Completing the IPD-KIR Database: Full-Length KIR Gene Characterization of Frequent Novel KIR Alleles

Kathrin Putke1, Marie Günther1, Steffen Klasberg1, Alexander H. Schmidt1, Gerhard Schöff1, Anja Klussmeier1, Vinzenz Lange1

1 DKMS Life Science Lab, Dresden, Germany; 2 DKMS Deutsche Knochenmarkspenderratei, 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 consideration of KIR genotypes for stem cell donor selection KIR was added to the genotyping profile for newly registered potential stem cell donors in 2014. Since then, millions of donors were KIR genotyped and thousands of novel sequences discovered (Fig. 1). To report those novel KIR alleles to the IPD-KIR database, a project for full-length phased sequence characterization was initiated.

Results

522 KIR allele sequences were submitted to IPD-KIR, including 374 unique sequences for novel alleles of KIR2DL1 (70), KIR2DL4 (38), KIR2DL5 (44), KIR2DS1 (21), KIR2DS2 (43), KIR3DL3 (77) and KIR3DP1 (81) (Fig. 3). These sequences will more than double the number of full-length sequences in the IPD-KIR database (Release 2.8.0: 323 alleles with full-length sequence information). Furthermore, another 380 novel alleles for the genes KIR2DL4, KIR2DS3, KIR2DS4, KIR2DS5 and KIR3DL2 have already been characterized and will be submitted soon.

Since initial sample selection was based on variations found in an exon-based typing workflow, the majority of sequenced alleles contained CDS variations. In addition we found many intronic variants.

Conclusion

Despite the high numbers of novel KIR alleles reported here, these correspond only to a fraction of the novel alleles discovered during our routine genotyping operation. Nevertheless, due to the focus on the more frequent variants, the addition of these sequences will represent a huge step forward for the IPD-KIR database towards a comprehensive representation of common allelic variation.

Methods

A gene-specific long-range PCR-approach was developed for the complete amplification (including the UTR regions) of each KIR gene (KIR2DL1-5, KIR2DP1, KIR2DS1-5, KIR3DL1-3, KIR3DP1, KIR3DS1), resulting in amplicon lengths between 5.5 and 17.5 kb. To characterize novel KIR alleles two independent PCR reactions for both short-read Illumina sequencing and long-read PacBio Sequel sequencing were performed. Subsequently, the dual redundant reference sequencing (DR2S) software was applied to generate fully phased high-quality error-corrected sequences (Fig. 2). Finally, the sequences were submitted to ENA and IPD-KIR with TypeLoader2.

References