Transfusion Medicine and Hemotherapy

Research Article

Transfus Med Hemother
DOI: 10.1159/000550249

Received: August 28, 2025
Accepted: December 23, 2025
Published online: January 8, 2026

Introduction of an Optimized Protocol for Long-Read Nanopore Sequencing of Blood Group Genes in Immunohematology Case Studies

Lea Wörner Gabriele Rink Xenia Merkle Peter Bugert

Institute of Transfusion Medicine and Immunology, Heidelberg University, Medical Faculty Mannheim, German Red Cross Blood Service Baden-Württemberg - Hessen, Mannheim, Germany

Keywords

Long-read sequencing · Nanopore sequencing · Long-range PCR · Blood group genes · Immunohematology case studies

Abstract

Introduction: In immunohematology case studies the knowledge about variants of the blood group gene of interest can facilitate antibody diagnosis. Known gene variants can be rapidly genotyped by specific methods, but the identification of unknown variants requires sequencing of the gene. High throughput next-generation sequencing (NGS) technologies represent important tools for DNA sequencing of many targets in larger numbers of samples but are less suitable for the analysis of one or few genes in single samples. Nanopore sequencing (Oxford Nanopore Technologies, ONT) is a fast sequencing technology that could fulfill the requirements for targeted gene sequencing in case studies. Here, we describe an optimized protocol for long-read nanopore sequencing of blood group genes that enables analysis of whole genes within less than 7 h from DNA extraction to genotype determination. ***Methods:*** Primers for long-range PCR (LR-PCR) were designed for the blood group genes *ACKR1*, *CD151*, *BCAM*, *KEL*, *SLC14A1*, *GYPA*, *GYPB*, *GYPE*, *RHD*, and *RHCE* with amplicon sizes in the range of 2.4–15.8 kilo base pairs (kbp). For evaluation of the sequencing data, 22 samples with 25 known gene variants were selected. The optimized sequencing workflow included DNA extraction from EDTA blood, LR-PCR amplification, library preparation, nanopore sequencing on the MinION Mk1D sequencing device with FLO-MIN114 (MinION) flow cells and data analysis including variant detection and genotyping. In addition, the workflow was tested on the MinION sequencing device with FLO-FLG114 (Flongle) flow cells and the PromethION 2 Solo sequencing device with FLO-PRO114 (PromethION) flow cells. ***Results:*** Using the outlined long-read nanopore sequencing protocol, sequencing data for reliable variant calling were obtained. All alleles were identified and the zygosity could be determined based on the read counts, except for *GYPB* in one sample. Besides sequencing on the MinION Mk1D sequencing device MinION flow cells, successful application of the sequencing protocol to the Flongle flow cells and the PromethION sequencing device using PromethION flow cells was demonstrated. As expected, mean coverage and mean Q scores varied between the flow cells and devices. ***Conclusion:*** The optimized nanopore sequencing protocol enabled the generation of long-read sequence data and identification of blood group gene variants within a working day. This approach is suitable for molecular analyses of different blood group genes in immunohematology case studies under the same LR-PCR and sequencing conditions.

© 2026 The Author(s).
Published by S. Karger AG, Basel

Introduction

In transfusion medicine, blood group genotyping becomes relevant when serological testing cannot clearly determine a person's blood phenotype. In terms of patient safety, there are various scenarios that make

karger@karger.com
www.karger.com/tmh

© 2026 The Author(s).
Published by S. Karger AG, Basel

Karger
OPEN ACCESS

This article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC) (https://karger.com/Services/OpenAccessLicense). Usage and distribution for commercial purposes requires written permission.

Correspondence to:
Peter Bugert, peter.bugert@medma.uni-heidelberg.de

blood group genotyping necessary [1]. A general limitation of serological testing is the availability of anti-sera, especially commercially available anti-sera, which is particularly relevant with regard to rare antigens or phenotypes [2–4]. Additionally, serological testing in pre-transfused or chronically transfused patients with a higher risk for alloimmunization, as well as testing of immunoglobulin-coated red blood cells is challenging [3–5]. Weak or unspecific antigen-antibody reactions in this context lead to ambiguous serological results [4, 5].

Common genotyping methods such as MALDI-TOF mass spectrometry, sequence-specific-priming PCR (PCR-SSP), TaqMan-PCR or BeadChip assays allow the targeted analysis of known variants but they lack the possibility of detecting novel variants [6–10]. The number of variants analyzed by these methods is limited to the assay design and the included variants. Application of sequencing techniques, such as Sanger sequencing or NGS, to the field of immunohematology overcomes this limitation which is especially important in case of resolving phenotype-genotype discrepancies or unknown alloantibody specificity. The ability of NGS technologies to sequence multiple blood group genes from different samples in parallel offers great potential in terms of donor screening or antibody specification in patients [5, 11]. Nevertheless, NGS techniques have disadvantages with regard to costs, time and the maximum read length of 400 base pairs (bp). Short-read sequencing is less suitable for the analysis of homologous regions or hybrid genes as found in the Rh or MNS blood group system [5, 11–13]. The analysis of short-read sequences is also less suitable for phasing of variants separated in distant gene regions, which is important because the *cis* or *trans* configuration of variants can affect the blood group phenotype [12, 14–16].

In the context of those restrictions, long-read sequencing, as a third-generation sequencing technology, offers a promising alternative for the analysis of novel (structural) variations and haplotype phasing. These techniques generate sequence reads with a length up to several megabase pairs [13, 17]. Currently, two long-read sequencing platforms are commercially available using different technologies: the single-molecule real-time sequencing technology (Pacific Biosciences, PacBio) and the nanopore sequencing technology (Oxford Nanopore Technologies [ONT]) [13, 17]. The basic principle of nanopore sequencing relies on nucleotide-specific voltage variations, from which sequencing data are generated. Voltage variations occur when a DNA molecule passes a nanopore, which functions as a voltage sensor [18]. Depending on the application and target, ONT provides different flow cell types and sequencing devices with varying output capacities [14]. Since the market introduction in 2014, nanopore sequencing is used for routine diagnostics in various fields, for example, in oncological diagnostics or other genetic disease, for fast pathogen identification in infectiology and also for HLA typing but not for blood typing yet [18].

In the past 5 years, first approaches using nanopore sequencing for the analysis of blood group genes were published. In 2020, Lane et al. [19] detected deletions in the *GYPB* gene of the MNS blood group system. Thun et al. [20] identified a novel regulation variation in the context of the A_3 phenotype in the ABO blood group system, while Tounsi et al. [21] tested the efficiency of this sequencing method for the analysis of the complete *RHD* gene locus. Gueuning et al. [22, 23] used the nanopore sequencing technique to resolve genotype-phenotype discrepancies in the Kidd blood group system and showed additionally the successful generation of blood group haplotype reference data for a collection of *AB0* alleles. A similar approach was published by Srivastava et al. [24] who generated haplotype reference data for the *ACKR1* gene of the Duffy blood group system.

Here, we developed an optimized protocol for long-read nanopore sequencing of the blood group genes *ACKR1*, *CD151*, *BCAM*, *KEL*, *SCL14A1*, *GYPA*, *GYPB*, *GYPE*, *RHD*, and *RHCE* (shown in Fig. 1). Our aim was the development of a cost-effective, time-efficient, and easy-to-implement sequencing workflow for single sample analysis, that generates sequencing data with a quality that allows reliable blood group gene variant detection and genotype determination.

Methods

Primer Design

Long Range (LR) PCR primers were designed using version 7.11 of the Clone Manager Suite 7 program (Sci Ed Software LLC, Westminster, CO, USA). Sequence specificity was tested using BLAST blastn suite. Up to one binding site in addition to the addressed target sequence was accepted. Primers were designed so that LR-PCR amplicons did not exceed a length of 16 kbp. The primer design for larger blood group genes considered two or more LR-PCR amplicons with at least 2 kbp overlap and a minimum of two SNVs with a minor allele frequency (MAF) of at least 0.2 in the overlapping region (tested using Variation Viewer 2.1.4; NCBI). The human genome assembly GRCh38 (hg38) was used as the reference sequence for primer design and for mapping sequence reads. LR-PCR primer sequences with location and amplicon length are listed in Table 1 with schematic illustrations shown in Figure 2.

Blood Samples and DNA Extraction

Anonymized EDTA blood samples from 22 blood donors and patients were selected with regard to known variant alleles of the blood group systems Duffy, Raph,

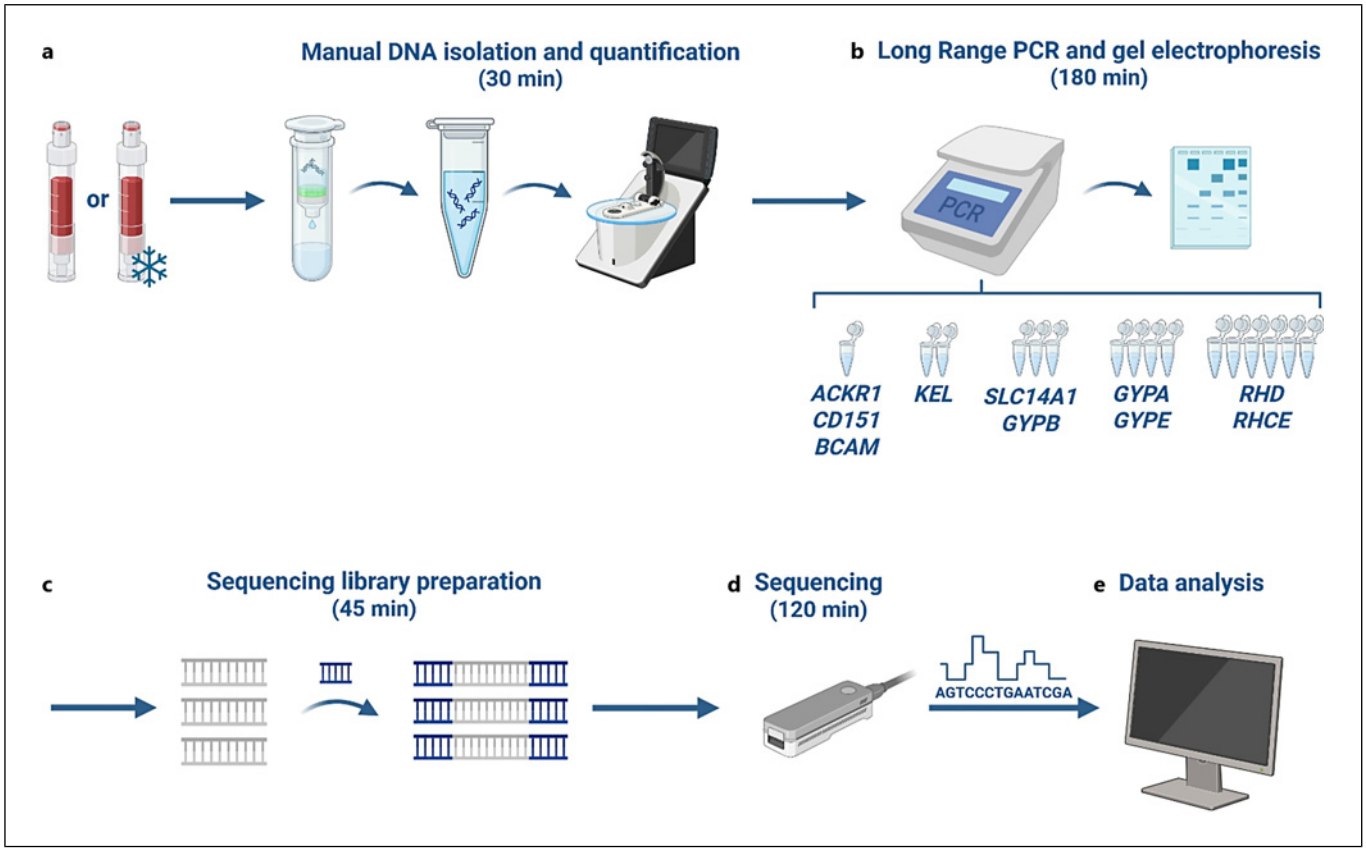

Fig. 1. a–**e** Schematic illustration of the optimized workflow for long-read nanopore sequencing of blood group genes. The single steps of the protocol starting with the manual DNA isolation from frozen or fresh blood samples (**a**) and ending with the analysis of generated long-read sequencing data (**e**). Created in BioRender. Wörner, L. (2026) https://BioRender.com/ac3yyg0.

Lutheran, Kidd, Kell, MNS, and Rh (shown in Table 2). Variant alleles were previously determined by PCR-SSP, TaqMan-PCR or NGS. Sample BGS_014 was referred to our lab with a Lu(a-b-) phenotype and unknown molecular background. Nanopore sequencing of this sample was performed first and the newly identified variant was confirmed by PCR-SSP afterward. All donors consented to the use of blood samples for research and development purposes. Genomic DNA was isolated from 200 µL blood using the QiAamp Blood Mini Kit (Qiagen, Hilden, Germany) with a final elution volume of 100 µL. Concentration of DNA was measured using photometry (NanoDrop Lite, Thermo Fisher Scientific, Darmstadt, Germany). DNA samples were stored at −20°C until used for LR-PCR.

Long-Range PCR

Singleplex LR-PCR was set-up as follows: (1) forward and reverse primer pairs were mixed and diluted to a final concentration of 2.5 µmol each; (2) genomic DNA was diluted to a final concentration of 20 ng/µL; (3) in a total volume of 25 µL the LR-PCR contained 5 µL genomic DNA, 2.5 µL primer mixture, 5 µL 5× PrimeStar GXL buffer, 2 µL PrimeStar dNTP mixture, 0.5 µL PrimeStar GXL DNA polymerase (Takara Bio Europe SAS, Saint-German-en-Laye, France). LR-PCR was performed on a Thermal Cycler (T100, Bio-Rad Laboratories, Feldkirchen, Germany) using a two-step cycling program with 30 cycles of 98°C for 10 s and 68°C for 5 min. Amplification was verified on a 2% agarose gel stained with GelRed Nucleic Acid Gel (Biotium, Fremont, CA, USA). For sequencing of blood group genes covered by only one LR-PCR amplicon, a minimum of two PCRs was set up to obtain a minimum of 40 µL PCR product for library preparation.

Library Preparation and Nanopore Sequencing

Library preparation was performed using the Ligation Sequencing Kit V14 SQK-LSK114 (ONT, Oxford, UK) with a modified version of the manufacturer's protocol (ACDE_9163_v114_revW_30Jan2025) to shorten library preparation time to 45 min. Hands-on time was significantly reduced by a consecutive workflow for DNA repair, End Prep, adapter ligation and clean-up without the intermediate washing steps. Briefly, the NEBNext Ultra II End Repair/dA-Tailing Module

Table 1. Primers for LR-PCR of blood group genes analyzed by nanopore sequencing

Blood group system (gene)	Primer name	Primer sequence 5′–3′	Amplicon location (on hg38)	Amplicon size (bp)
Duffy (*ACKR1*)	FY-LR-F	ACCACCACCACCTTTCTCC	chr1: 159,204,404–159,206,837	2,434
	FY-LR-R	GCTCTGGGCTTTGTCATCAC		
Raph (*CD151*)	CD151-LR-F	CTCCCACTCTTGGCGTCTTGG	chr11:832,364–838,949	6,579
	CD151-LR-R	GGCTCTGCCCTTCTTTGTCCTG		
Lutheran (*BCAM*)	BCAM-LR-F	TAAGCCCGGACAGCTCTAGG	chr19:44,808,895–44,820,985	12,091
	BCAM-LR-R	GCAGGGCTGGTTCTCTGATG		
Kidd (*SLC14A1*)	SLC14-LR1-F	GAGGCAGCAGAGTATGTCCAAG	chr18:45,723,583–45,735,119	11,537
	SLC14-LR1-R	GGAGGGAGGAGTCAAGTGTATG		
	SLC14-LR2-F	CTGCCTTGTCAGGTCAATCAATC	chr18:45,732,828–45,745,474	12,647
	SL14C-LR2-R	TGAGCAATGGAGGTTCTACAC		
	SLC14-LR3-F	CCTTGTGATTACATTGGGTCC ATTC	chr18:45,743,213–45,752,586	9,375
	SLC14-LR3-R	CGATACCTACTTATGCCTGTG ATTG		
Kell (*KEL*)	KEL-LR1-F	CTTCCCAATAAATGCTTGCCTCTG	chr7: 142,962,684–142,950,581	11,914
	KEL-LR1-R	CAGAGATGCCACTGATCCTTC		
	KEL-LR2-F	ATGGCAGTGAGCAGGCTTAC	chr7: 142,953,054–142,940,696	12,397
	KEL-LR2-R	GGTTCAGCAGCAGAGATACC		
MNS (*GYPA*)	GYPA-LR1-F	TGTCCTGCTGCTCAGATTACC	chr4: 144,141,810–144,132,971	8,860
	GYPA-LR1-R	CCAGAAAGCACTTAGACTTTC TATG		
	GYPA-LR2-F	TCTTACCCATTAGCAGCCACTTC	chr4: 144,136,015–144,123,205	12,833
	GYPA-LR2-R	TGCTGTCTGGGAGATAGTGATTTC		
	GYPA-LR3-F	GTGCCTCACCACCTCTTCAG	chr4: 144,125,336–144,114,600	10,756
	GYPA-LR3-R	CCCTCCTCCCTCAAATCAGAAAG		
	GYPA-LR4-F	AGAACGCTGGGTCTGGAATC	chr4: 144,117,224–144,108,211	9,004
	GYPA-LR4-R	GGGCTTTAGTCTGAGGGAAAC		
MNS (*GYPB*)	GYPB-LR1-F	TGTCCTCACATTTCCATACCTTTC	chr4: 144,020,057–144,011,200	8,858
	GYPB-LR1-R	TGAAACCCCGTGTCTACTGAA ATAC		
	GYPB-LR2-F	TATCTCCCGATGCTATCCCTTCC	chr4: 144,013,809–144,004,398	9,412
	GYPB-LR2-R	GCACAGAGGCTTGCTATTGG		
	GYPB-LR3-F	TCAAACCAGCAGCTTATCACATC	chr4: 144,006,633–143,995,166	11,468
	GYPB-LR3-R	TCCTCGTCATCAGACCTACC		
MNS (*GYPE*)	GYPE-LR1-F	ACGGCACTGCTTCCTCACAC	chr4: 143,908,001–143,897,269	10,733
	GYPE-LR1-R	CCCTGGGCAACATAAGAAGAC		
	GYPE-LR2-F	TTCCAGGATTGGCTCCGTTAC	chr4: 143,899,238–143,886,740	12,499
	GYPE-LR2-R	CTTTGCCTTTGGTGCTCTCAG		
	GYPE-LR3-F	GTCCCAGTTGCTCAGTAGAATG	chr4: 143,889,639–143,878,642	10,998
	GYPE-LR3-R	CTTCAGGGAAACGATGGACAAG		
	GYPE-LR4-F	TAAACCTCTGCCATGCAGAAATTC	chr4: 143,880,743–143,870,631	10,113
	GYPE-LR4-R	CACATCCAGGTAATGCTGTTG		
Rh (*RHD*)	RHD-LR1-F	GGCCCTGGCTTTATTCTCTG	chr1:25,271,437–25,279,511	8,075
	RHD-LR1-R	CAGGAGGATTCAAACCCTAC		
	RHD-LR2-F	AGCCCTCTTGGTGATTCTTC	chr1:25,277,481–25,290,778	13,316
	RHD-LR2-R	CTGATGACCATCCTCAGGTTG		
	RHD-LR3-F	GTGCCACTTGACTTGGGACT	chr1:25,287,556–25,301,290	13,734
	RHD-LR3-R	TCCTGAACCTGCTCTGTGAAGTGC		
	RHD-LR4-F	GTCAGCAGAGCTTTCTTAGGC	chr1:25,297,757–25,309,721	11,966
	RHD-LR4-R	CAATGTTTTGTGGCTTGGACTGG		
	RHD-LR5-F	CATCCCCTTTGGTGGCC	chr1:25,306,494–25,322,282	15,788
	RHD-LR5-R	CACCCGCATGTCAGACTATTTGGC		
	RHD-LR6-F	GTCCTTTTTGTCCCTGATGACC	chr1:25,316,931–25,329,215	12,284
	RHD-LR6-R	CTTAGCTTACTGGATGACCACCA		

Table 1 (continued)

Blood group system (gene)	Primer name	Primer sequence 5′–3′	Amplicon location (on hg38)	Amplicon size (bp)
Rh (*RHCE*)	RHCE-LR1-F	GTATCCACTTTCCACCTTCCAC	chr1:25,421,917–25,413,852	8,107
	RHCE-LR1-R	CAGGAGGATTCAAACCCTACC		
	RHCE-LR2-F	GGGAGTGAAGCTGGTTCTGC	chr1:25,414,370–25,402,683	11,688
	RHCE-LR2-R	CTTCCCCAAGACAGCACCCG		
	RHCE-LR3-F	CACGCCTAGCTGATTTTTCTGTC	chr1:25,405,199–25,391,153	14,069
	RHCE-LR3-R	CTGCTCTGTGATGGTCAAATGC		
	RHCE-LR4-F	GAGAACTCCCTGATTCCCAG	chr1:25,395,648–25,382,719	12,930
	RHCE-LR4-R	CCAATGTTTTGTGGCTCGGAATG		
	RHCE-LR5-F	CCATCCCCATTTGGTGGCG	chr1:25,385,947–25,370,113	15,835
	RHCE-LR5-R	GGTTTCTTCCAGCTTTTGCATTGC		
	RHCE-LR6-F	GGCGCTATCTCGGCTCAACG	chr1:25,372,478–25,361,797	10,682
	RHCE-LR6-R	GCCTGGTGATCAAGACTCTCC		

(E7546, New England Biolabs, Frankfurt am Main, Germany) was used for end repair and the Quick T4 DNA Ligase of the NEBNext Quick Ligation Module (E6056, New England Biolabs, Frankfurt am Main, Germany) instead of the in the manufacturer's protocol recommended Salt-T4 DNA Ligase (M0467, New England Biolabs, Frankfurt am Main, Germany) for adapter ligation. The corresponding number of LR-PCR products from one DNA sample was pooled and 40 µL of the PCR pool were used for further processing: (1) 20 µL AMPure Beads were added, incubated at room temperature (RT) for 3 min and placed on a magnetic stand to pull down the DNA-bound magnetic beads; (2) after removing the supernatant, beads were centrifuged at 500 *g* for 1 min and again placed on the magnetic plate to completely remove remaining supernatant; (3) beads were resuspended in 20 µL End Prep mix containing 2.5 µL Ultra II End Prep reaction buffer, 1 µL Ultra II End Prep Enzyme mix, and 16.5 µL ddH_20 and incubated at RT for 3 min; (4) 20 µL Adapter Ligation mix (8 µL Ligation Buffer, 4 µL Quick T4 DNA Ligase, 8 µL ddH_20) and 2 µL Ligation Adapter were added and incubated at RT for 10 min; (5) another 20 µL AMPure Beads were added, incubated at RT for 3 min and placed on a magnetic stand; (6) after removal of the supernatant the beads were resuspended in 100 µL Long Fragment Buffer (LFB) and again placed on the magnetic stand; (7) step (6) was repeated once and after centrifugation at 500 *g* for 1 min remaining LFB was removed completely; (8) beads were resuspended in 15 µL (35 µL for sequencing on the PromethION device) Elution Buffer and incubated at RT for 3 min; (9) the supernatant containing the DNA library was transferred to a 0.5 mL low binding tube and either directly used for sequencing or stored at 4°C for up to 24 h.

For nanopore sequencing, primarily the MinION Mk1D device with FLO-MIN114 flow cells (MinION flow cell; ONT) was used. For testing the sequencing workflow on other ONT sequencing devices we also used the MinION device with the FLO-FLG114 flow cell in the Flongle adapter (Flongle flow cell; ONT) and the PromehtION 2 Solo device with the FLO-PRO114 flow cell (PromethION flow cell; ONT). Priming and loading of the flow cells using the Flow Cell Priming Kit EXP-FLP004 (ONT) were performed according to the manufacturer's protocol with two exceptions: (1) no bovine serum albumin was added to the Flush mix for flow cell priming of the MinION flow cell; (2) loading of the Flongle flow cell was performed rapidly with 20 µL priming mix being pipetted onto the loading port three times. Since MinION and PromethION flow cells were reused up to five times as standard with the exception of BGS_007, sequenced on a MinION flow cell, that has been used for the sixth time, washing and storage preparation using the Flow Cell Wash Kit EXP-WSH004 (ONT) were performed according to the manufacturer's protocol. The sequencing run was set up for a duration of 2 h with a minimum read length of 300 bp using the super accurate (SUP) model for basecalling with the Dorado algorithm version 7.6.8 in the MinKNOW software version 24.11.10 (ONT). Mapping the unfiltered sequencing data (fastq) was performed using the Geneious Prime software (Version 2025.1.3; Dotmatics, Boston, USA) with the MiniMap2 algorithm (Version 2.24, asm10) [27]. A minimum coverage of 50 reads was set as a cut-off for variant calling. Based on previous sequencing results, an allele proportion of 0.8 was set as the cut-off for homozygosity and 0.2 as the cut-off for heterozygosity.

The applicability of the sequencing protocol was additionally tested on other ONT sequencing devices or flow cells. For this purpose, the three blood group genes *ACKR1*, *SLC14A1* and *RHD* with three biological

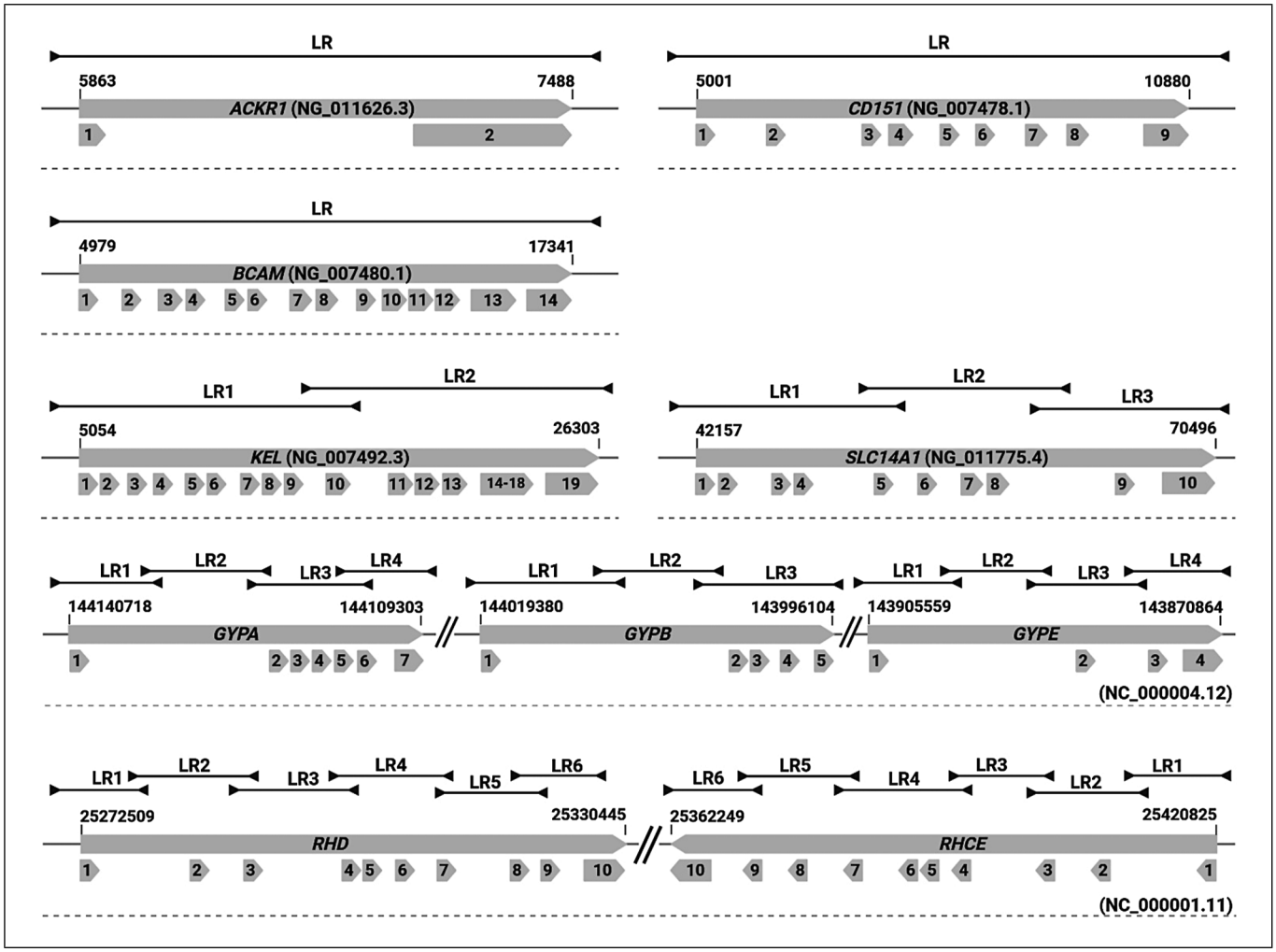

Fig. 2. Schematic illustration of the blood group genes and location of the LR-PCR amplicons with regard to the reference sequence. Blood group genes are depicted as grew arrows according to the orientation with their respective exons as grew arrows underneath numbered consecutively. The location on the reference sequence (shown in brackets behind gene name or on the right below the exons depicted for the MNS and the Rh genes) is represented by numbers left and right of each gene locus. The position of the LR-PCR amplicons is depicted by black arrows above the genes. Created in BioRender. Wörner, L. (2026) https://BioRender.com/lkqmd1e.

replicates each were sequenced using the MinION device with Flongle flow cells or MinION flow cells (reference), and the PromethION Solo 2 device with PromethION flow cells. Sequencing of the nine samples was performed on flow cells of different reuse cycles as referred to in Table 3. For comparability of sequencing results, LR-PCR was performed three times for each sample to generate a sufficient amount of LR-PCR products. The three libraries were prepared from one LR-PCR pool per blood group gene and sample.

Evaluation of Sequencing Data

Mean Q scores were calculated from the per-base Q scores calculated by the Dorado algorithm during basecalling in the MinKNOW software for each analyzed sample and compared between blood group genes. Additionally, the correlation between passed reads, pore availability, and repeated use of the minion flow cell was examined. Values of mean coverage with standard deviation and minimum coverage were determined after read mapping by using the Geneious Prime software. The mean coverage refers to the coverage averaged over all LR-PCR amplicons according to the number of generated LR-PCR amplicons for the different blood group genes. The minimum coverage refers to the nucleotide position showing the lowest number of reads for single LR-PCR amplicons. To evaluate the usability of the sequencing workflow on the three different sequencing device and flow cell combinations, mean Q scores and the minimum coverage of biological replicates were calculated as previously described and compared.

Table 2. Patient and donor samples with known gene variants of the analyzed blood group systems

Sample	Blood group gene	SNV	SNV zygosity	Amino acid change	dbSNP	Variant allele[a]
BGS_007*	*ACKR1*	c.125G>A c.265C>T c.298G>A	Heterozygous Heterozygous Heterozygous	p.Gly42Asp p.Arg89Cys p.Ala100Thr	rs12075 rs34599082 rs13962	*FY*01/*02W.01*
BGS_008*	*ACKR1*	c-67T>C c.125G>A	Homozygous Homozygous	(none) p.Gly42Asp	rs2814778 rs12075	*FY*02N.01*
BGS_009*	*ACKR1*	c.125G>A	Heterozygous	p.Gly42Asp	rs12075	*FY*01/*02*
BGS_012	*CD151*	c.346C	Homozygous	p.Gln116	–	*RAPH*01*
BGS_013	*CD151*	c.346C	Homozygous	p.Gln116	–	*RAPH*01*
BGS_015	*CD151*	c.346C>T	Homozygous	p.Gln116Ter	–	*RAPH*01 (unreported)* [25]
BGS_014	*BCAM*	c.133C>T c.230G	Homozygous Homozygous	p.Arg45Ter p.Arg77	rs199665533 rs28399653	*LU*02 (unreported)*
BGS_018	*BCAM*	c.230G c.1615A>G c.1340C>T c.1742A>T	Homozygous Heterozygous Heterozygous Heterozygous	p.Arg77 p.Thr539Ala p.Ser447Leu p.Gln581Leu	rs28399653 rs1135062 rs117737673 rs28399659	*LU*02.19/*02.-13*
BGS_019	*BCAM*	c.230G>A	Heterozygous	p.Arg77His	rs28399653	*LU*01/*02*
BGS_010	*KEL*	c.578C>T	Heterozygous	p.Thr193Met	rs8176058	*KEL*01/*02*
BGS_012	*KEL*	c.578C>T c.1268C>T c.1334C>T	Heterozygous Heterozygous Heterozygous	p.Thr193Met p.Ala423Val p.Ala445Val	rs8176058 rs61728831 rs150993188	*KEL*01.01/*02 (unreported)*
BGS_013	*KEL*	c.578C c.877C>T	Homozygous Heterozygous	p.Thr193 p.Arg293Trp	rs8176058 rs557358978	*KEL*02/*02.39*
BGS_002*	*SLC14A1*	c.838G c.334C>T	Homozygous Homozygous	p.Asp280 p.Gln112Ter	rs1058396 –	*JK*01 (unreported)*
BGS_003*	*SLC14A1*	c.130G>A c.588A>G c.838G>A	Heterozygous Homozygous Heterozygous	p.Glu44Lys p.Pro196= p.Asp280Asn	rs2298720 rs2298718 rs1058396	*JK*01/*02W.03*
BGS_010*	*SLC14A1*	c.838G>A	Heterozygous	p.Asp280Asn	rs1058396	*JK*01/*02*
BGS_011	*GYPA* *GYPB* 	c.59C>T c.71G>A c.72T>G c.143C>T	Heterozygous Heterozygous Heterozygous Homozygous	p.Ser20Leu p.Gly24Glu p.Gly24Glu p.Thr48Met	rs7682260 rs7687256 rs7658293 rs7683365	*GYPA*01/*02* *GYPB*03*
BGS_022	*GYPA* *GYPB* 	c.59C c.71G c.72T c.143C>T	Homozygous Homozygous Homozygous Heterozygous	p.Ser20 p.Gly24 p.Gly24 p.Thr48Met	rs7682260 rs7687256 rs7658293 rs7683365	*GYPA*01* *GYPB*03/*04*
BGS_023	*GYPA* *GYPB* 	c.59C c.71G c.72T c.143C>T	Homozygous Homozygous Homozygous Homozygous	p.Ser20 p.Gly24 p.Gly24 p.Thr48Met	rs7682260 rs7687256 rs7658293 rs7683365	*GYPA*01* *GYPB*03*
BGS_004*	*RHD*	c.48G	Homozygous	p.Trp16	rs772865539-	*RHD**01*
BGS_005*	*RHD*	c.8C>G c.809T>G	Heterozygous Heterozygous	p.Ser3Cys p.Val270Gly	rs144969459 rs121912763	*RHD*01W.3/*01W.1*
BGS_006**	*RHD*	c.739G>C	n.d.	p.Val247Leu	–	*RHD (unreported)* [26]
BGS_016	*RHD*	c.1170T>C c.1193A>T	n.d.	p.Leu390= p.glu398Val	rs1132772 rs45549244	*RHD*01W.41.0.1*
BGS_017	*RHCE*	c.48G>C c.254C>G c.1025C>T	Heterozygous Heterozygous Heterozygous	p.Trp16Cys p.Ala85Gl p.Thr342Ile	rs586178 rs57992529 rs1053374	*RHCE*01.02.01/*01.06.01*

Table 2 (continued)

Sample	Blood group gene	SNV	SNV zygosity	Amino acid change	dbSNP	Variant allele[a]
BGS_020	*RHCE*	c.307C	Homozygous	p.Pro103	rs676785	*RHCE*03*
		c.676G>C	Homozygous	p.Ala226Pro	rs609320	
BGS_021	*RHCE*	c.48G>C	Homozygous	p.Trp16Cys	rs586178	*RHCE*02*
		c.150C>T	Homozygous	p.Val50=	rs200955066	
		c.178C>A	Homozygous	p.Leu60Ile	rs181860403	
		c.201A>G	Homozygous	p.Ser67=	rs1053343	
		c.203A>G	Homozygous	p.Asn68Ser	rs1053344	
		c.307C>T	Homozygous	p.Pro103Ser	rs676785	
		c.676G	Homozygous	p.Ala226	rs609320	

[a]Variant allele according to ISBT allele tables (unreported: not reported in ISBT tables); the new RAPH allele was recently reported [25]; the *RHD* c.739G>C variant was described before [26] but was not reported as variant allele in the ISBT tables. *Samples were also used for comparison of ONT sequencing devices and flow cells. **Sample was used only for comparison of ONT sequencing devices and flow cells.

To confirm known genotypes, the variants that contribute to a particular allele and that were genotyped by other methods were evaluated.

Results

Sequencing Efficiency

To examine the efficiency of sequencing runs with sequencing libraries generated with the optimized protocol, the number of passed reads (reads that matched the filtering criteria set in the MinKNOW software) and the alteration regarding MinION flow cell reuse were analyzed. The proportion of passed reads within total reads was in the range of 23.6–85.0% (shown in online suppl. Table S1; for all online suppl. material, see https://doi.org/10.1159/000550249). A reduction in the number of passed reads was observed for repeated flow cell use on average from 15.4 ± 7.4 k reads on first time used flow cells (n = 3) to 2.9 ± 1.8 k reads on five times used flow cells (n = 6). As expected, a significant correlation was seen between the number of passed reads and the number of nanopores available for sequencing independent of the blood group gene and the flow cell reuse cycle (r = 0.986; p < 0.0001; n = 20). No significant change in the ratio of passed to failed reads for the different flow cell reuse cycles was observed.

Sequencing Data Quality

To examine sequencing data quality, the mean Q score of sequencing runs as well as the mean coverage of blood group genes and the (mean) minimum coverage of single LR-PCR amplicons were considered. As expected, the data from sequencing 93 amplicons revealed a highly significant correlation of the mean coverage and minimum coverage (r = 0.8792; p < 0.0001). Mean Q scores varied between but also within blood group genes with a median of 18.9 (range 17.4–20.3) corresponding to an error rate of 1.8%–0.9% (shown in Fig. 3). A tendency of higher Q scores was found for larger genes sequenced by more than one amplicon (r = 0.6289; p = 0.0948; n = 8). A variable mean coverage was observed between the different blood group genes, but also between the biological replicates of single blood group genes with a tendency of low coverage values to be more frequent in larger blood group genes covered by a higher number of LR-PCR amplicons (shown in Table 3). For the larger blood group genes covered by more than one amplicon the coverage of the different amplicons was unbalanced (shown exemplary for *KEL* and *RHD/CE* in Fig. 4). When considering the minimum coverage of single LR-PCR amplicons (n = 93) a median of 255 reads was observed. The LR-PCR amplicon 2 of all three samples sequenced for *KEL*, the *BCAM* LR-PCR amplicon in all three samples, the LR-PCR amplicons 2 of two samples and LR-PCR amplicon 3 of one sample sequenced for *GYPE*, the LR-PCR amplicon 5 of one *RHD* sample and the LR-PCR amplicons 2 of two samples and LR-PCR amplicon 6 of all three samples sequenced for the *RHCE* gene showed less than 50 reads.

Confirmation of Variant Alleles

For 24 of the 25 known variant alleles analyzed in 22 samples the nanopore sequencing data revealed the correct genotype (examples of read mapping to *BCAM* shown in online suppl. Fig. S1). Patient sample BGS_014 was included in this study because of a Lu(a-b-) phenotype. Nanopore sequencing of the *BCAM* gene revealed homozygosity for a novel null allele c.133C>T (p.Arg45Ter; rs199665533) and the mutation was confirmed by PCR-SSP (shown in online suppl. Fig. S2). The *GYPB* c.143C>T (*GYPB*03/GYPB*04*) heterozygosity for sample BGS_022

Table 3. Mean coverage over all LR-PCR amplicons of three samples per blood group gene analyzed on the MinION sequencing device

Blood group gene	Sample	Mean coverage±SD	Flow cell reuse cycle
ACKR1	BGS_007	10,750±203	6
	BGS_008	4,525±101	5
	BGS_009	5,763±109	4
CD151	BGS_012	1,033±39	2
	BGS_013	5,424±273	3
	BGS_015	2,850±110	n.d.
BCAM	BGS_014	350±20	3
	BGS_018	98±6	4
	BGS_019	24±2	5
KEL	BGS_010	222±61	5
	BGS_012	683±52	5
	BGS_013	4,506±119	1
SLC14A1	BGS_002	265±153	5
	BGS_003	577±456	5
	BGS_010	1,297±510	4
GYPA	BGS_011	328±146	4
	BGS_022	1,652±394	2
	BGS_023	842±193	4
GYPB	BGS_011	380±102	4
	BGS_022	1,963±347	2
	BGS_023	967±153	4
GYPE	BGS_011	206±76	4
	BGS_022	974±389	2
	BGS_023	487±205	4
RHD	BGS_004	243±74	n.d.
	BGS_005	1,016±223	1
	BGS_016	758±274	2
RHCE	BGS_017	591±551	n.d.
	BGS_021	425±360	1
	BGS_022	845±772	3

n.d., not determined.

could not be confirmed since analysis of the nanopore sequencing data showed 668 of 832 reads (0.803) for c.143T indicating homozygosity. In order to exclude a strand bias, the symmetric odds ratio (SOR) for this nucleotide position was calculated (Strand Odds Ratio test from GATK, version 4.1.8.1). A mild non-significant bias was found (SOR 1.042; Fisher Exact Test, $p = 0.274$). Furthermore, other mapping algorithms (Geneious, Bowtie2) revealed the same result. The DNA sample was re-typed for the c.143C>T variant by using TaqMan-PCR and heterozygosity was confirmed. For the samples BGS_011 and BGS_023, both homozygous c.143T, nanopore sequencing confirmed homozygosity with higher values for the T allele (0.93 and 0.91, respectively).

Application of the Nanopore Sequencing Protocol to Other ONT Sequencing Devices and Flow Cells

For the evaluation of further ONT sequencing devices and flow cells, we applied the sequencing protocol to the MinION device with Flongle flow cells and the PromethION with PromethION flow cells in addition to sequencing on the MinION device with MinION flow cells. A sub-analysis was performed for *ACKR1* (one LR-PCR amplicon), *SLC14A1* (three LR-PCR amplicons), and *RHD* (six LR-PCR amplicons) with three samples each. Mean Q scores calculated for each analyzed sample showed a median of 16.1 for the Flongle flow cell, 18.6 for the MinION flow cell and 19.0 for the PromethION flow cell. As expected, the numbers of generated passed reads were significantly different between the Flongle and MinION flow cells as well as the Flongle and PromethION flow cells while no significant difference was seen between the MinION and PromethION flow cell (shown in Fig. 5; online suppl. Table S2). Interestingly, the number of passed reads relative to the number of available pores was higher with the Flongle flow cells compared to PromethION flow cells (shown in Fig. 5). Analysis of the minimum coverage of the LR-PCR amplicons revealed significantly different values between the flow cells (shown in Fig. 5; online suppl. Table S2).

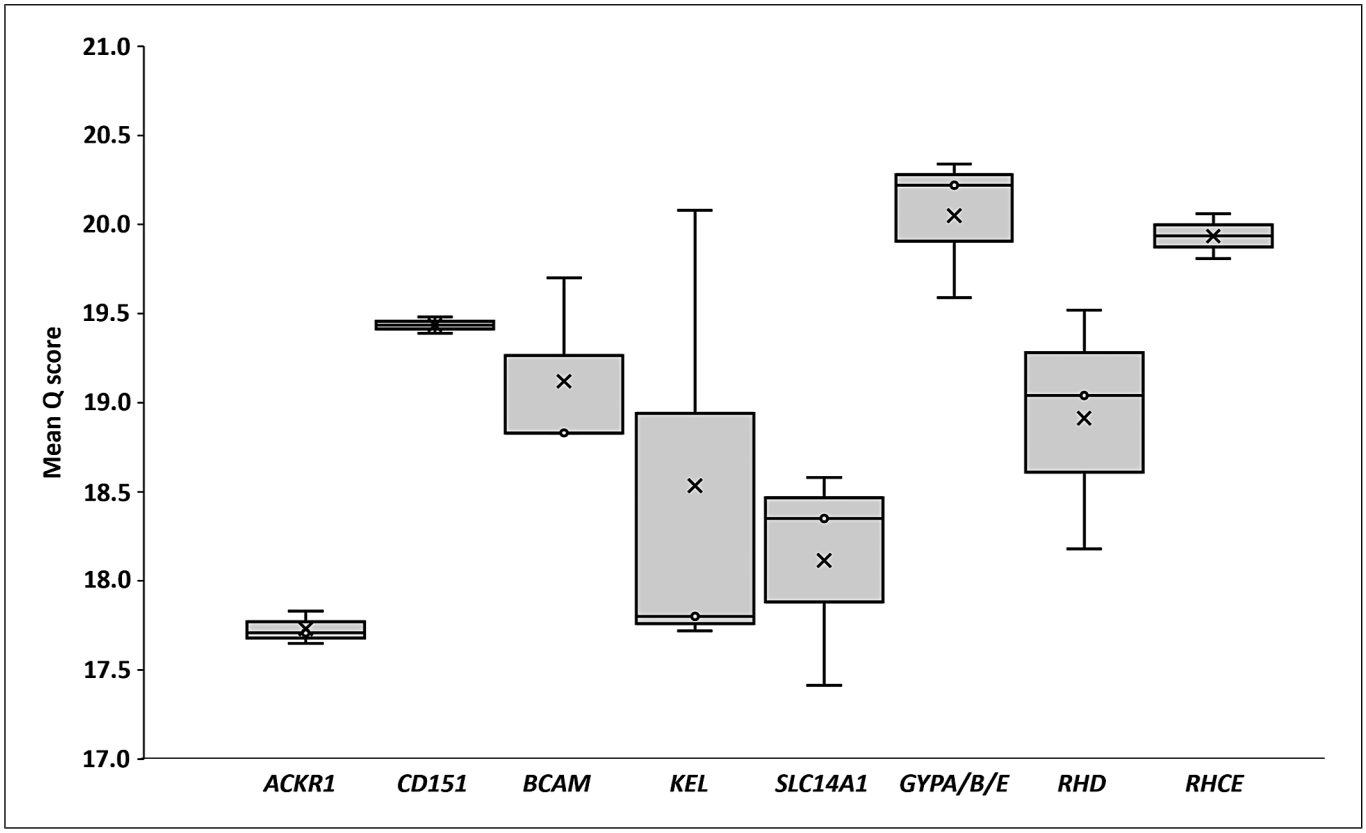

Fig. 3. Mean Q score over all samples and LR-PCR amplicons for each blood group gene. Shown are mean Q scores with $n = 3$ sequencing runs for each blood group gene, except for *CD151* and *RHCE*, each with $n = 2$. The median mean Q score is indicated by a line within the box, while the mean of the displayed values is indicated by ×.

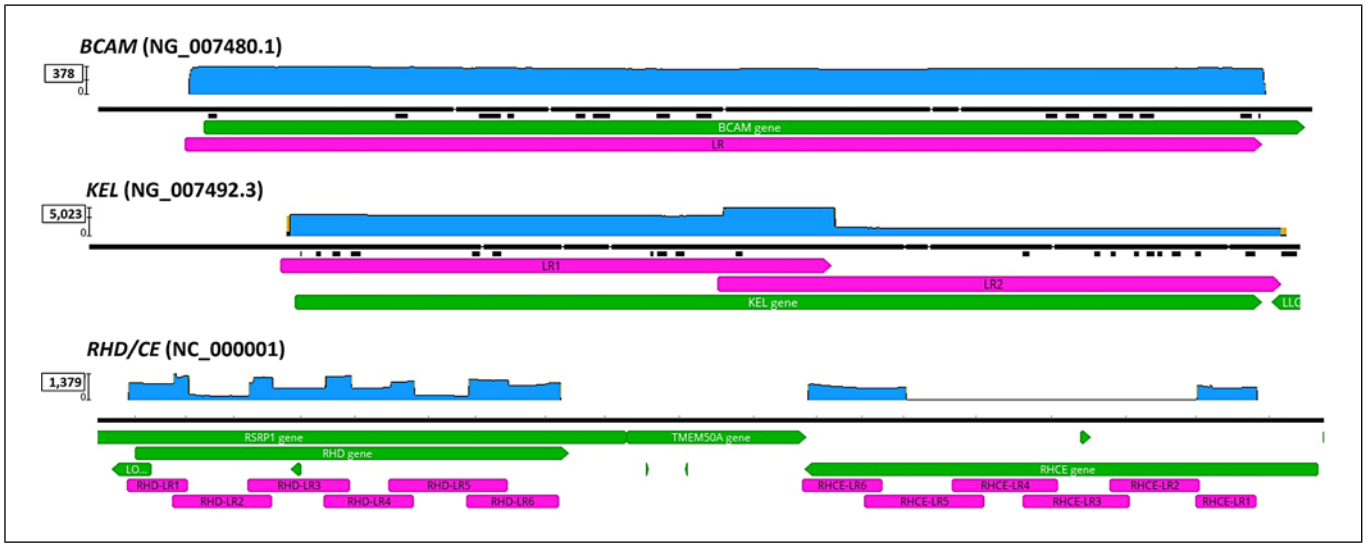

Fig. 4. Representative examples (BGS_014 for *BCAM*; BGS_13 for *KEL*; BGS_005 for *RHD*) of LR-PCR amplicon coverage following read mapping of sequencing data for genes of different sizes. *BCAM* was covered by one LR-PCR amplicon, *KEL* by two overlapping amplicons and *RHD* by six overlapping amplicons. Reads mapped to the respective reference sequence (shown in brackets behind gene name) are displayed in blue bars with the corresponding maximum coverage in boxes on the left side. Gene loci (green) and LR-PCR amplicon locations (magenta) are shown below. For the *RHD* sequencing alignment, the *RHCE* gene locus is also displayed to show the reads mapped to this homologous gene locus.

The PromethION flow cells produced the highest absolute minimum coverage values, however, with a high variability as indicated by the asymmetric distribution of the mean and median minimum coverage. The Flongle flow cells showed the lowest coverage with a high stability. The genotypes for all samples were correctly confirmed independent of the used sequencing device and flow cell.

Discussion

In this study, we introduce an optimized long-read nanopore sequencing protocol for single samples to be used for further investigations on blood group genes and for immunohematology case studies. So far, the protocol was established for ten blood group genes from seven systems. With a corresponding primer design for LR-PCR, the sequencing protocol can be adapted for further blood group genes or even for other genes of interest.

The nanopore sequencing technology generates data with a higher sequencing error compared to other technologies such as Illumina or PacBio [14]. The mean Q score for the R10.4.1 chemistry with R10.4.1 flow cells is considered to be greater than 20 [28]. We observed marginally lower Q scores in the range of 17.4–20.3 corresponding to an error rate of up to 1.8%. However, the correct variant detection was not affected by the error rate. With our optimized nanopore sequencing protocol the analysis of blood group genes is time- and cost-efficient. The analysis time of 6–7 h from blood sample to the result makes it feasible to implement the protocol in the routine work of molecular immunohematology laboratories. The generation of sequence data was limited to 2 h and a sufficient number of passed reads was obtained. As expected, with the repeated use of MinION flow cells the number of nanopores available for sequencing and the number of passed reads decreased significantly. A MinION flow cell could be used up to five times reliably with a sequencing time of 2 h per run. This should be considered when calculating the costs of the analysis per sample.

There are different parameters that can have an impact on Q scores such as the used basecalling algorithm, the used sequencing chemistry or the analyzed DNA sequence itself (e.g., homopolymeric regions are more error-prone). We rather exclude the basecalling algorithm and sequencing chemistry since both parameters were the same for all analyzed samples. The negative correlation between pore availability in the flow cell reuse cycles and corresponding mean Q scores indicated a decrease in pore health. The Q scores also differed between genes and biological replicates. The tendency of higher Q scores for larger genes sequenced by more than one amplicon could be due to lower Q scores in the first minutes of sequencing that stabilized at a higher level over time. We assume that the formation of the ionic environment for optimal sequencing requires some time. During the 2 h sequencing larger genes with a higher number of LR-PCR amplicons could have more bases sequenced with a stabilized Q score reducing the impact of sequenced bases of lower quality. Differences in Q scores could also be due to flow cell reuse and pore availability (pore health). We found negative correlation between pore availability in the flow cell reuse cycles and corresponding mean Q scores. In our study, we used both fresh and frozen blood samples, whereby the sample type had no impact on the Q score or sequencing result. As long as LR-PCR is successful, no pre-analytical aspects need to be taken into account.

The mean coverage varied between DNA samples and between the different gene loci. A variability in the coverage was also observed between the LR-PCR amplicons for the larger genes (e.g., mean coverage of *RHCE* LR-PCR amplicon 1 with 1,713 reads and LR-PCR amplicon 6 with 97 reads). This is due to several factors: (1) quality of the template DNA; (2) efficiency of LR-PCR; (3) GC content of target sequences; (4) nanopore availability on the flow cell and flow cell reuse cycle. The quantification of LR-PCR products and equimolar use for library preparation could circumvent the imbalanced sequencing data. However, we did not consider that in our protocol in order to reduce hands-on-time. For a highly accurate detection of alleles in diploid genomes using nanopore sequencing, a minimum coverage of 30× to 50× (15× to 25× per allele) should be considered. In the context of a higher sequencing error rate of ONT sequencing data compared to other methods, the determination of an unknown variant as genuine or artificial mainly depends on the coverage of the position of interest. In the majority of our analyses the sequencing time of 2 h was sufficient to achieve a coverage of >50×. We assume that the sequencing error rate has no impact on variant calling at a coverage of >50×. The definition of 50× minimum coverage as cut-off is reasonable, especially for the detection of unknown variants in an optimized protocol for routine diagnostics. However, even lower coverages (<10×) were reported for correct allele detection [29]. With the application of our sequencing protocol, the mean coverage was highly variable and low for some samples. The minimum coverage for some samples and genes was low as well. Nevertheless, we achieved the identification of the known gene variants in our sample set, e.g., we clearly identified the heterozygous BCAM allele *LU*01* (c.230G>A) in sample BGS_019 with 16 of only 26 reads and a minimum coverage of 0. In such cases, the coverage could be easily increased by extension of the sequencing time. LR-PCR amplicons with a minimum coverage below 50× will be taken into account in further optimization to ensure a

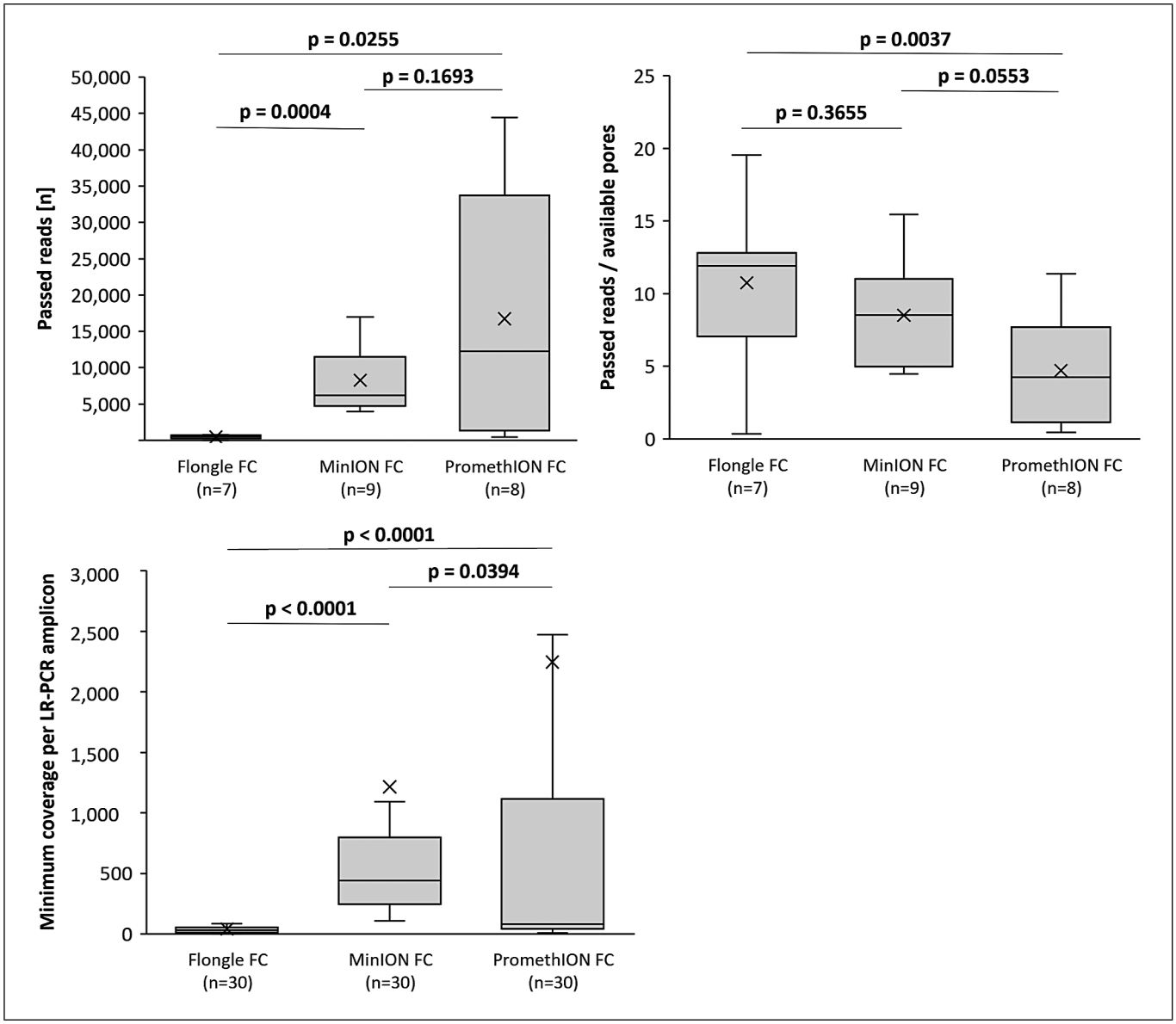

Fig. 5. Comparative analysis of sequencing data generated within 2 h sequencing on different ONT sequencing devices and flow cells (MinION with Flongle FC, MinION with MinION FC, PromethION with PromethION FC). *ACKR1*, *SLC14A1*, and *RHD* were sequenced in three samples each (9 runs total on each device). Left: total number of passed reads; middle: ratio of passed reads per available pores; right: minimum coverage per LR-PCR amplicon for all samples and blood group genes per device/flow cell (outliers are not displayed for scaling reasons of the *y*-axis). All values are listed in online supplementary Table S2. Median (line within the box) and mean (×) are indicated. *p* values were calculated based on all values using the following statistical tests: two-tailed *T* test for normally distributed samples (passed reads, passed reads/available pores); two-tailed Mann-Whitney-U test for non-normally distributed samples (minimum coverage).

sufficient coverage of every single nucleotide position. We like to point out that it is reasonable to consider both, minimum coverage and mean coverage to evaluate the amplicon performance. Based on our data of sequencing a total of 93 amplicons we found a highly significant correlation of the minimum coverage and mean coverage. A minimum coverage discrepant from the mean coverage can be explained by using reads of different length for read mapping since no further length filtering or trimming was performed other than described in the methods section.

All but one genotypes were determined correctly. Sample BGS_022, previously genotyped heterozygous *GYPB*03/*04* (c.143C>T) by PCR-SSP, revealed homozygosity for *GYPB*03* by nanopore sequencing. With 668 of 832 reads (0.803), the T allele was over-represented and the defined value for homozygosity (>0.8) was reached by a narrow margin showing a non-significant mild strand bias for

nucleotide position c.143. Two additional samples pre-typed c.143C>T heterozygous, were sequenced and confirmed the overrepresentation (>0.8) of the T allele (data not shown). An imbalanced PCR amplification of the *GYPB* alleles or preferential sequencing could be a reason as well as a genetic alteration such as a copy number variation of the sequence carrying the T allele. A larger number of pre-typed samples need to be sequenced in order to confirm the cut-off values in such cases. Additionally, the use of a specific variant calling tool for long-read sequences, such as DeepVariant, clair3 or the FreeBayes plugin for the Geneious Prime software could be beneficial in this context but also for variant determination in general. Another challenge is the specific sequence analysis of highly homologous genes such as *GYPA* and *GYPB* or *RHD* and *RHCE*. A specific primer design and amplification, presumably, has no impact on variant calling regardless of using only the targeted gene or including also the homologous gene(s). However, some unspecific amplification of the homologous genes cannot completely be excluded. We, therefore, recommend to include also the homologous genes as reference (shown in Fig. 4 for RHD and RHCE) in order to verify the mapping algorithm. Furthermore, rearrangements between homologous genes could be identified more easily.

During this study, 2 patient samples were referred to our laboratory for molecular investigation. For patient BGS_015 an alloantibody against a high prevalence antigen was suspected by serology. We performed panel sequencing of the blood group genes and identified a novel nonsense mutation in *CD151* (RAPH system) leading to a MER2-negative phenotype [25]. The *CD151* gene was included in the long-read nanopore sequencing protocol and we confirmed homozygosity for the nonsense mutation c.346C>T. Patient sample BGS_014 with a Lu(a-b-) phenotype revealed homozygosity for a novel null allele c.133C>T (p.Arg45Ter; rs199665533) of the *BCAM* gene.

Comparison of the ONT sequencing devices and flow cells showed that the sequencing protocol developed on the MinION sequencing device using MinION flow cells can also be used in combination with the Flongle flow cell and on the PromethION sequencing device using PromethION flow cells. The lower number of passed reads generated with the Flongle flow cell can be explained by the smaller sequencing capacity of this flow cell (564 nanopores) compared to the MinION flow cell (2,048 nanopores) or the PromethION flow cell (10,700 nanopores). However, with Flongle flow cells the number of passed reads relative to the number of available pores was higher compared to MinION or PromethION flow cells. The median minimum coverage produced by MinION flow cells was higher and less variable compared to PromethION flow cells. We speculate that under the defined conditions (library quality and concentration, 2 h sequencing, gene size and number of LR-PCR amplicons) the varying effectiveness of the flow cells is substantially affected by the number of reuse cycles of the used flow cell. With a prolonged sequencing time, the data output could be more robust. The Flongle flow cell with the sequencing time of 2 h is more suitable for smaller genes amplified by a single LR-PCR amplicon. Since the Flongle flow cells have more favorable costs, adjustment of the protocol to this single-use flow cell type by extending the sequencing time could be taken into consideration for sequencing larger genes. When sequencing time is limited, the MinION or PromethION flow cells should be preferentially used.

Compared to the nanopore sequencing approach, the use of other methods such as Sanger sequencing is significantly more laborious, time-consuming and costly. Moreover, with such a short-read sequencing technology, the analysis is usually focused on the coding exons, whereas, long-read sequencing data are suitable for the characterization of structural gene variants and haplotype phasing beyond exon boundaries. Whether the LR-PCR amplicon-based approach can be used for the characterization of hybrid genes will be considered in further studies. Nanopore adaptive sampling could be an alternative approach to generate long-read sequence data. However, some disadvantages compared to the amplicon-based method should be taken into account: (1) a significantly longer sequencing time is required to obtain a sufficient number of on-target reads; (2) depending on the power of the graphic processor unit (GPU) basecalling can also take several hours; (3) flow cells can only be used once. This is contradictory to the time and cost efficiency of the amplicon-based approach.

Immunohematology case studies are mostly initiated by the identification of an irregular antibody or unclear blood group phenotype by using serological methods. The molecular analysis of the blood group genes of interest at least supports or even enables the diagnosis. In conclusion, this work showed that the optimized long-read nanopore sequencing protocol is suitable to generate sequencing data for the reliable detection of variants in the blood group genes *ACKR1*, *CD151*, *BCAM*, *KEL*, *SLC14A1*, *GYPA*, *GYPB*, *GYPE*, *RHD*, and *RHCE*. The analyses were performed for single samples and single or few genes per analysis and we assume that implementation of additional blood group genes such as ABO is possible [22]. Depending on the size of a gene, primers for overlapping amplicons must be designed and the size of the amplicons should not exceed 16 kbp. With the limited hands-on-time and total analysis time of less than 7 h as well as the cost effectiveness, the workflow fulfills the requirements for implementation in laboratories of routine diagnosis. Amplicon-based nanopore sequencing is the method of choice for the rapid and complete analysis of single genes in single samples. Further developmental work in sequence data analysis is required for the application of the optimized

sequencing protocol to the routine setting in order to facilitate variant detection, allele calling and haplotype determination.

Statement of Ethics

The donors of this study gave written informed consent to use anonymized blood samples for research and development purposes. The Ethics Committee II of the Medical Faculty Mannheim, Heidelberg University, approved the use of anonymized blood samples for research and development purposes (Approval No. 2016-615N-MA).

Conflict of Interest Statement

The authors have no conflicts of interest to declare. Peter Bugert was a member of the journal's Editorial Board at the time of submission.

Funding Sources

This study was funded by the German Red Cross Blood Service Baden-Württemberg – Hessen gGmbH.

Author Contributions

L.W. performed the laboratory work, evaluated the data, and wrote the manuscript; G.R. and X.M. assisted in laboratory work and reviewed the manuscript; and P.B. designed the study, evaluated the data, and wrote the manuscript.

Data Availability Statement

All data generated or analyzed during this study are included in this article and its supplementary material files. Further inquiries can be directed to the corresponding author.

References

1 Westhoff CM. Blood group genotyping. Blood. 2019;133(17):1814–20. https://doi.org/10.1182/blood-2018-11-833954
2 Denomme GA. Prospects for the provision of genotyped blood for transfusion. Br J Haematol. 2013;163(1):3–9. https://doi.org/10.1111/bjh.12476
3 Reid ME. Transfusion in the age of molecular diagnostics. Hematol Am Soc Hematol Educ Program. 2009;2009(1):171–7. https://doi.org/10.1182/asheducation-2009.1.171
4 Wagner FF. Molecular testing in transfusion medicine. Expert Opin Med Diagn. 2010;4(5):411–28. https://doi.org/10.1517/17530059.2010.506509
5 Fichou Y, Férec C. NGS and blood group systems: state of the art and perspectives. Transfus Clin Biol. 2017;24(3):240–4. https://doi.org/10.1016/j.tracli.2017.06.002
6 Denomme GA, Van Oene M. High-throughput multiplex single-nucleotide polymorphism analysis for red cell and platelet antigen genotypes. Transfusion. 2005;45(5):660–6. https://doi.org/10.1111/j.1537-2995.2005.04365.x
7 Hopp K, Weber K, Bellissimo D, Johnson ST, Pietz B. High-throughput red blood cell antigen genotyping using a nanofluidic real-time polymerase chain reaction platform. Transfusion. 2010;50(1):40–6. https://doi.org/10.1111/j.1537-2995.2009.02377.x
8 Hashmi G, Shariff T, Seul M, Vissavajjhala P, Hue-Roye K, Charles-Pierre D, et al. A flexible array format for large-scale, rapid blood group DNA typing. Transfusion. 2005;45(5):680–8. https://doi.org/10.1111/j.1537-2995.2005.04362.x
9 Beiboer SH, Wieringa-Jelsma T, Maaskant-Van Wijk PA, van der Schoot CE, van Zwieten R, Roos D, et al. Rapid genotyping of blood group antigens by multiplex polymerase chain reaction and DNA microarray hybridization. Transfusion. 2005;45(5):667–79. https://doi.org/10.1111/j.1537-2995.2005.04319.x
10 Karpasitou K, Drago F, Crespiatico L, Paccapelo C, Truglio F, Frison S, et al. Blood group genotyping for Jk(a)/Jk(b), Fy(a)/Fy(b), S/s, K/k, Kp(a)/Kp(b), Js(a)/Js(b), Co(a)/Co(b), and Lu(a)/Lu(b) with microarray beads. Transfusion. 2008;48(3):505–12. https://doi.org/10.1111/j.1537-2995.2007.01555.x
11 Tilley L, Grimsley S. Is next generation sequencing the future of blood group testing? Transfus Apher Sci. 2014;50(2):183–8. https://doi.org/10.1016/j.transci.2014.02.013
12 Fürst D, Tsamadou C, Neuchel C, Schrezenmeier H, Mytilineos J, Weinstock C. Next-generation sequencing technologies in blood group typing. Transfus Med Hemother. 2020;47(1):4–13. https://doi.org/10.1159/000504765
13 Warburton PE, Sebra RP. Long-read DNA sequencing: recent advances and remaining challenges. Annu Rev Genomics Hum Genet. 2023;24:109–32. https://doi.org/10.1146/annurev-genom-101722-103045
14 Thun GA, Gueuning M, Mattle-Greminger M. Long-read sequencing in blood group genetics. Transfus Med Hemother. 2023;50(3):184–97. https://doi.org/10.1159/000530652
15 Browning SR, Browning BL. Haplotype phasing: existing methods and new developments. Nat Rev Genet. 2011;12(10):703–14. https://doi.org/10.1038/nrg3054
16 Tewhey R, Bansal V, Torkamani A, Topol EJ, Schork NJ. The importance of phase information for human genomics. Nat Rev Genet. 2011;12(3):215–23. https://doi.org/10.1038/nrg2950
17 Pollard MO, Gurdasani D, Mentzer AJ, Porter T, Sandhu MS. Long reads: their purpose and place. Hum Mol Genet. 2018;27(R2):R234–41. https://doi.org/10.1093/hmg/ddy177
18 Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. Nanopore sequencing technology, bioinformatics and applications. Nat Biotechnol. 2021;39(11):1348–65. https://doi.org/10.1038/s41587-021-01108-x
19 Lane WJ, Gleadall NS, Aeschlimann J, Vege S, Sanchis-Juan A, Stephens J, et al. Multiple GYPB gene deletions associated with the U– phenotype in those of African ancestry. Transfusion. 2020;60(6):1294–307. https://doi.org/10.1111/trf.15839
20 Thun GA, Gueuning M, Sigurdardottir S, Meyer E, Gourri E, Schneider L, et al. Novel regulatory variant in ABO intronic RUNX1 binding site inducing A(3) phenotype. Vox Sang. 2024;119(4):377–82. https://doi.org/10.1111/vox.13580
21 Tounsi WA, Lenis VP, Tammi SM, Sainio S, Haimila K, Avent ND, et al. Rh blood group D antigen genotyping using a portable nanopore-based sequencing device: proof of principle. Clin Chem. 2022;68(9):1196–201. https://doi.org/10.1093/clinchem/hvac075
22 Gueuning M, Thun GA, Wittig M, Galati AL, Meyer S, Trost N, et al. Haplotype sequence collection of ABO blood group alleles by long-read sequencing reveals putative A1-diagnostic variants. Blood Adv. 2023;7(6):878–92. https://doi.org/10.1182/bloodadvances.2022007133
23 Gueuning M, Thun GA, Trost N, Schneider L, Sigurdardottir S, Engström C, et al. Resolving genotype-phenotype discrepancies of the kidd blood group system using long-read nanopore sequencing. Biomedicines. 2024;12(1):225. https://doi.org/10.3390/biomedicines12010225
24 Srivastava K, Khil PP, Sippert E, Volkova E, Dekker JP, Rios M, et al. ACKR1 alleles at 5.6 kb in a well-characterized renewable US Food and Drug Administration (FDA) reference panel for standardization of blood group genotyping. J Mol Diagn, 2020;22(10):1272–9. https://doi.org/10.1016/j.jmoldx.2020.06.014

25 Worner L, Rink G, Stürtzel A, Rothenberger-Mürb S, Seyboth S, Küthe-Rieger J, et al. A novel null allele of the RAPH blood group system. Transfusion. 2025;65(8):E31–3. https://doi.org/10.1111/trf.18304
26 Lyu H, Wang K, Feng Z, Xu T. A novel RHD allele caused by c.767 C>T mutation was identified in a Chinese individual. Transfusion. 2024; 64(5):E18–20. https://doi.org/10.1111/trf.17820
27 Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018; 34(18):3094–100. https://doi.org/10.1093/bioinformatics/bty191
28 Sanderson ND, Kapel N, Rodger G, Webster H, Lipworth S, Street TL, et al. Comparison of R9.4.1/Kit10 and R10/Kit12 Oxford Nanopore flowcells and chemistries in bacterial genome reconstruction. Microb Genom. 2023;9(1): mgen000910. https://doi.org/10.1099/mgen.0.000910
29 Lamb HJ, Hayes BJ, Randhawa IAS, Nguyen LT, Ross EM. Genomic prediction using low-coverage portable Nanopore sequencing. PLoS One. 2021;16(12): e0261274. https://doi.org/10.1371/journal.pone.0261274