IDENTIFICATION OF A NOVEL DQB1*03 ALLELE IN PSEUDO EXON 5 BY NEXT GENERATION SEQUENCING HLA TYPING

Memorial Sloan Kettering Cancer Center

Nancy Yeh, Peter McManus, Byoungkyu Kim, Eunsil Kim, Christina Savchik, Jonathan Barone, Amanda G. Blouin, Fei Ye

Department of Pathology and Laboratory Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, United States

BACKGROUND:

Next-Generation Sequencing (NGS) has become a major technology in HLA typing and provides high-throughput, high-resolution and high-accuracy genotypes at 11 classical HLA genes. Currently, two main techniques targeting HLA region include long-range PCR NGS (LR-NGS) and hybridization-capture NGS (HC-NGS). One of the advantages of NGS is the ability to discover novel sequences through routine testing. However, LR-NGS may miss a novel allele if the variant is within or close to primer binding region, but HC-NGS is less likely to miss variants due to lower stringency of capture probes and multiple probe panel to overcome any potential variant.. In addition, due to the nature of HLA genes existing multigene families, which consist of a cluster of closed related genes located at the same chromosome (6p21) with multiple variants for each gene, the significantly similar sequences of other HLA genes or pseudogenes (non-functional gene copy) can lead to inaccurate allele identification during analysis.

In our lab, we perform HLA typing using both NGS techniques and examine their performance to determine an effective workflow incorporating both to address limitations of each assay.

CASE:

Here we describe the identification of a novel allele DQB1*03:01 carrying one nucleotide substitution in pseudo exon 5 by NGS, discovered in a healthy stem cell donor.

Briefly, total genomic DNA was extracted from the peripheral blood of a donor, which was recruited by National Bone Marrow Donor registry (NMDP), using EZ1 advanced XL (Qiagen). HLA typing was performed on the DNA sample using HC-NGS (AlloSeq Tx, CareDx) and analyzed by Assign v1.0.3, IMGT db v3.49. The donor was found to have a novel allele DQB1*03:01 with one nucleotide substitution in codon 232 at pseudo exon 5, changed from CCA to CGA (**Figure 1**). For majority of characterized DQB1 alleles including DQB1*03:01, a single base substitution from G>A at 3' acceptor splice site in intron 4 leads to elimination of exon 5 and no protein expression at exon 5 **(Figure 2**). Therefore, this variant does not result in a protein change.

CONTACTS:

yef@mskcc.org; blouina@mskcc.org

We performed additional testing on the DNA sample using LR-NGS (MiaFora MFlex, Werfen) and analyzed by MIA FORA v5.2, IMGT db v3.50. This variant was not identified by MFlex covering of exons 1 through 5 of DQB1, but an imbalanced SNP at this position was observed (**Figure 3**), which may be due to either imbalanced amplification or inability of MiaFora software to identify this heterozygous position. To further confirm this novel allele, this DNA sample was sent to reference laboratories and was confirmed by AllType One Lambda along with TypeStream vTSV 3.0, IMGT db 3.53.0 (LR-NGS with coverage extending through exon 6 of DQB1) and by MFlex expanded (XP) kit (with new expanded coverage on DQB1 exon 1-6).

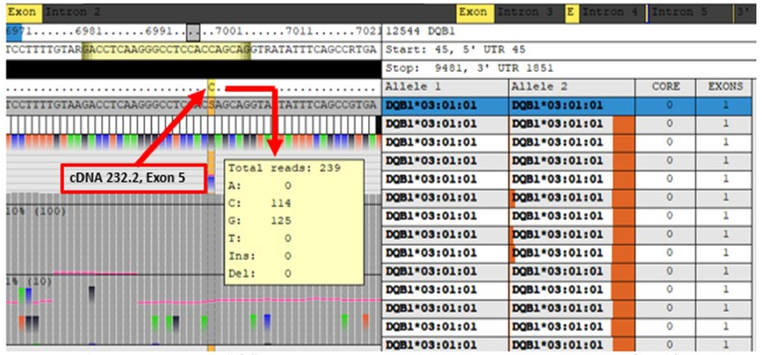

Figure 1. The variant "G" at the position 232.2 at exon 5 on DQB1*03:01 was identified by AlloSeq Tx.

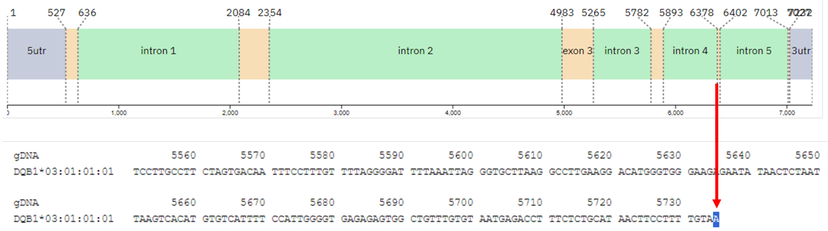

Figure 2. No protein change due to a nucleotide substitution G>A at exon 5, 3' splice site present in DQB1*03:01

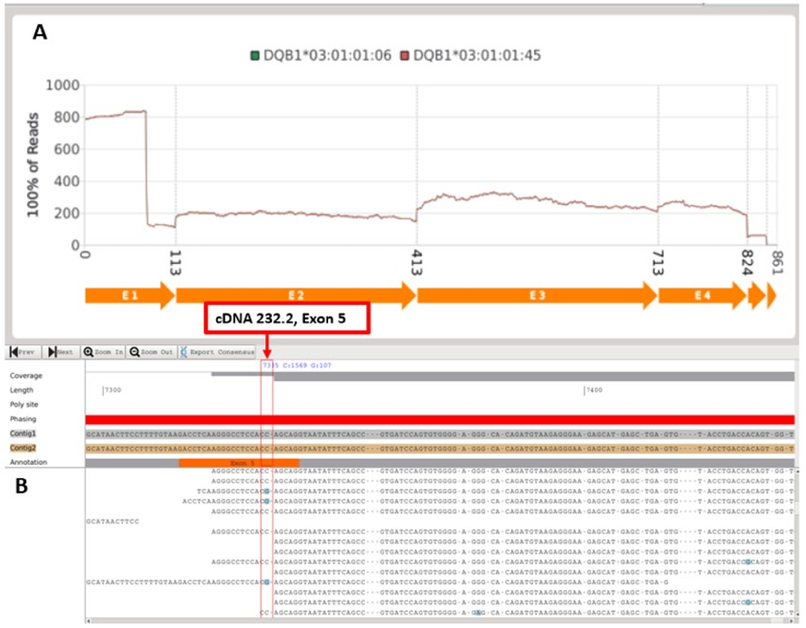

Figure 3. The variant "G" at the position 232.2 at exon 5 on DQB1*03:01 was not identified by MiaFora MFlex. **A**. Coverage plot in all exons; **B**. Sequencing alignment in exon 5. The variant "G" at the position 232.2 at exon 5 only had 107x reads, but "C" at same position had 1569x reads.

CONCLUSIONS:

The novel allele DQB1*03:01 at pseudo exon 5 was identified by AlloSeq Tx but missed by MiaFora MFlex due to either amplification bias or software limitation in distinguishing heterozygous position from sequencing noise. HC-NGS was more robust to identify novel allele than the LR-NGS with less effective coverage possibly due to proximity of variant to primer binding site. The complementary advantages and limitations of the two methods (HC-NGS and LR-NGS) may allow laboratories to recognize novel alleles more effectively in HLA typing.

REFERENCES:

1. Ralazamahaleo M, Elsermans V, Top I, Guidicelli G, Visentin J. Characterization of the novel HLA-DQB1*03:02:01:08 allele by sequencing-based typing. HLA. 2019 Sep;94(3):335-336. PMID: 31081257
2. Tran JN, Sherwood KR, Mostafa A, Benedicto RV, ElaAlim A, Greenshields A, Keown P, Liwski R, Lan JH. Novel alleles in the era of next-generation sequencing-based HLA typing calls for standardization and policy. Front Genet. 2023 Oct 13;14:1282834.. PMID: 37900182