Long-range whole-gene amplification approach for the characterization of novel KIR alleles

Marie Günther[1], Kathrin Lang[1], Kuntal Worah[1], Steffen Klasberg[1], Gerhard Schöfl[1], Alexander H. Schmidt[1,2], André P. Mäurer[1‡], Vinzenz Lange[1]
[1] DKMS Life Science Lab, Dresden, Germany; [2] DKMS, Tübingen, Germany

Introduction

The human killer-cell immunoglobulin-like receptors (KIR) have been reported to effect HLA-matched hematopoietic stem cell transplantation outcomes. To facilitate the selection of potential stem cell donors based on KIR genotypes we established an exon-based NGS workflow for KIR typing in 2015. Using this workflow, 2.6 million donors have been typed and thousands of novel KIR allele sequences have been discovered.
For the submission of novel alleles to the IPD-KIR database, we prefer to characterize the full genes including introns and UTR regions. Therefore, we developed a long-range amplification PCR and applied it for the full-length characterization of KIR alleles.

Methods

Based on haplotype KIR sequencing data (Pyo et al., 2010), we developed a comprehensive long-range PCR-approach resulting in amplicons up to 17 kb in length. For the separate amplification of each of the 16 KIR genes (KIR2DL1-5, KIR2DP1, KIR2DS1-5, KIR3DL1-3, KIR3DP1, KIR3DS1), a set of 19 forward and 22 reverse primers was used in specific combinations to amplify distinct KIR genes including the UTR regions (**Fig 1**).

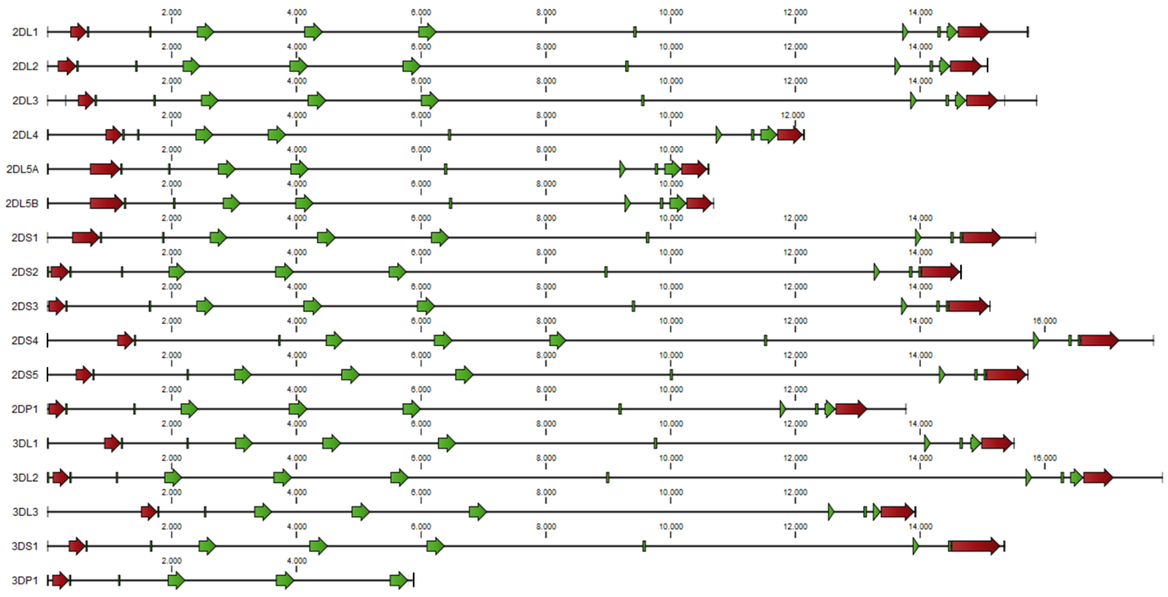

Figure 1: Amplified regions of KIR2DL1-5, KIR2DS1-5, KIR2DP1, KIR3DL1-3, KIR3DS1, and KIR3DP1. All amplicons include all exonic regions and the 5' and 3'UTRs. [red arrows: 5´/3´UTR; green bar or arrow: exon]

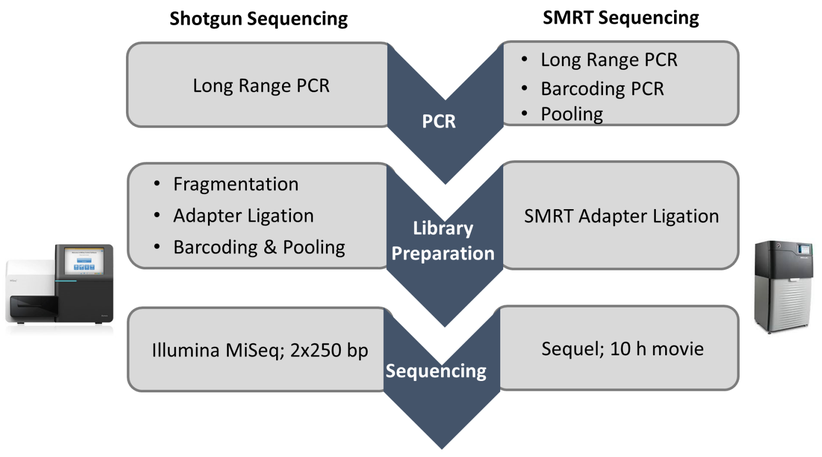

Figure 2: Sequencing workflows for the characterization of novel KIR alleles. Two independent long-range PCR products are sequenced by two different kinds of sequencing technologies, Shotgun Sequencing by Illumina and Single Molecule Real Time (SMRT) Sequencing by Pacific Biosciences.

The developed long-range PCR approach was used for the characterization of novel KIR alleles. To achieve this, we applied a dual sequencing workflow by combining data from long- and short-read technologies: Shotgun sequencing on Illumina MiSeq instruments and Single Molecule Real Time (SMRT) sequencing on Pacific Biosciences Sequel Systems (**Fig. 2**). Novel KIR alleles characterized with this dual workflow will be submitted to the IPD-KIR database.

Results & Conclusion

To evaluate the specificity of the PCR approach, we generated long sequencing reads using Pacific Biosciences Sequel SMRT technology for each KIR gene specific primer combination. On average 87% and at least 70% of the reads mapped to the targeted KIR gene (**Fig. 3**). These PCR assays have been successfully adapted to different sample types including DNA extracted from buccal swabs which are known to be a challenging target for long-range PCR.

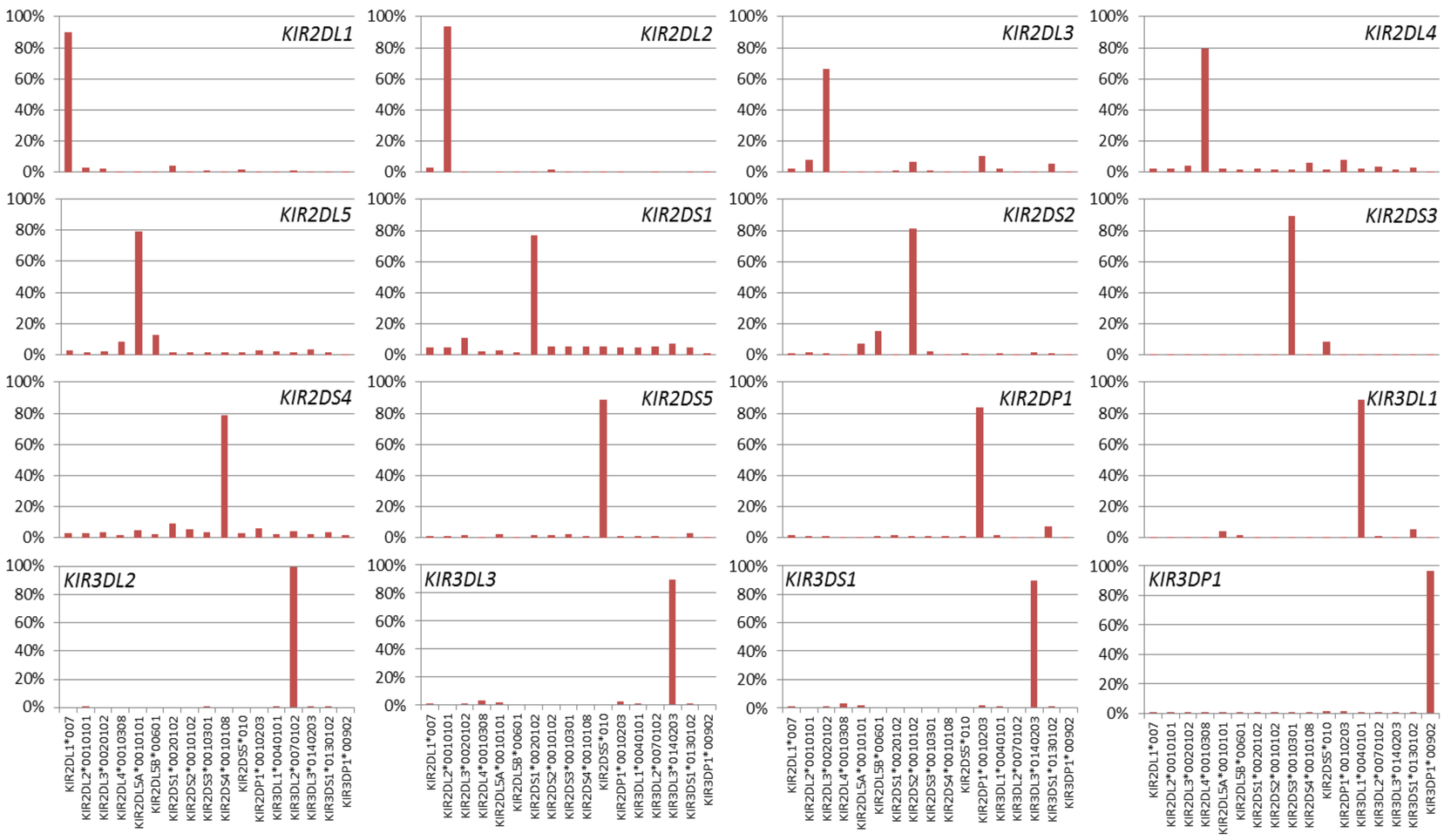

Figure 3: Specificity of KIR amplification. KIR-gene specific PCR was performed for up to 5 samples per KIR gene. Amplification products were sequenced by SMRT sequencing and mapped against a database with one full-length allele sequence per KIR gene.

Table 1: Overview of the successfully analyzed KIR samples with so far undescribed sequences.

	Analyzed sequences	Unique sequences
KIR3DP1	300	87
KIR2DL1	75	60
KIR2DL4	141	45
KIR2DS2	56	44
KIR3DL3	80	65
Σ	**652**	**301**

We successfully employed this long-range amplification approach for a dual sequencing workflow. So far we amplified and sequenced novel KIR alleles of more than 600 samples (**Tab. 1**). Submitting these sequences to the IPD-KIR database will contribute considerably to our knowledge of KIR diversity.

DKMS Life Science Lab

DKMS Life Science Lab

Blasewitzer Str. 43 * 01307 Dresden * www.dkms-lab.de