38th International Congress of the International Society of Blood Transfusion (ISBT), 23.-27. Juni 2024, Barcelona, Spain

Full-length *RHD* and *RHCE* haplotypes by Nanopore sequencing of five overlapping, generic long-range PCR amplicons

G. A. Thun[1], M. Gueuning[1], S. Sigurdardottir[2], N. Trost[2], K. Engehausen[2], C. Engström[3], S. Meyer[2], M. P. Mattle-Greminger[1]

[1] Blood Transfusion Service Zurich (SRC), Research and Development, Schlieren, Switzerland
[2] Blood Transfusion Service Zurich (SRC), Molecular Diagnostics and Cytometry, Schlieren, Switzerland
[3] Blood Transfusion Service Zurich (SRC), Immunohematology, Schlieren, Switzerland

Background

Elucidating the allelic composition of the long and highly homologous *RHD* and *RHCE* genes on the haplotype level, which can be transfusion-relevant, is a challenging endeavor. So far, it required very laborious techniques like cloning the allele into a suitable vector or performing transcript (cDNA) analysis in reticulocytes. Thus, it is common to only deduce the most likely genetic allele from phenotype data. For instance, in a case of weak RhC and normal Rhc antigen expression, an identified heterozygous non-synonymous variant would be assigned to the *RHCE*C* allele. Newest long-read sequencing methodology finally facilitates *RHD* and *RHCE* haplotype re-construction by single-molecule sequencing of long-range PCR amplicons.

Aim

Here, we developed a long-range PCR approach using generic primers, designed to co-amplify the homologous *RHD* and *RHCE* genes. The use of generic primers reduces the risk of unnoticed allelic dropout in case of *RHD*/*CE* hybrid alleles.

Methods

To fully cover exons 1 to 10 from *RHD* (~57 kb) and *RHCE* (~59 kb), we designed five generic primer pairs to co-amplify both genes. PCR amplicon lengths ranged from 12.3 to 15.2 kb for *RHD*, and 13.3 to 15.2 kb for *RHCE*. Overlaps of PCR amplicons comprised at least 1.1 kb and were chosen, whenever possible, to lie within polymorphic regions as haplotype reconstruction requires heterozygous variants. The five PCR products per sample were equimolarly pooled and barcoded to allow Oxford Nanopore sequencing of multiple samples on the same flow cell. Bioinformatic analysis involved reference-based variant calling and subsequent phasing using an extracted 165-kb HG38 reference sequence of the entire *RHD* and *RHCE* gene locus. We tested the method with two donors known to carry heterozygous *RHCE* C/c, E/e and c.733C>G variants, the latter causative for V+ and VS+ antigen formation. Phenotypically, both donors showed DCcEe, but the C antigen was weak in one of them.

Results

Both samples showed good distributions of read depths along both genes for all PCR-amplicons (least covered amplicons with 250x and 150x, respectively). Allelic distribution patterns of variants clearly showed that the generic PCR amplicons of both genes were sufficiently different to unequivocally map to the respective gene reference sequence. Genetic phasing worked well, which not only allowed the re-construction of full-length haplotypes for both genes, but also elicited converse results regarding the *RHCE* allele backgrounds of the c.733C>G variant in both donors. While the donor with the serological weak C showed a classical ce^{s} allele (*RHCE*01.20.01* | *RHCE*04*), the other donor showed the V and VS-causing variant on a *RHCE*C* background (*RHCE*02.30* | *RHCE*03*). The complete sequence information for *RHD* allowed to rule out alternative complex alleles like (C)ce^{s}. Based on these results, we speculate that the observed weak C in the ce^{s}/CE sample is most likely a consequence of the R(z)-haplotype where DCE all lie on the same allele.

Summary
Here, we present an amplicon-based Nanopore sequencing method suitable to gain full-length *RHD* and *RHCE* haplotypes. By relying on generic PCR primers, the method not only reduces the number and complexity of PCRs but also the risk of unnoticed allelic dropout in case of hybrid alleles, which are prevalent in the Rh blood group system.