Swisstransfusion 2023, Rorschach

Adaptive sampling to resolve complex structural variants in blood group genes by nanopore sequencing

Auflösung komplexer struktureller genetischer Varianten im Blutgruppengenom mittels Adaptive Sampling Nanopore Sequenzierung

M. Gueuning[1], Y. Merki[2], G. A. Thun[1], N.Trost[2], C. Gassner[3], B. Veldhuisen[4], C. Engström[5], S. Meyer[2], M. P. Mattle-Greminger[1]

[1] Blood Transfusion Service Zurich (SRC), Department of Research and Development, Schlieren, Switzerland

[2] Blood Transfusion Service Zurich (SRC), Department of Molecular Diagnostics and Cytometry, Schlieren, Switzerland

[3] Institute of Translational Medicine, Private University in the Principality of Liechtenstein (UFL), Triesen, Liechtenstein

[4] Department of Diagnostic Immunohematology, Sanquin Diagnostic Services, Amsterdam, The Netherlands

[5] Blood Transfusion Service Zurich (SRC), Department of Immunohematology, Schlieren, Switzerland

Introduction: Adaptive sampling is a software-driven enrichment method unique to Oxford Nanopore Technologies (ONT) and allows simultaneous sequencing of the entire blood group genome. Without any restrictions on read length, it can excel in resolving complex structural variants (SVs) like hybrid alleles, which are a hallmark of some blood group systems like the RH. Here, we used adaptive sampling to tackle RHCE genotype-phenotype discrepancies unresolved by other methods and suspected to harbor a SV.

Methods: We reassessed three RHCc genotype-phenotype discrepancies, which we observed in routine high-throughput donor typing, by PCR-SSP, MLPA, and extensive serological analyses. Suspecting a SV, we employed high density MALDI-TOF MS analysis along the *RHCE* gene to get hints on copy number variation. As unresolvable, we took advantage of latest developments of ONT's adaptive sampling for sequencing the entire *RHCE* gene as single reads for one of the samples. The sequencing library was built with high-molecular weight gDNA. Beside *RHCE* we targeted the entire blood group genome (48 genes; ~8 Mb incl. 50 kb flanking regions) in adaptive sampling. Structural variant breakpoints in all three samples were confirmed by bridge-PCRs and Sanger sequencing.

Results: All three donors were serologically typed as C+c- but genotyped as *RHCE*Cc.* The high-density MALDI-TOF MS assay pointed toward an identical potential SV located near the end of the *RHCE* gene. The type or location of the SV, however, could not be determined. ONT sequencing of one sample exhibited high mean read length (N50 > 30 kb) with a median coverage of ~15x across all 48 genes. The availability of particularly long reads (20 reads > 50 kb; max. 250 kb) spanning the RH locus allowed identifying a large deletion. Specifically, we identified a novel ~8.6 kb deletion spanning from intron 8 to intron 9 of *RHCE*. Bridge-PCRs and Sanger sequencing confirmed the exact breakpoint locations, with an identical deletion in the other two samples.

Conclusions: Adaptive sampling is a promising enrichment method for simultaneously targeting all blood group genes without restrictions on read length, which opens new avenues for identifying SVs. It has allowed us to resolve long-tackled RHCE genotype-phenotype discrepancies by discovering a large novel deletion. Although SVs are a hallmark of RH, to our knowledge only two other large deletions have yet been reported for *RHCE*, suggesting the need for suitable approaches like the one presented here.