GENDX

personalizing diagnostics

A novel NGS-based high-resolution genotyping strategy for HLA-E

HLA typing in stem cell transplantation diagnostics requires high-resolution typing of a combination of Class I and Class II HLA genes of donor and recipient. There is evidence that the genotype of the non-classical HLA-E gene may also be clinically significant and may impact transplantation outcome (Tsamadou et al, Haematologica, 2017, 102(11):1947-1955). GenDx has developed an NGS-based typing strategy for whole-gene HLA-E typing at high-to-allelic level resolution, supporting sequencing of all exons and introns.

Materials & Methods

IHWG gDNA samples (International Histocompatibility Working Group, Fred Hutchison, USA) and UCLA gDNA samples (UCLA immunogenetics Center, USA) were PCR-amplified using NGSgo®-AmpX HLA-E primers. Amplicons were verified by gel electrophoresis and processed in the NGSgo® library preparation procedure (GenDx). Libraries were paired-end sequenced (2x151bp) on a MiSeq platform (Illumina). Sequences were analyzed in NGSengine® HLA typing software (GenDx), using IMGT/HLA database 3.31.0.

Results

The performance of the assay was evaluated using a panel of IHWG and UCLA gDNA samples, including nine samples with HLA-E pre-type information as reference samples. The amplification of HLA-E was robust, with successful amplification of all samples (Figure 1). Analysis of the reference samples showed that the typing results obtained in NGSengine® was fully concordant with the pre-type (Table 1).

The majority of the HLA-E alleles belonged to the E*01:01 or E*01:03 group (Table 2). These are known to be the predominant allele groups for HLA-E. The two most common alleles observed are E*01:01:01:01/02 (45.0%) and E*01:03:02:01 (26.1%). For two alleles (E*01:01:01:01 and E*01:01:01:02), the fourth field could not be unambiguously typed due to the fact that the discriminating SNP was located outside the amplicon in the 5'UTR. All other alleles could be distinguished at the fourth field, a.k.a. allele level resolution, in NGSengine.

The HLA-E sequence data was of overall high quality, with 1) full coverage of the whole gene, 2) an allele ratio of ~50% as expected for heterozygous samples and 3) low noise levels (<3.5%) (Figure 2). The data was characterized by a low density of polymorphic positions, typical for HLA-E. Phasing of the core exons 2 and 3 was accomplished (indicated with the red line in Figure 2), but phasing over longer distances was not possible. For example, see sample IHWG9270 in Figure 2. This sample had two heterozygous positions in exon 3 and intron 6, which were >2 kb apart and therefore could not be phased. This resulted in a genotype ambiguity in the 4th field: E*01:01:01:01/02, *01:03:01:04 could not be distinguished from E*01:01:01:03, 01:03:01:01. NGSengine enables easy detection of new alleles, such as demonstrated in sample IHWG9220 (Figure 3). This sample contains a new allele with a new SNP in exon 5, codon 297. All alleles currently present in the IMGT/HLA database have a proline at this position. The new SNP is a coding difference, changing amino acid proline (P) into a glutamine (Q). This new allele is the first example with a glutamine at this position, and would therefore represent a new HLA-E allele group with a new second-field typing. As the SNP is located outside the core region of HLA-E, the functional relevance of this glutamine remains unknown. Since phasing could not be accomplished, this sample will be resequenced with long-read sequencing technology, such that the sequence can be submitted to the IMGT/HLA database to expand the list of HLA-E alleles.

Conclusion

Typing results obtained with NGSgo-AmpX HLA-E are at high-to-allelic level resolution when analyzed in NGSengine. The amplification is highly robust and gives typing results concordant with the expected pre-types. Phasing of HLA-E is accomplished in exon 2 and exon 3, but not over longer distances, which can result in ambiguities in the 4th-field. This new typing strategy extends the existing NGSgo® workflow for Class I and Class II with HLA-E, such that ultimately HLA typing can be based on the full spectrum of clinically relevant HLA loci in one assay.

Figure 1: HLA-E amplicons of 35 representative samples, as analyzed by gel electrophoresis. Each sample shows a strong locus-specific amplicon of the expected size (3.6 kb), when compared to the DNA ladder (DL) and negative control (NC).

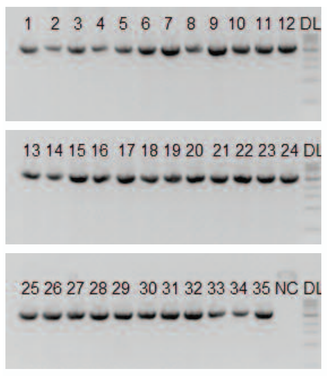

Figure 2: Three examples of HLA-E typing results of samples analyzed in NGSengine®.
Top panel shows typing results in NGSengine with data quality parameters of core+ region, exon+ region and amplicon respectively. a = mappability, b = lowest read depth, c = number of heterozygous positions, d = maximum noise level, e = delta signal to noise, f = second allele percentage, g = number of mismatches, h = IMGT/HLA database, i = number of best matching genotypes, j = homozygous typing result, k = best match, l = number of phasing regions. Bottom panel shows the alignment view of the three samples.

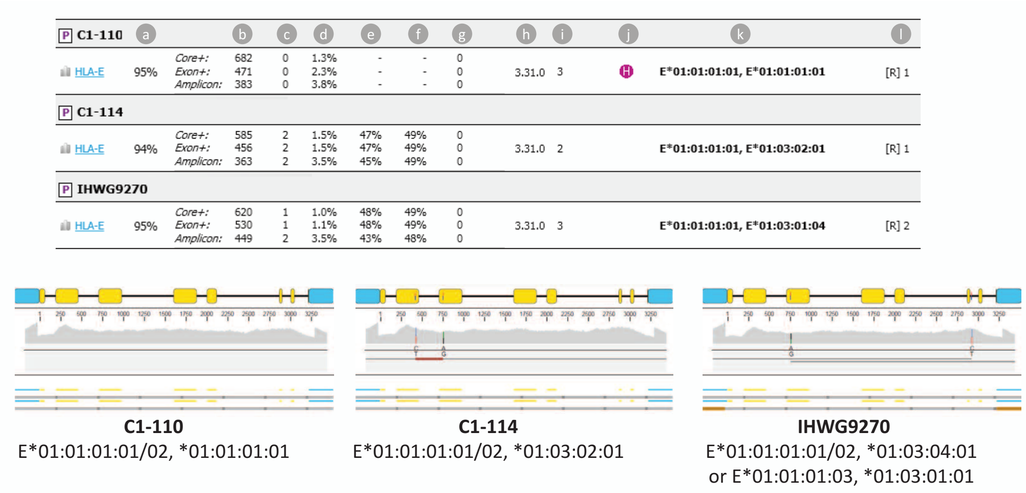

Figure 3. New HLA-E allele detected in sample IHWG9220. Left panel shows the alignment view in NGSengine. Right panel shows the close-up of exon 5 codon 297 with the new SNP, changing proline (CCG) into glutamine (CAG).

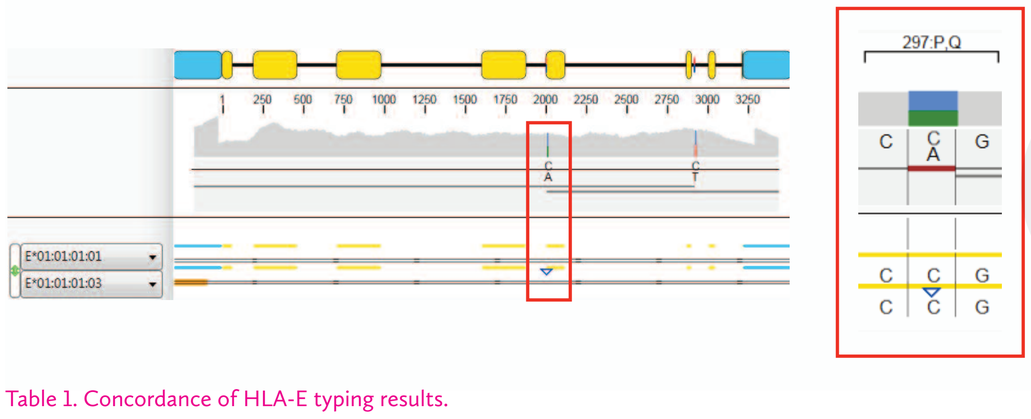

Table 1. Concordance of HLA-E typing results.

Sample ID	Pre-type	NGS-based typing result	Concordant
IHW01141	E*01:01:01:01	E*01:01:01:01/02, *01:01:01:01/02	yes
IHW01143	E*01:01:01:01	E*01:01:01:01/02, *01:01:01:01/02	yes
IHW01175	E*01:01:01:01, E*01:03:04	E*01:01:01:01/02, *01:03:04	yes
IHW01181	E*01:03:04, E*01:03:02:01	E*01:03:04, *01:03:02:01	yes
IHW01182	E*01:01:01:01, E*01:03:02:01	E*01:01:01:01/02, *01:03:02:01	yes
IHW01184	E*01:01:01:01	E*01:01:01:01/02, *01:01:01:01/02	yes
IHW09035	E*01:01	E*01:01:01:01/02, *01:01:01:01/02	yes
IHW09043	E*01:01	E*01:01:01:01/02, *01:01:01:01/02	yes
IHW09052	E*01:01:01:01	E*01:01:01:01/02, *01:01:01:01/02	yes

Table 2: HLA-E allele frequencies.

Group	Allele	Count	Frequency (%)
*01:01	01:01:01:01/02	98	45.0
	01:01:01:03	9	4.1
	01:01:01:04	1	0.5
	01:01:01:06	3	1.4
	01:01:01:08	1	0.5
	01:01:01:10	1	0.5
*01:03	01:03:01:01	23	10.6
	01:03:01:02	1	0.5
	01:03:01:04	7	3.2
	01:03:02:01	57	26.1
	01:03:02:02	5	2.3
	01:03:04	2	0.9
	01:03:05	1	0.5
Other	01:06	1	0.5
	01:10	2	0.9
New	01:XX	1	0.5
	01:01:01:XX	2	0.9
	01:01:01:XX	1	0.5
	01:01:01:XX	1	0.5
	01:03:XX	1	0.5
Total		218	100

Van de Pasch L.A.L., Van Heck M., Malestein R., Penning M.T., Rozemuller E.H.
GenDx, Utrecht, The Netherlands

GenDx.com