

Evaluation of Single Nucleotide Variants in Intron 1 of the ABO Gene as Diagnostic Markers for the A₁ Blood Group

Peter Bugert Gabi Rink Harald Klüter

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

Keywords

ABO genotyping · ABO blood group · Diagnostic marker · Sensitivity · Specificity

Abstract

Introduction: The molecular diagnosis of the A₁ blood group is based on the exclusion of ABO gene variants causing blood groups A₂, B, or O. A specific genetic marker for the A₁ blood group is still missing. Recently, long-read ABO sequencing revealed four sequence variations in intron 1 as promising markers for the ABO*A1 allele. Here, we evaluated the diagnostic values of the 4 variants in blood donors with regular and weak A phenotypes and genotypes. **Methods:** ABO phenotype data (A, B, AB, or O) were taken from the blood donor files. The ABO genotypes (low resolution) were known from a previous study and included the variants c.261delG, c.802G>A, c.803G>C, and c.1061delC. ABO variant alleles (ABO*AW.06, *AW.08, *AW.09, *AW.13, *AW.30, and *A3.02) were identified in weak A donors by sequencing the ABO exons before. For genotyping of the ABO intron 1 variants rs532436, rs1554760445, rs507666, and rs2519093, we applied TaqMan assays with endpoint fluorescence detection according to a standard protocol. Genotypes of the variants were compared with the ABO phenotype and genotype. Evaluation of diagnostic performance included sensitivity, specificity, positive (PPV), and negative predictive value (NPV). **Results:** In 1,330 blood donors with regular ABO phenotypes and genotypes, the intron 1 variants were significantly associated with the proposed A₁ blood group. In 15 donors, we found discrepancies to the genotype of at least

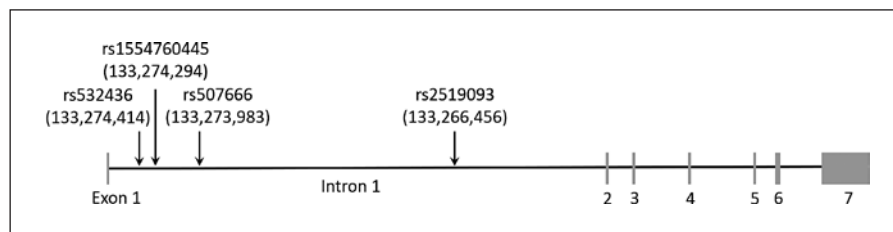
one of the 4 variants. For the diagnosis of the ABO*A1 allele, the variants showed 98.79–99.48% sensitivity, 99.66–99.81% specificity, 98.80–99.31% PPV, and 99.66–99.86% NPV. Regarding the A phenotype, the diagnostic values were 99.02–99.41% sensitivity, 99.63–99.76% specificity, 99.41–99.61% PPV, and 99.39–99.63% NPV. The *A1 marker allele of all intron 1 variants was also associated with the *AW.06, *AW.13, and *AW.30 variants. Samples with *AW.08, *AW.09, and *A3.02 variants lacked this association. **Conclusion:** The ABO intron 1 variants revealed significant association with the ABO*A1 allele and the A phenotype. However, the intron 1 genotype does not exclude variant alleles causing weak A phenotypes. With the introduction of reliable tag, single nucleotide variants for the A₁, A₂, B, and O blood groups and the genotyping instead of phenotyping of the ABO blood group are getting more feasible on a routine basis.

© 2023 The Author(s).
Published by S. Karger AG, Basel

Introduction

The ABO blood group system (ISBT 001) includes the two major antigens A and B. The further antigens A₁, B₁, and A₁B₁ are less common. The major phenotypes A, B, AB, and O are defined by the presence or absence of the A and B antigens [1]. The A phenotype is divided into A₁ and A₂ with approximately 5 times more A epitopes per red cell for A₁. Furthermore, a significantly lower number of epitopes per red cell is the most important cause of additional phenotypes such as weak A and weak B.

Fig. 1. Genomic structure of the *ABO* gene on chromosome 9 and location of rs532436, rs1554760445, rs507666, and rs2519093. Chromosomal position of the variations is given according to the reference genome GRCh38.p13.



The *ABO* gene located on the long arm of human chromosome 9 encodes the ABO blood group system. Compared to the reference allele *ABO***A1.01*, most of the sequence variations were identified in exon 6 and 7, encoding the catalytic domain of the glycosyltransferase [2, 3]. According to the ISBT allele table (version 1.2, released on October 21, 2017), 84 alleles are associated with the *A*₁, *A*₂, and weak *A* phenotypes [4]. Most of the **A2* alleles for the *A*₂ phenotype have the deletion of C at position c.1061 in common. The 49 listed alleles for the B and weak B phenotypes are defined by specific sequence variations such as c.803G>C compared to the *ABO***A1.01* reference allele. The O phenotype is most commonly caused by *ABO* alleles with deletion of G at position c.261. More than 40 deletional *ABO***O.01* alleles with sequence variations in addition to c.261delG have been described. The most common non-deletional allele *ABO***O.02* is characterized by the missense variation c.802G>A in exon 7. In summary, genotyping of the most relevant ABO sequence variations c.261delG, c.802G>A, c.803G>C, and c.1061delC enabled diagnosis of the *A*₂, B, and O phenotypes [5]. Exclusion of these proposed the **A1* allele for the *A*₁ phenotype. This low-resolution genotype does not consider rare variants with mutations at other positions.

In a recent study by Gueuning et al. [6], long-read sequencing of the entire *ABO* gene led to the identification of four sequence variations in intron 1 significantly associated with the *ABO***A1* allele. These variants represent promising diagnostic targets for the *A*₁ blood group. Here, we genotyped the four intron 1 variants in a representative sample of 1,330 blood donors with known ABO phenotype and genotyped for **A2*, **B*, **O.01*, and **O.02* alleles. The diagnostic values were calculated for each variant with regard to the identification of the **A1* allele and the A blood group.

Materials and Methods

Blood Donors

The geographical origin of blood donors of our transfusion service is the southwestern part of Germany. A DNA bank was established as a representative sample of our blood donor cohort and encompasses 1,330 blood donors with a mean age of 46.8 ± 15.3 years (range 18.0–68.8 years) and 1:1 gender distribution. DNA was isolated from buffy coats of CPD-anticoagulated whole blood

donations using a commercial system (FlexiGene DNA Kit; Qiagen, Hilden, Germany). Data about ABO blood groups were taken from the blood donor files, and all donors had regular ABO blood groups (A, B, AB, or O) without discrimination of the *A*₁ and *A*₂ phenotypes. The low-resolution *ABO* genotypes were determined in a previous study [7] and included the variants c.261delG (for deletional **O.01* alleles), c.802G>A (for the non-deletional O alleles **O.02*), c.803G>C (for **B* alleles), and c.1061delC (for **A2* alleles). The *A*₁, *A*₂, B, and O phenotypes were deduced from the low-resolution genotype. Furthermore, DNA samples from blood donors and patients with weak A phenotype and *ABO* variant alleles previously identified by sequencing exons 1 to 7 were included. All donors gave written consent to use their blood samples for research purposes. The anonymous DNA bank was approved by the Ethics Committee of the Heidelberg University, Medical Faculty Mannheim.

Genotyping of *ABO* Intron 1 Variants

According to the database (dbSNP), the four variants are located in intron 1 of the *ABO* gene (Fig. 1). The rs1554760445 is a CA>T insertion/deletion (INDEL) variant, whereas rs532436 (G>A), rs507666 (G>A), and rs2519093 (C>T) are single nucleotide variants (SNVs). Allelic discrimination was achieved by end-point fluorescence detection using custom TaqMan™ SNP genotyping assays (assay IDs: AN9HYTF for rs2519093; ANAAP46 for rs532436; ANCFJP3 for rs507666; and ANDKEAZ for rs1554760445) designed and produced by Thermo Fisher Scientific (Darmstadt, Germany). For all variants, the minor allele was the proposed tag for the *ABO***A1* allele.

Results

Allelic discrimination was achieved for all variants and enabled determination of the genotypes in most of the 1,330 samples (Fig. 2). Genotyping failed for rs532436 in 3 samples, for rs1554760445 in 4 samples, for rs507666 in 8 samples, and for rs2519093 in 3 samples. The minor allele frequencies were 0.2163 for rs532436, 0.2177 for rs1554760445, 0.2179 for rs507666, and 0.218 for rs2519093.

For most of the samples, the genotype of the intron 1 variants corresponded to the low-resolution *ABO* genotype and the phenotype. The proposed **A1* markers showed false positive or false negative results in 15 samples (Table 1). In 9 samples, at least one of the intron 1 variants was wrong, and in 6 samples, all 4 markers showed a genotype not corresponding to the proposed *ABO* genotype. With regard to the diagnosis of blood

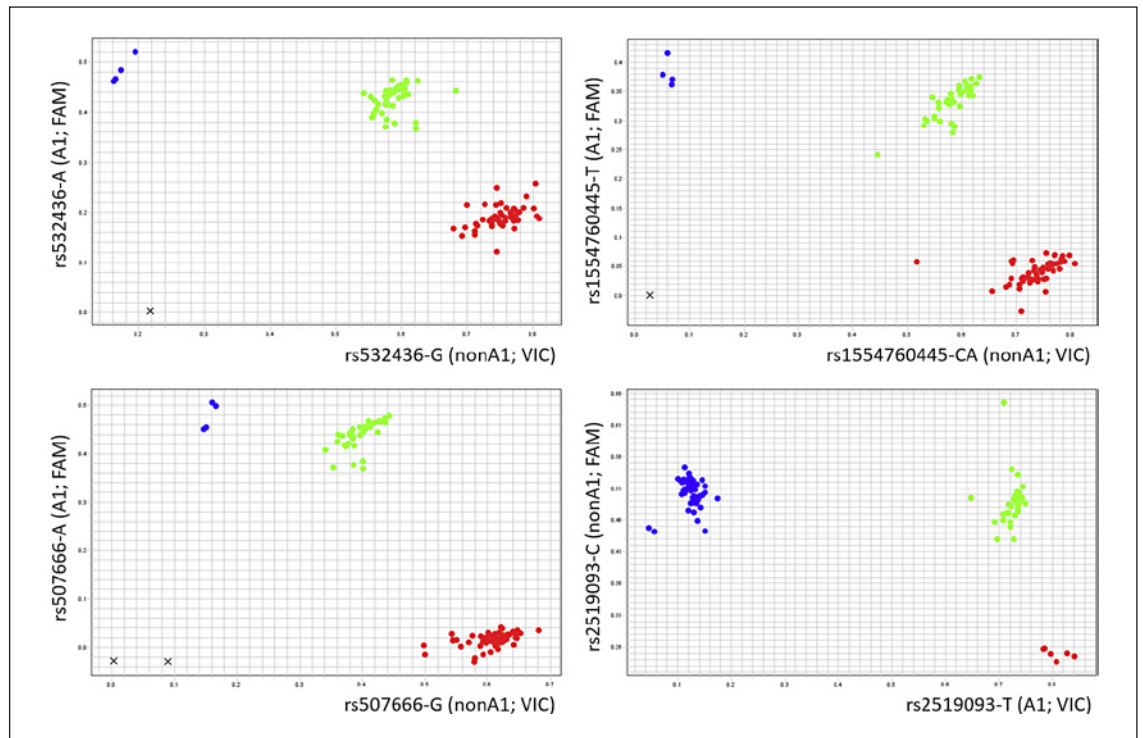


Fig. 2. Representative results from genotyping of 95 DNA samples for rs532436, rs1554760445, rs507666, and rs2519093 by using TaqMan™ assays. Allelic discrimination was achieved by endpoint fluorescence detection of FAM- and VIC-labeled allele-specific probes. All samples group into homozygous for VIC allele (red dots), homozygous for FAM allele (blue dots), and heterozygous for both (green dots).

Table 1. In 15 samples, the genotypes of the intron 1 variants revealed discrepancy to the ABO phenotype or low-resolution genotype

Sample	Phenotype*	Genotype	Deduced phenotype [#]	rs532436	rs1554760445	rs507666	rs2519093
M01E10	AB	*A1/*B	A1B	nonA1 (FN)	A1/nonA1	A1 (FP)	A1/nonA1
M02B09	A	*O.01*A2	A2	A1/nonA1 (FP)	A1/nonA1 (FP)	A1/nonA1 (FP)	nonA1
M03F08	A	*O.01/*A1	A1	nonA1 (FN)	nonA1 (FN)	nonA1 (FN)	nonA1 (FN)
M05C04	A	*O.01/*A1	A1	nonA1 (FN)	nonA1 (FN)	nonA1 (FN)	nonA1 (FN)
M05G09	A	*A2/*A2	A ₂	A1/nonA1 (FP)	A1/nonA1 (FP)	A1/nonA1 (FP)	A1/nonA1 (FP)
M07C07	A	*A1	A ₁	A1/nonA1 (FN)	A1/nonA1 (FN)	A1/nonA1 (FN)	A1
M07G07	A	*A1	A ₁	A1/nonA1 (FN)	A1/nonA1 (FN)	A1/nonA1 (FN)	A1
M10A06	A	*O.02/*A1	A ₁	A1/nonA1	A1/nonA1	A1/nonA1	A1 (FP)
M10F09	A	*O.01/*A1	A1	nonA1 (FN)	nonA1 (FN)	A1/nonA1	A1/nonA1
M11B01	O	*O.01	O	A1/nonA1 (FP)	A1/nonA1 (FP)	A1/nonA1 (FP)	A1/nonA1 (FP)
M11D12	O	*O.01	O	A1/nonA1 (FP)	A1/nonA1 (FP)	A1/nonA1 (FP)	A1/nonA1 (FP)
M12E12	A	*O.02/*A1	A ₁	A1/nonA1	A1/nonA1	A1/nonA1	A1 (FP)
M13D08	AB	*A/*B	A1B	nonA1 (FN)	nonA1 (FN)	nonA1 (FN)	nonA1 (FN)
M13G10	A	*O.01/*A1	A ₁	A1/nonA1	A1/nonA1	A1/nonA1	A1 (FP)
M14G12	A	*O.02/*A1	A ₁	A1/nonA1	A1/nonA1	A1/nonA1	A1 (FP)

* ABO phenotype from the donor files without discrimination of A₁ and A₂ phenotype. [#] ABO phenotype deduced from the low-resolution genotype; gray background indicates false ABO genotype proposed from the genotype of the intron 1 variants: FN, false negative; FP, false positive for the A1 allele; in 8 samples given in bold letters the genotypes of the intron 1 variants indicated a wrong A phenotype.

group A, the intron 1 variants failed in 8 of the 15 samples (Table 1). In two samples with blood group O, all intron 1 variants were heterozygous indicating blood group A₁. In three samples with blood group A, the absence of the

*A1 allele at all intron variants revealed false negative results. In all 15 samples sequencing of the ABO exons, 1 to 7 confirmed the supposed presence or absence of the *A1.01 allele.

Table 2. Calculation of the diagnostic values for the intron 1 variants with regard to the diagnosis of the *ABO**A1 allele and the A phenotype

	Intron 1 variants			
	rs532436	rs1554760445	rs507666	rs2519093
Diagnosis of the <i>ABO</i> *A1 allele				
True positive, <i>n</i>	570	573	571	574
False positive, <i>n</i>	4	4	5	7
True negative, <i>n</i>	2,073	2,069	2,063	2,070
False negative, <i>n</i>	7	6	5	3
Sensitivity, %	98.79	98.96	99.13	99.48
Specificity, %	99.81	99.81	99.76	99.66
PPV, %	99.30	99.31	99.13	98.80
NPV, %	99.66	99.71	99.76	99.86
<i>Diagnosis of the A phenotype</i>				
True positive, <i>n</i>	505	507	504	505
False positive, <i>n</i>	2	2	3	2
True negative, <i>n</i>	815	813	812	816
False negative, <i>n</i>	5	4	3	4
Sensitivity, %	99.02	99.22	99.41	99.21
Specificity, %	99.76	99.75	99.63	99.76
PPV, %	99.61	99.61	99.41	99.61
NPV, %	99.39	99.51	99.63	99.51

Sensitivity: true positive/(true positive + false negative); specificity: true negative/(true negative + false positive); positive predictive value (PPV): true positive/(true positive + false positive); negative predictive value (NPV): true negative/(true negative + false negative).

Table 3. Association of intron 1 variants with *ABO* variant alleles causing the weak A phenotype

Variant*	Samples, <i>n</i>	Genotype	rs532436	rs1554760445	rs507666	rs2519093
*AW.06	12	*O.01/*AW.06	A1/nonA1	A1/nonA1	A1/nonA1	A1/nonA1
	1	*AW.06/*B.01	A1/nonA1	A1/nonA1	A1/nonA1	A1/nonA1
*AW.08	10	*O.01/*AW.08	nonA1	nonA1	nonA1	nonA1
*AW.09	4	*O.01/*AW.09	nonA1	nonA1	nonA1	nonA1
*AW.13	11	*O.01/*AW.13	A1/nonA1	A1/nonA1	A1/nonA1	A1/nonA1
	4	*AW.13/*B.01	A1/nonA1	A1/nonA1	A1/nonA1	A1/nonA1
	2	*AW.13/*B.01	nonA1	nonA1	nonA1	nonA1
*AW.30	25	*O.01/*AW.30	A1/nonA1	A1/nonA1	A1/nonA1	A1/nonA1
	1	*O.02/*AW.30	A1/nonA1	A1/nonA1	A1/nonA1	A1/nonA1
	1	*AW.30/*B.01	A1/nonA1	A1/nonA1	A1/nonA1	A1/nonA1
*A3.02	5	*O.01/*A3.02	nonA1	nonA1	nonA1	nonA1
	1	*A3.02/*B.01	nonA1	nonA1	nonA1	nonA1

ABO variants defined by mutations: c.502C>G (*AW.06), c.488C>T and c.526C>G (*AW.08), c.46G>A (*AW.09), c.2T>C (*AW.13), c.646T>A (*AW.30 and others), and c.829G>A (*A3.02).

The diagnostic values including sensitivity, specificity, positive (PPV), and negative predictive value (NPV) were calculated for the intron 1 variants (Table 2). Overall, the diagnostic values for the 4 intron 1 variants were similar. For the diagnosis of the *ABO**A1 allele, the intron 1 variants showed a sensitivity of 98.79–99.48%, a specificity of 99.66–99.81%, a PPV of 98.80–99.31%, and a NPV of 99.66–99.86%. For the diagnosis of the A phenotype, we found 99.02–99.41% sensitivity, 99.63–99.76% specificity,

99.41–99.61% PPV, and 99.39–99.63% NPV. None of the intron 1 variants showed complete association with the *ABO**A1 allele or the A phenotype.

Furthermore, 77 samples with known *ABO* variant alleles causing a weak A phenotype were genotyped for the intron 1 variants (Table 3). The *AW.06 and *AW.30 variants were associated with the *A1 allele of all intron 1 variants in 13 and 27 samples, respectively. For the *AW.13 variant, we found an association with the *A1 al-

lele of the intron 1 variants in 15 of 17 samples. Two samples with **AW.13* and all samples with **AW.08*, **AW.09*, and **A3.02* variants lacked this association.

Discussion

The identification of sequence characteristics of the *ABO* gene as specific markers for the A, B, and O blood groups enabled the molecular diagnosis of ABO phenotypes. Because of the lack of a specific marker, the A₁ phenotype was diagnosed by exclusion of alleles encoding A₂, B, or O phenotypes. Based on the genotyping of the 4 major variants c.261delG, c.802G>A, c.803G>C, and c.1061delC, a genotype-phenotype correlation was achieved for 99.7% of the samples [7]. However, the significant diversity of the *ABO* gene and the increasing number of mutations associated with aberrant blood group phenotypes impeded the molecular diagnosis. The introduction of additional DNA markers, especially for the A₁ blood group, is important to increase the specificity and sensitivity of the molecular diagnosis of the ABO blood group.

The rs507666 and rs2519093 have been used as a marker for the A₁ blood group in studies on the association of the *ABO* locus with ICAM-1 plasma levels and risk of venous thrombosis, respectively [8, 9]. Further analyses linking the lead SNVs to *ABO* allele groups have not been undertaken in these studies. Thus, proof of the diagnostic value of the markers as tags for blood group A₁ was lacking. Long-read sequencing of the *ABO* locus uncovered putatively *ABO***A1* diagnostic variants in the intron 1 region [6]. Three SNVs (rs532436, rs1554760445, and rs507666) are located in close proximity within a 431 bp region at the 5' end of intron 1. The rs2519093 is located more than 7.5 kbp downstream from rs507666. In many of the samples with discordant results, we observed a linkage of the first 3 SNVs, i.e., concordant genotypes, whereas rs2519093 revealed a different genotype. This could indicate recombination events in intron 1 as a cause of discordant genotyping results. A multi-ethnic validation approach indicated that the three SNV-based variants rs532436, rs507666, and rs2519093 significantly tagged the *ABO***A1* allele across ethnicities. The INDEL variant rs1554760445 specifically tagged the *ABO***A1.01* allele in populations outside Africa. The sensitivities ranged from 97.55% (rs507666) to 97.99% (rs2519093) and the specificities from 99.41% (rs532436) to 99.72% (rs1554760445). The geographic origin of the blood donors in our study is Southwestern Germany; however, the ethnicity is not known in detail. It can be assumed that the majority of the donors are European, and only a minor proportion is of African descent.

We evaluated the performance of the intron 1 variants for the diagnosis of the *ABO***A1* allele. All 4 variants

showed a strong but not complete association with the **A1* allele in 1,330 samples with regular ABO phenotypes. For the diagnosis of the *ABO***A1* allele, we found sensitivities ranging from 98.79% (rs532436) to 99.48% (rs2519093) and specificities from 99.66% (rs2519093) to 99.81% (rs532436), similar to the values reported previously [6]. For the diagnosis of the A blood group, the rs507666 showed the highest sensitivity (99.41%), whereas, rs532436 and rs2519093 had the highest specificity (99.76%). In 6 of the 1,330 samples (0.45%), none of the 4 intron 1 variants showed correct tagging of the **A1* allele with 3 false negative and 3 false positive results. The combination of all 4 intron variants did not increase the diagnostic performance. According to our data from blood donors with regular ABO phenotypes, the genotyping of the intron 1 variants would lead to a wrong phenotype prediction in 0.6% of the donors. To increase the diagnostic reliability, ABO genotyping should include the markers for **A2*, **B*, **O.01*, and **O.02* alleles in addition to the **A1* markers in intron 1.

Further limitation of ABO genotyping is given by the large number of variant alleles. Most of these variants are very rare (<0.01%), and the association with the intron 1 variants is not known. In our study, we found that the rare *ABO***AW.06*, **AW.13*, and **AW.30* alleles causing a weak A phenotype were associated with the **A1* markers, i.e., the intron 1 variants indicated a regular A₁ phenotype. Other variant alleles such as **AW.08*, **AW.09*, and **A3.02* lacked this association, i.e., intron 1 genotyping indicated the absence of a regular A₁ phenotype. The prevalence of variant alleles could be different in the populations. The **AW.09* allele tagged by c.49G>A (rs55917063) is rare in Europeans (0.02%) but frequent in Africans (2%) [10]. The frequency of the **AW.08* allele tagged by c.488C>T (rs55756402) is 0.1% in Europeans and South Asians but not observed in Africans. Thus, depending on the population, the association of the **A1* markers with rare *ABO* gene variants could significantly delimit the correct diagnosis of the A₁ blood group.

Blood group genotyping projects usually include numerous clinically relevant blood group systems but not ABO. This is because ABO phenotyping is cost effective and well established in routine labs. The diagnosis of the ABO blood group based on genotyping is getting more relevant since molecular techniques are developed toward high-throughput screening. For example, the genotype screening project at our institute includes 37 blood groups as well as platelet and neutrophil antigens [11]. The inclusion of *ABO* genotyping with tag SNVs could be cost and time effective. Thereby, the availability of reliable diagnostic markers for the main phenotypes is an important prerequisite. The genotyping platform for blood donors introduced by the Blood Transfusion Genomics Consortium (BGC) comprised a large number of

determinants of blood group, platelet, and leukocyte antigens as well as variants in genes relevant to antigen expression [12]. The genotyping results demonstrated concordance of 99.91% with antigen typing data for 28 blood group antigens including ABO. In 7 of 7,449 samples (0.09%), the ABO genotyping results were discordant with the phenotype.

In conclusion, with the introduction of reliable molecular markers, genotyping instead of phenotyping of the ABO blood group is getting more feasible also on a routine basis. This could be reached by a combination of SNVs that tag the A₁, A₂, B, and O phenotypes and the exclusion of ABO variants encoding aberrant phenotypes. However, as long as discordance of genotypes is observed, the antibody-based ABO phenotyping remains the only method to ensure ABO-matched blood transfusion.

Statement of Ethics

This study was conducted ethically in accordance with the World Medical Association Declaration of Helsinki. The donors gave written consent to provide blood samples for research purposes. The anonymous DNA bank was approved by the Ethics Committee of the Heidelberg University, Medical Faculty Mannheim (Ref. 87/04).

References

- 1 Storry JR, Olsson ML. The ABO blood group system revisited: a review and update. *Immunohematology*. 2009;25(2):48–59.
- 2 Yamamoto F, McNeill PD, Hakomori S. Genomic organization of human histo-blood group ABO genes. *Glycobiology*. 1995;5(1): 51–8.
- 3 Chester MA, Olsson ML. The ABO blood group gene: a locus of considerable genetic diversity. *Transfus Med Rev*. 2001;15(3):177–200.
- 4 The International Society of Blood Transfusion. *ABO (ISBT 001) blood group allele table*.
- 5 Gassner C, Schmarda A, Nussbaumer W, Schonitzer D. ABO glycosyltransferase genotyping by polymerase chain reaction using sequence-specific primers. *Blood*. 1996;88(5): 1852–6.
- 6 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*. 2022 Sep 21: bloodadvances.2022007133.
- 7 Bugert P, Rink G, Kemp K, Klüter H. Blood group ABO genotyping in paternity testing. *Transfus Med Hemother*. 2012 Jun;39(3): 182–6.
- 8 Paré G, Chasman DI, Kellogg M, Zee RY, Rifai N, Badola S, et al. Novel association of ABO histo-blood group antigen with soluble ICAM-1: results of a genome-wide association study of 6,578 women. *PLoS Genet*. 2008 Jul 4;4(7):e1000118.
- 9 Goumidi L, Thibord F, Wiggins KL, Li-Gao R, Brown MR, van Hylckama Vlieg A, et al. Association between ABO haplotypes and the risk of venous thrombosis: impact on disease risk estimation. *Blood*. 2021 Apr 29;137(17): 2394–402.
- 10 *gnomAD database* (accessed December 06, 2022).
- 11 Portegys J, Rink G, Bloos P, Scharberg EA, Klüter H, Bugert P. Towards a regional registry of extended typed blood donors: molecular typing for blood group, platelet and granulocyte antigens. *Transfus Med Hemother*. 2018 Oct;45(5):331–40.
- 12 Gleadall NS, Veldhuisen B, Gollub J, Butterworth AS, Ord J, Penkett CJ, et al. Development and validation of a universal blood donor genotyping platform: a multinational prospective study. *Blood Adv*. 2020 Aug 11; 4(15):3495–506.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

No funding was received for this study.

Author Contributions

Peter Bugert coordinated the study, analyzed the data, and wrote the manuscript. Gabi Rink performed genotyping and summarized the data. Harald Klüter coordinated the study and wrote the manuscript.

Data Availability Statement

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