Diversity of the non-classical MHC class II genes HLA-DMA, -DMB, -DOA, and -DOB

Madlen Pahlke[1], Viviane Albrecht[1], Christin Paech[1], Kathrin Putke[1], Alexander H. Schmidt[1,2], Vinzenz Lange[1], Anja Klussmeier[1]

[1]DKMS Life Science Lab, St. Petersburger Str. 2, 01069 Dresden, Germany

[2]DKMS, Kressbach 1, 72072 Tübingen, Germany

Introduction

HLA-DM and HLA-DO are non-classical MHC class II genes and modulate the peptide loading to the classical MHC class II proteins inside endosomes. While HLA-DM catalyzes the peptide-exchange in favor to a highly stable MHC class II/peptide-complex, HLA-DO acts as its natural antagonist. However, only few alleles of the HLA-DM/DO genes HLA-DMA, HLA-DMB, HLA-DOA and HLA-DOB are known and data about allelic frequencies or functional differences are sparse. Therefore, the aim of this study was to provide an overview about their allelic diversity using nanopore sequencing on 1880 samples from potential stem cell donors registered with DKMS Germany, DKMS UK, DKMS Poland and DKMS Chile.

Diversity

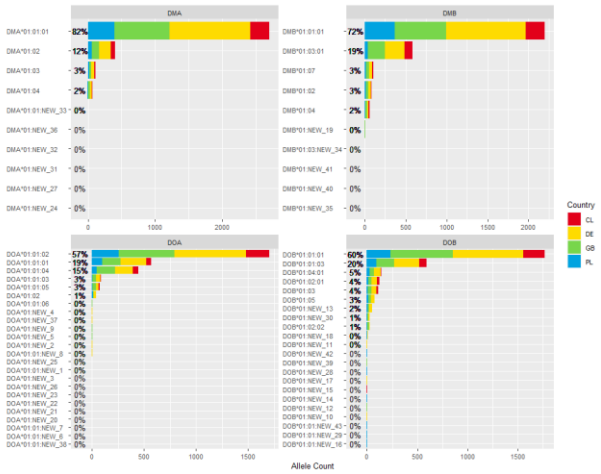


Figure 1: Diversity of HLA-DMA, -DMB, -DOA and -DOB in three-field resolution. Total allele counts are plotted on the x-axis. The corresponding allele frequencies (rounded) are next to the bars. Novel sequences, which are not yet officially named by IPD-IMGT/HLA, carry random numbers. Colors indicate the country, in which the sample has been collected.

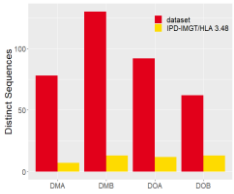


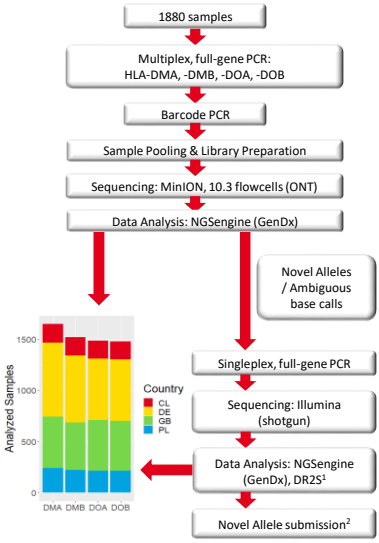
Figure 2: Total number of distinct sequences in the samples and in the current IPD-IMGT/HLA release (3.48) in four-field resolution.

Some previously described proteins were not detected in our samples:

- DMB*01:05, DMB*01:06
- DOA*01:03, DOA*01:04N

One explanation might lie in the ethnicity: While our sample set was restricted to Europe and Chile, the missed alleles had all been initially reported from oriental ethnicities.

Method



Submissions to IPD-IMGT/HLA

Almost all sequences with novel exon bases and the most frequent sequences with novel intron bases were submitted to IPD-IMGT/HLA. Furthermore, sequences were selected to extend partially described alleles in the database. If available, sequences from two independent samples were submitted as confirmation.

Overall, we have identified, analyzed and characterized 31 novel proteins. This will double the number of known proteins (currently 4, 7, 3 and 5 for HLA-DMA-, DMB, -DOA and -DOB; release 3.48).

	Total Submissions	Novel Protein Submissions
HLA-DMA	75	5
HLA-DMB	109	2
HLA-DOA	122	12
HLA-DOB	83	12

Table 1: Submissions to IPD-IMGT/HLA. Total submissions include sequence confirmations and extensions of known alleles.

Conclusion

We analyzed the diversity of HLA-DMA, -DMB, -DOA and -DOB genes in 1880 samples, which were collected in Germany, Poland, UK and Chile. In all four genes, we identified one major protein (frequency > 70%) with only marginal differences between the four countries. However, we detected more than 250 distinct and currently undescribed variations, including 31, which represent novel proteins. These high numbers suggest the existence of even more variations in the population, even though they seem to be rare.

References

[1] Klasberg S, Schmidt AH, Lange V, Schöfl G. DR2S: an integrated algorithm providing reference-grade haplotype sequences from heterozygous samples. BMC Bioinformatics. 2021 May 10;22(1):236.

[2] Schöne B, Fuhrmann M, Surendranath V, Schmidt AH, Lange V, Schöfl G. TypeLoader2: Automated submission of novel HLA and killer-cell immunoglobulin-like receptor alleles in full length. HLA. 2019;93(4):195-202. doi:10.1111/tan.13508

DKMS Life Science Lab

GitHub *Download from GitHub:* https://github.com/DKMS-LSL

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