HLA-DRB4*03:01N: absent or well hidden? Genotyping an allele lacking exon 2

Grit Wagner[1], Christin Paech[1], Jürgen Sauter[2], Annett Heidl[1], Bianca Heyn[1], Vinzenz Lange[1], Anja Klussmeier[1]

[1] DKMS Life Science Lab, Dresden, Germany
[2] DKMS Group, Tübingen, Germany

Introduction

The allele HLA-DRB4*03:01N was first described in 1997 and is a challenge for PCR-based genotyping. Its HLA-DRB1*07-linked haplotype is characterized by a 16 kb long deletion that includes HLA-DRB4 exons 1 and 2 as well as primer binding regions commonly used for genotyping (Steiner et al., 2018). Therefore, HLA-DRB4*03:01N might fail to amplify and remains undetected.

Our high-throughput workflow for potential stem cell donors genotypes exons 2 and 3 of HLA-DRB4 using two independent PCR amplicons. Since 2020, we identified more than 100 samples meeting the criteria for a potential presence of HLA-DRB4*03:01N:

- presence of at least one HLA-DRB1*07 allele
- amplification failure of exon 2
- sequence of exon 3 concordant with HLA-DRB4*03:01N

However, we were still uncertain about reporting HLA-DRB4*03:01N based on these criteria alone and aimed for confirmation.

PCR & Sequencing

76 samples (DNA isolated from buccal swabs) were subjected to two separate PCR reactions that differed only in the used primer pairs.

HLA-DRB4*03:01N-specific primers:

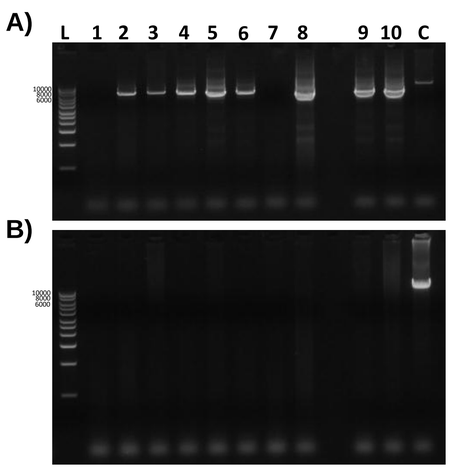
- CCAATCACTGAGATGACAATGAGTACTGCC (DRB4_31N_NA_F1)
- AGTAACAACCTGGTCCTGACAAAGC (DRB4_NA_R1)

HLA-DRB4 full-length primers:

- TGCTAGTGAACTCAGGTGCTGAC (DRB4_NA_F2)
- TGCTAGTGAACTCAGATGCTGAC (DRB4_NA_F3)
- AGTAACAACCTGGTCCTGACAAAGC (DRB4_NA_R1)

All primers were tagged with 5´ overhangs for later sample barcoding and ONT sequencing as described by Putke et al..

4 µL DNA (range 2-100 ng/µL) were mixed with primers (200 nM each), dNTPs (0.4 mM each), 2.5 µg BSA, 10x Advantage Genomic LA buffer, and 0.25 U Advantage Genomic LA Polymerase Mix (Takara Bio, Mountain View, California) in a 25 µL PCR reaction. The following PCR conditions were used: 94°C 1 min, 35 cycles: 98°C 10 sec / 67°C 13 min, 72°C 10 min. PCR products were checked by agarose gel electrophoresis.



PCR and agarose gel electrophoresis of 10 samples (lanes 1-10), in which the presence of HLA-DRB4*03:01N was suspected and no other HLA-DRB4 allele was present. **A) HLA-DRB4*03:01N-specific primers. B) HLA-DRB4 full-length primers**. Lane 1 and 7: suspected PCR failures, L: DNA ladder (500-10,000 bp), C: positive control for HLA-DRB4 full-length primers (mixture of high-quality genomic DNA).

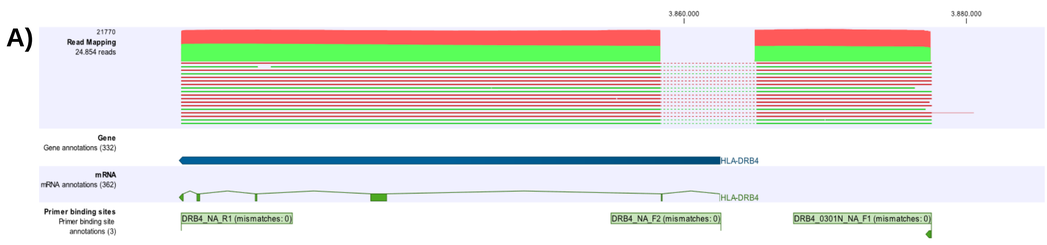
HLA-DRB4*03:01N	other HLA-DRB4	primers	expected amplicon size
-	+	HLA-DRB4*03:01N-specific	none (or 23 kb)
		HLA-DRB4 full-length	15 kb
+	-	HLA-DRB4*03:01N-specific	7 kb
		HLA-DRB4 full-length	none
+	+	HLA-DRB4*03:01N-specific	7 kb (+ 23 kb)
		HLA-DRB4 full-length	15 kb
-	-	HLA-DRB4*03:01N-specific	none
		HLA-DRB4 full-length	none

HLA-DRB4*03:01N-specific primers generate a 7 kb amplicon in presence of an HLA-DRB4*03:01N allele. Additionally, a 23 kb amplicon can be generated with all other HLA-DRB4 alleles. This was not expected with our buccal swab samples, but might occur if DNA of higher quality is used.
HLA-DRB4 full-length primers produce a 15 kb amplicon for standard HLA-DRB4 alleles. Since the binding site for the forward primers is missing, no amplicon is expected for HLA-DRB4*03:01N.

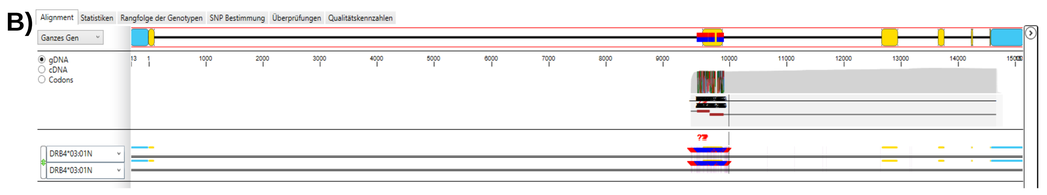
After index PCR (see Putke et al. for sequences and protocol), all generated amplicons were sequenced on a MinION 10.4 flowcell (Oxford Nanopore Technologies, Oxford, GB).

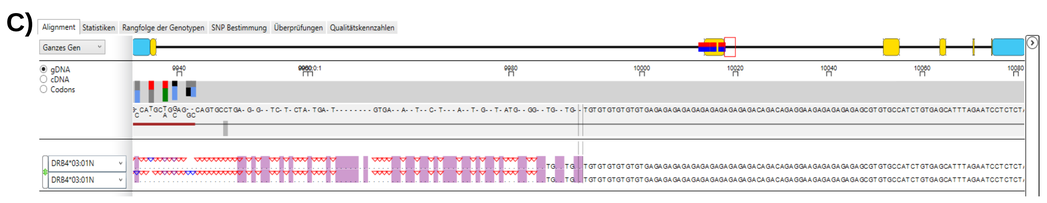
Analysis

Sequencing reads from the HLA-DRB4*03:01N-specific PCR were analyzed with CLC Genomics Workbench 24 (Qiagen, Hilden, Germany) (A) and NGSengine (GenDx, Utrecht, The Netherlands) (B and C).



ONT sequencing reads were mapped against the NT_167249_SSTO reference allowing read splicing (Map Long Read to Reference, Automatic Spliced). This read splicing (dashed region in read mapping) highlights the 16 kb deletion region, that was first described by Steiner et. al.. Forward and reverse reads are depicted in red and green. Locations of the HLA-DRB4 gene (blue arrow), exons/mRNA (green) and primer binding sites (primer DRB4_NA_F3 binds to the same region as DRB4_NA_F2 and is not shown) are depicted as additional tracks to the read mapping.





In NGSengine, sequencing reads were mapped against HLA-DRB4. HLA-DRB4*03:01N needs to be chosen manually as reference allele (two HLA-DRB4*03:01N alleles are shown because NGSengine cannot work with only one allele). Obviously, NGSengine cannot map the sequencing reads to the non-existent exon 2 correctly (colorful region with question marks and triangles). Zooming into the breakpoint region right behind exon 2 (C: red box in gene organization overview) shows a region rich in GT and GA repeats. This supports the proposed breakpoint from Steiner et al..

Out of the 76 tested samples, 58 were confirmed to be HLA-DRB4*03:01N without any mismatches. 12 samples failed in both long-range PCRs (HLA-DRB4*03:01N-specific and HLA-DRB4 full-length) or did not generate enough sequencing reads for analysis. The remaining six samples had very low read counts for exon 2 in the initial genotyping (as opposed to the other 70 samples that had no reads in exon 2) and were included to challenge the HLA-DRB4*03:01N identification. As expected, they only yielded PCR products in the HLA-DRB4 full-length PCR, but not in the HLA-DRB4*03:01N-specific PCR. Sequencing revealed that four of these six samples had mismatches in the exon 2 primer sites of our high-throughput workflow. These novel alleles will be submitted to IPD-IMGT/HLA. For the remaining two samples the reason for an extremely weak exon 2 amplification in the high-throughput workflow remains unclear.

Conclusion

We successfully implemented a protocol for the full-length identification of HLA-DRB4*03:01N. Furthermore, we demonstrate that all samples that amplified successfully and fulfilled certain criteria after genotyping of HLA-DRB4 exons 2 and 3 (HLA-DRB1*07 present, no amplification of exon 2, exon 3 sequence concordant with HLA-DRB4*03:01N) were confirmed to be HLA-DRB4*03:01N. Consequently, we feel confident that in these cases HLA-DRB4*03:01N could be reported with high accuracy even without running an additional confirmation PCR. However, cases with very low amplification of exon 2 have been found to be caused by other reasons, including mutations in the primer binding regions.

Considering both DRB4-positive and -negative haplotypes, we determined an allele frequency of 0.0014% for HLA-DRB4*03:01N (46 cases) in the German population based on 1.67 million samples.

References

- Steiner NK, Hou L, Hurley CK. Characterizing Alleles with Large Deletions Using Region Specific Extraction. Hum Immunol. 2018 Mar 23; doi: 10.1016/j.humimm.2018.03.005
- Putke K, Albrecht V, Paech C, Pahlke M, Schöne B, Klasberg S, et al. Full-Length Characterization of Novel HLA-DRB1 Alleles for Reference Database Submission. In: Boegel S, editor. HLA Typing: Methods and Protocols. New York, NY: Springer US; 2024. p. 145–56. Available from: https://doi.org/10.1007/978-1-0716-3874-3_10

DKMS WE DELETE BLOOD CANCER

DKMS Life Science Lab
St. Petersburger Str. 2
01069 Dresden, Germany
www.dkms-lab.de