GENDX

personalizing diagnostics

NGS-based KIR typing at high resolution: challenges and solutions for the KIR framework genes

Killer-cell Immunoglobulin-like Receptors (KIRs) are encoded by 17 genes in the Leukocyte Receptor Complex (LRC). These genes are highly polymorphic on the level of structural variants, copy-number variation and sequence homology. These characteristics make it challenging to design specific amplifications, and to perform reliable analysis resulting in accurate identification of each KIR allele present in a sample. To enable high-resolution KIR genotyping, we are developing NGS-based strategies that involves whole-gene amplification of KIR genes, using KIR-specific primers that amplify one or more KIR genes.

Materials & Methods

Eight gDNA samples (Coriell) were amplified with gene-specific primers for KIR2DL1, 2DL2, 2DL3, 2DL4, 3DS1, 3DL1, 3DL2, 3DL3 and 3DP1. Samples were selected based on the availability of KIR genotype information (www.allelefrequencies.net/kir6007a.asp). Amplicons were checked by gel electrophoresis to score for absence/presence. Amplicons were processed by library preparation with NGSgo® (GenDx) and sequenced on an Illumina MiSeq platform (2x151 bp). Sequences were genotyped with NGSengine® software (GenDx), using KIR database IPD-KIR 2.7.0.

Results

We were able to successfully develop amplifications for the nine KIR genes (Figure 1). For the framework genes 3DL2, 3DL3, 2DL4, and 3DP1, all samples were found to be positive, as expected. For the other genes, KIR gene content was variable and therefore negative samples were expected and observed (Figure 1). Amplification of KIR-positive (POS) samples was successful, but KIR-negative (NEG) samples generated a weak false positive band (Figure 1, see 3DS1 and 2DL2).

Pilot studies in which multiple KIR genes were pooled in one assay were hampered by homologous reads which could not be unambiguously annotated to a specific gene. The assay setup was therefore simplified such that each amplicon was barcoded separately, allowing genotyping of each individual KIR in NGSengine (Figure 2). Only the KIR2DL1 amplification was found to co-amplify 2DL5. Full coverage of the analyzed genes was accomplished (Figure 3). For the majority of samples, allele-level resolution typings were obtained for the KIR-positive samples (Table 1). Most KIR genotypes obtained with NGS were concordant with the pre-type (Table 1, shown in green).

Many intron and exon variants were identified, representing new alleles. No samples with more than 2 copies were encountered yet.

One discordant typing result (Table 1, shown in red) was observed for KIR3DL1, showing a 3DL1*005 dropout. In two other samples, 3DL1*005 was amplified correctly, indicating that the primer is specific for 3DL1*005. Some discordances were also seen in 3DL2 and 2DL4. However, the NGS data was of such high quality that we considered the NGS-based typing result to be correct, probably because the technology used for the pre-type did not allow to type to this detail. For example, samples with a 2DL4*002 allele as pre-type, actually showed a 2DL4*0080101 or 2DL4*0080102 allele by NGS. These alleles differ by a single nucleotide in a homopolymer (10A vs 9A), that may discriminate between a membrane-bound and truncated soluble protein. This suggests that the 2DL4*008 allele group is more frequent than 2DL4*002.

Conclusion

- This proof-of-concept study has demonstrated feasibility of KIR genotyping, using the short read sequencing technology of Illumina.
- We have NGS typing workflows for 9 KIR genes and achieve high-to-allelic level typing results.
- The KIR2DL4*008 allele may be more frequent than KIR2DL4*002.
- This new typing strategy may be an attractive alternative to existing KIR genotype assays that only determine KIR gene content at a limited resolution, or complex sequencing methods aiming to sequence the entire LCR.

Figure 1. Locus-specific KIR amplicons of eight samples, analyzed by gel electrophoresis. The presence/absence of amplicons agrees with the expected KIR genotype (see Table 1).

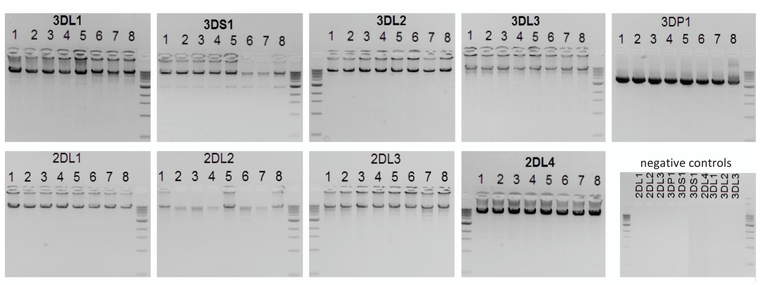

Figure 2. KIR genotype results of one representative sample (NA10833). Figure shows typing results in NGSengine with data quality parameters of core+ region, exon+ region, and amplicon respectively. a = mappability, b = number of best matching genotypes, c = homozygous typing result, d = best match, e = number of exon and intron mismatches, f = number of phasing regions.

	a	b	c	d	e	f
NA10833-2DL4						
KIR2DL4	156397/165964 (94%)	1		KIR2DL4*00501, KIR2DL4*0080101	[Ex] 0 [In] 0	[R] 4
NA10833-3DL1						
KIR2DL5	19100/19383 (98%)	>10		KIR2DL5A*0050101, KIR2DL5B*0020101	[Ex] 0 [In] 0	[R] 1
KIR3DL1	79314/80162 (98%)	3	H	KIR3DL1*0010101, KIR3DL1*0010101	[Ex] 0 [In] 4	[R] 1
NA10833-3DL2						
KIR3DL2	108386/117805 (92%)	1		KIR3DL2*0070103, KIR3DL2*01101	[Ex] 0 [In] 4	[R] 12
NA10833-3DL3						
KIR3DL3	67232/103885 (64%)	1		KIR3DL3*0030101, KIR3DL3*00904	[Ex] 0 [In] 0	[R] 4
NA10833-2DL1						
KIR2DL1	46501/49862 (93%)	1		KIR2DL1*0030204, KIR2DL1*0040101	[Ex] 0 [In] 0	[R] 3
NA10833-2DL2						
KIR2DL2	43797/54036 (81%)	6	H	KIR2DL2*0010102, KIR2DL2*0010102	[Ex] 0 [In] 0	[R] 1
NA10833-2DL3						
KIR2DL3	56965/66885 (85%)	6	H	KIR2DL3*0010101, KIR2DL3*0010101	[Ex] 0 [In] 0	[R] 1
NA10833-3DP1						
KIR3DP1	56199/60447 (92%)	1		KIR3DP1*0030102, KIR3DP1*006	[Ex] 1 [In] 0	[R] 4
NA10833-3DS1						
KIR3DLS1	42368/51332 (82%)	1	H	KIR3DS1*0130101, KIR3DS1*0130101	[Ex] 0 [In] 0	[R] 1

Figure 3. Alignment view and base variation % for each KIR gene, as analyzed in NGSengine. Full coverage was obtained with balanced allele ratios and low noise levels.

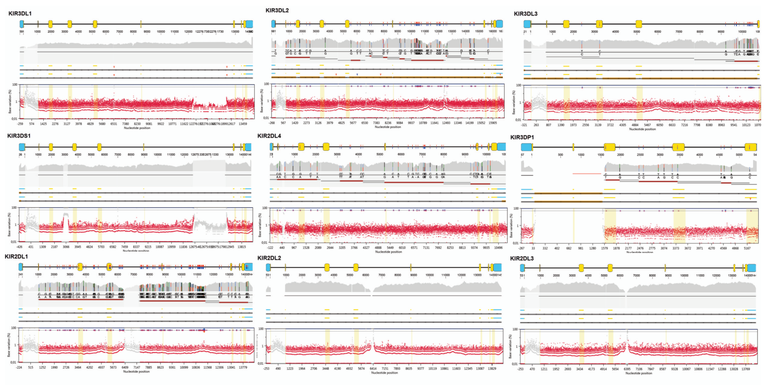

Table 1. Comparison of KIR pre-type with the obtained NGS-based genotype results. Green = concordant. Red = discordant.

	Sample	3DL1	3DS1	3DL2	3DL3	3DP1	2DL1	2DL2	2DL3	2DL4
Pre-type	NA10858	NEG	*013	*007	POS	POS	*00302, *00401	*001	*001	*005
	NA10830	*002	*013	*002, *007	POS	POS	*002, *00302	NEG	*001, *002	*00102, *005
	NA10861	*005	*013	*001, *007	POS	POS	*00302	NEG	*001	*00202, *005
	NA10853	*00101	*013	*001, *007	POS	POS	*001, *00302	NEG	*001, *005	*00202, *005
	NA10833	*00101	*013	*007, *011	POS	POS	*00302, *00401	*001	*001	*00202, *005
	NA11990	*00101, *005	NEG	*001	POS	POS	*001, *00302	NEG	*001, *005	*00202
	NA10851	*004, *005	NEG	*001, *005	POS	POS	POS	NEG	*001, *002	*00201, *005
	NA10857	*002, *008	NEG	*002, *009	POS	POS	*00302	*003	*001	*00202
NGS-based typing	NA10858	NEG	*0130101	*0070103	*0090101 (i.v.), *0140201 (i.v.)	*0030102, *0030202 (i.v.)	*0030205 (i.v.), *0040101 (i.v.)	*0010102 (i.v.)	*0010101 (i.v.)	*00501
	NA10830	*002 (i.v.)	*0130101	*0020101, *0070102 (i.v.)	*00206 (e.v.), *019	*0030202 (i.v.), *006	*0020101, *0030205 (i.v.)	NEG	*0010101 (i.v.), *0020101 (i.v.)	*0010201, *00501
	NA10861	*0050101	*0130101 (i.v)	*0010302, *0070102 (i.v.)	*00202 (i.v.), *00802 (i.v.)	*0030102, *0030202 (i.v.)	*0030204, *0030205 (i.v.)	NEG	*0010101/09/10, *0010111 (i.v.)	*0080101, *00501
	NA10853	*0010101 (i.v.)	*0130101	*0010101, *0070103	*00801 (i.v.), *0250101	*0030202 (i.v.), *014	*001, *0030204	NEG	*0010111, *00501	*0080101, *00501
	NA10833	*0010101 (i.v.)	*0130101	*00709, *01101	*0030101, *00904	*0030102, *006 (e..v.)	*0030204, *0040101	*0010102/06/07	*0010101/09/10	*0080101, *00501
	NA11990	*0010101	NEG	*00301 (e.v.), *017	*00601, *00904	*01001 (i.v.), *014	*001, *0030204	NEG	*0010101 (i.v.), *00501	*0080101
	NA10851	*0040101, *0050101	NEG	*0010302, *00501	*00101, *00206 (i.v.)	*0030202 (i.v.), *006	*0020103, *0030205 (i.v.)	NEG	*0010101 (i.v.), *0020103	*0080201 (i.v.), *011 (i.v.)
	NA10857	*002 (i.v.), *008	NEG	*0020101 (i.v.), *00901	*0030101 (i.v.), *0090101	*0090102, *0090102 (i.v.)	*0030204	*0030101/02/03	*0010103	*0010201, *0010305

i.v. = intron variant
e.v. = exon variant

Van de Pasch L.A.L., Vendelbosch S., Van Deutekom H., van Ham K., Penning M.T., Rozemuller E.H.
GenDx, Utrecht, The Netherlands

GenDx.com