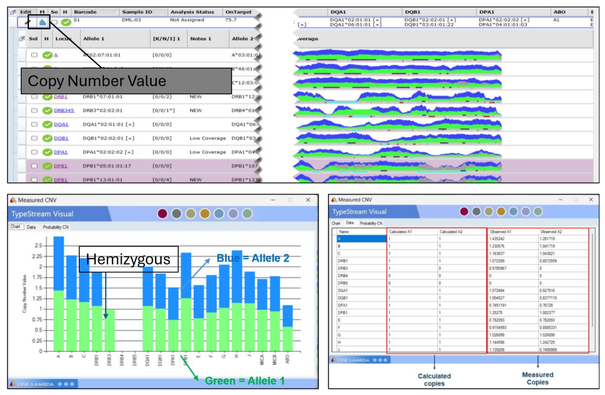


Figure 1: The calculated whole-number copies for allele 1 and allele 2 for each locus and the measured number of copies of allele 1 and allele 2 are used for the implication of calculated whole-number copies for each allele. The measured copies values are used to assess the probability

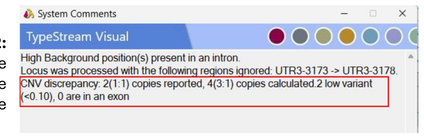


Figure 2: Copy Number Value Discrepancy message on TSV software

Copy Number Value Discrepancy

- Occurs when probable CNV combination is different than the known biological CNV
- One allele may be over-represented
- Useful for LOH cases
- Imbalance in read coverage can trigger the discrepancy

CONCLUSIONS

The mismatches and high CNV allowed to determine that the DRB1*14:06 exon 2 was an extraneous amplicon and removing it from analysis allowed determination of the correct typing (the hidden DRB1*08:02) (Figure 7). This observation signifies the importance of CNV analysis in the HLA typing. While less common in Class I, CNVs are critical in Class II, especially HLA-DR and must be resolved to prevent misclassification, ambiguity, and potential clinical errors in high-resolution HLA typing.

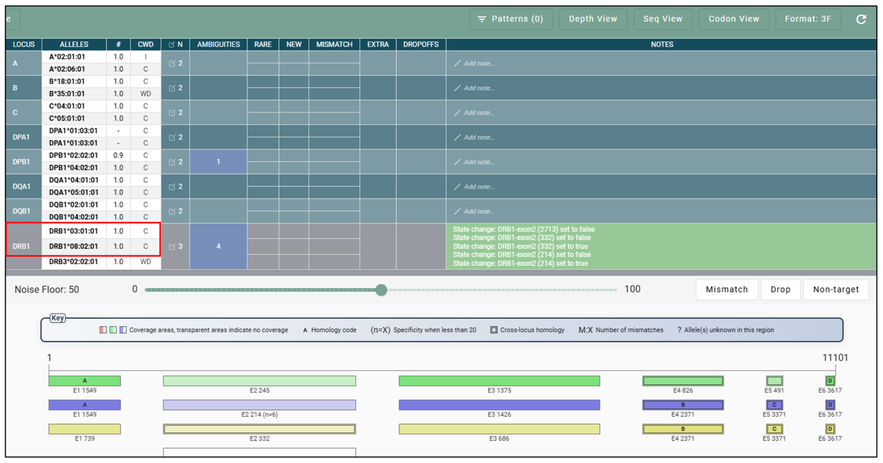


Figure 7: After re-analysis on Scisco Genetics platform: since there was a message 'Extra targets in E2' (shown in figure 1), it was reanalyzed and Sciscloud software resulted out in correct DRB1 alleles (DRB1*03:01 and DRB1*08:02)

KASHI CLINICAL LABORATORIES

CONTACT
(503) 206-4989 | info@kashilab.com
www.KashiLab.com
10101 SW Barbur Blvd. Portland, OR 97219