Review Article

Next generation sequencing and blood group genotyping: a narrative review

Agnieszka Orzińska

Department of Hematological and Transfusion Immunology, Institute of Hematology and Transfusion Medicine, Warsaw, Poland
Correspondence to: Agnieszka Orzińska. Institute of Hematology and Transfusion Medicine, 14 Indiry Gandhi Street, 02-776 Warsaw, Poland. Email: aorzinska@ihit.waw.pl.

Objective: The work is a narrative review on blood group antigen (BGA) genotyping using next generation sequencing (NGS).

Background: The introduction of genotyping technologies for BGA typing has changed the practice of transfusion medicine facilitating the identification of antigen-negative blood donors compatible with immunized patients against common antigens and extended BGA matching for transfusion dependent patients.

Methods: The review presents data from the articles published as a full-length papers in PubMed and reported in conference abstracts.

Conclusions: The massive parallel sequencing has opened up new possibilities in the immunohematological field giving the opportunity to screen numbers of donors and alleles as well as determine any nucleotide variants in the tested sequences, quantify sequences for estimation of copy number variation (CNV) or chimerism and map of haplotype and *cis/trans* location. NGS technology is used for blood donor BGA screening, to puzzle out the background of complicated serological cases and for non-invasive prenatal testing (NIPT) of fetal antigens. The correlation between BGA NGS and phenotyping/genotyping results varies from 92% to 99%. The current state of BGA NGS requires a deep standardization and validation of all the steps including bioinformatic tools as well as ethical guidelines before it can be implemented in routine BGA diagnostics.

Keywords: Next generation sequencing (NGS); blood group antigens (BGAs)

Received: 10 May 2021; Accepted: 26 September 2021; Published online: 26 October 2021.
doi: 10.21037/aob-21-39
View this article at: https://dx.doi.org/10.21037/aob-21-39

Introduction

The introduction of genotyping technologies for blood group antigen (BGA) typing has changed the practice of transfusion medicine (1-3). Molecular methods were at first a complement to serological tests, then became an alternative or the only solution. They have a great many useful applications which include the blood grouping of recently transfused blood recipients, autoimmune hemolytic anemia patients or during some monoclonal therapies, individuals with rare blood antigens, inconclusive serological results or multi-specific complex antibodies and finally, non-invasive fetal antigen genotyping and paternal *RHD* zygosity (4-9). Also, genotyping facilitates the identification of antigen-negative blood donors compatible with immunized patients against common antigens and extended BGA matching for transfusion dependent patients (10,11). But the currently routinely-used genotyping workflows determine a limited number of the most probable single nucleotide variations (SNV) using predefined primers/probes and frequently not solving complex cases which then require further laborious and time-consuming investigations with the classical sequencing of many exons (11-15). Thus, BGA genotyping workflow is evolving from SNV typing for a single patient to fully automatic high-throughput massive parallel sequencing of a whole

genome or several exomes or targeted regions for patients and donor screening (11,16). Among the techniques from the genetic toolbox, sequencing has been an expensive and time-consuming solution used as a last resort but it has proved to be highly helpful in investigating the causes of the most atypical serological reactions and learning about underlying phenomena (17-19). Undoubtedly, there are several advantages of next generation sequencing (NGS) that overcome the limitations of current BGA genotyping platforms (20,21). The crucial issue is the revealing of the direct sequence of the entire gene or genes, exons, the promotors or intronic regions as well as the associated genes in one test in a relatively short time for a different number of individuals. Also, NGS technology determines any nucleotide variants in the tested sequences, quantifies sequences for estimation of copy number variation (CNV) or chimerism and maps of haplotype and *cis/trans* location (22,23). Moreover, the capacity of NGS is cost-effective and useful for high throughput screening (24). In the field of transfusion medicine there are nearly 50 genes known to be involved in expressing BGAs that are an ideal goal for the high throughput and capacity of massive parallel sequencing.

Sequencing of the genotype of BGAs has evolved in the last 30 years, since the discovery and sequencing of genes encoding the basic blood group systems (25). Over these years Sanger sequencing has been applied to describe single fragment/exon or gene encoding BGAs in single cases with discrepant serological and genetic results (1-3,17,18). The massive parallel sequencing has opened up new possibilities in the immunohematological field giving the opportunity to screen numbers of donors and alleles. The high throughput and huge capacity was firstly applied to HLA-typing and now it is a routine technique in these diagnostics (24). A characteristic feature of this method is the sequencing of millions of relatively short reads in one run. Researchers decide for themselves how deeply to sequence and how many fragments to read from how many people at the same time, which allows for a balance between multiplexing the desired features in research with BGA panels for hundreds of individuals (e.g., donor screening) or testing millions of reads with extremely high sensitivity for several patients (e.g., in non-invasive fetal testing from the maternal plasma of immunized pregnant women) or extracting BGA alleles from existing WES/WGS databases (9,16,26,27). Different sequencing kits offer a high flexibility fitting the workflow to the aims and budgets of studies. This study discusses the progress made in BGA typing using NGS technology. The review presents data from the articles published as a full-length papers in PubMed and reported in conference abstracts. We present the following article in accordance with the Narrative Review reporting checklist (available at https://aob.amegroups.com/article/view/10.21037/aob-21-39/rc).

NGS platforms applied for BGA genotyping

Currently, two NGS platforms are used for BGA typing known as the second generation: Sequencing-by-synthesis offered by Illumina and semiconductor sequencing on Ion Torrent from Thermofisher (28,29). The first technology is similar to the classical solution with the detection of fluorescence of each incorporated nucleotide. A flowcell coated with clusters of oligo-adaptors is flushed by a sequenced library. After capture by the adapters, cluster synthesis is generated and sequencing is followed by the detection of dye-labelled nucleotides. The solution makes sequencing from both ends possible (paired-end sequencing). The capacity of the system depends on the instrument used for testing (e.g., for the MiSeq it is about 45 milion paired reads per run). The time of each run is reported to be the main limitation of the system. Also, the error rate is estimated to be about 0.1% mainly due to non-specific substitutions. But the Illumina system is reported as being the main technology applied in BGA sequencing by various researchers.

The second semiconductor Ion Torrent method detects changes in electric potentials generated during the incorporation of complementary nucleotides. Sequencing beads coated with clonal templates are loaded into microwells covering a chip. The synthesis of each nucleotide releases protons changing the potential that is measured during each cycle. The main limitation is associated with about 1% error rate in the detection of changes during the sequencing of homopolymer fragments. But, on the other hand, such manners of detection take only a few hours per run. The capacity of the system is about 200–400 bp reads with a depth of 0.4–5.5 million reads. The solution is dedicated for panel sequencing and it was the first to be used for massive screening of regions of interest (ROI) encoding clinically relevant BGAs for blood donors.

BGA NGS strategies

Since 2011, when the first report of applying NGS for *RHD* typing was published, many NGS-based BGA typings

Table 1 Examples of NGS strategies applied in the field of blood group (BG) antigen genotyping

NGS strategy	Tested blood group system, authors (ref)
NGS panels	ABO, Lang *et al.* (30)
	ABO, Fox *et al.* (31)
	ABO, Wu *et al.* (32)
	18 BG, Fichou *et al.* (33)
	17 BG, Fichou *et al.* (34)
	6 BG/HPA, Orzinska *et al.* (35)
	28 BG, Schoeman *et al.* (36)
	43 genes for 36 BG, 6 HPA, 5 HNA, Roulis *et al.* (37)
	Rh, Dezan *et al.* (38)
	Rh, El Wafi *et al.* (39)
	Rh, Tounsi *et al.* (40)
	Rh, Wheeler *et al.* (41)
	Rh, Dinardo *et al.* (42)
	Kell, Duffy, Kidd Dinardo *et al.* (43)
	Rare BG, Jakobsen *et al.* (44)
	Rare BG, Boccoz *et al.* (45)
	HPA, Vorholt *et al.* (46)
WES	36 BGs, Schoeman *et al.* (47)
	38 BGAs in 12BGs /22 HPA, Lane *et al.* (48)
	Rh, Lopez (49)
	Rh, Chou *et al.* (50)
	Rh, Stef *et al.* (51)
	Kell, Duffy, Kidd, Montemayor *et al.* (52)
	Kell, Millard (53)
	AUG, Millard *et al.* (54)
	In(Lu), Fraser *et al.* (55)
	Cromer Vrignaud (56)
	MAM Thorton (57)
	PEL, Azouzi *et al.* (58)
	CTL2 Vrignaud C (59)
WGS	45 BGs, Lane *et al.* (60)
	38 BGs/HPA, Lane *et al.* (61)
	9 BGAs/DRB1, Paganini *et al.* (62)
	Rare BGs, Lane *et al.* (63)
	MNS, Lane *et al.* (64)
	Rh, Halls *et al.* (65)
	Rh, Chang *et al.* (66)
Third generation long-read NGS	*SLC14A1,* Montemayor *et al.* (52)
	GYPE, *GYPB*, *GYPA,* Lane *et al.* (64)
	ACKR1, Fichou *et al.* (67)
	ACKR1, Srivastava *et al.* (68)
	ABO, Lang *et al.* (30)

NGS, next generation sequencing; WES, whole exome sequencing; WGS, whole genome sequencing; HPA, human platelet antigens.

have been described presenting a wide range of efforts to adjust and implement this new technology into routine immunohematological diagnostics (see *Table 1*) (69). Currently, four NGS strategies are applied for BGA genotyping often combined with platelet antigen (HPA) and even neutrophil antigen (HNA) genotyping, also useful in transfusion medicine.

The most common is targeted NGS focused on a single locus or employing a designed panel for several regions/genes/exons of interest encoding BGAs/HPAs. Fichou *et al.* were the first to prove that NGS using a customized panel for the Ion Torrent platform makes genotyping of 18 regions encoding BGAs for four individuals possible in a single run (33). Similarly, in our work we applied a panel for sequencing ROIs identifying six blood groups and five platelet antigen SNVs in 55 cases and due to an extremely high depth of reads determining natural chimerism in one donor (35). Also Jakobsen *et al.* used a panel designed for 15 alleles encoding rare BGAs not included in routine commercial assays (44). Dezan *et al.* who applied the separate *RHD*/*RHCE* libraries identified *RH* variants in 35 SCD patients clarifying their Rh phenotypes and improving their transfusion support since in 62% of the cases Rh typings were classified inaccurately by serology (38). Also Tounsi *et al.* focused on *RHD* and *RHAG* sequencing but using long-product PCR reconstructed haplotypes by identifying an intronic SNV pattern linked to the DcE haplotype but not to DCe, DCE or Dce (40). They described two reference sequences of the whole *RHD*. El Wafi *et al.*, also focused on the *RHD* gene, revealing new intronic SNVs affecting mRNA splicing in 9 RhD negative patients (39). Wheeler *et al.* presented a report on the identification of *RH* structural variants using the NGS panel on the Illumina platform (41). Wu *et al.* identified novel *ABO* variants as well as *ABO* structural variants and microchimerism using an *ABO* targeted panel for 24

Table 2 Examples of novel findings in the field of blood group (BG) antigen genotyping using NGS

Blood group system/gene	Novel findings	Reference
45 BG genes, 6 HPA genes	Novel *GYPA*, *A4GALT*, *CR1*, *CD109* variants	Lane *et al.* (60)
ABO	5 novel structural *ABO* variants	Fox *et al.* (31)
17 BG genes	Novel *ACHE*, *ART4* variants	Fichou *et al.* (34)
ABO (exons 6 and 7)	Novel 287 *ABO* variants	Lang *et al.* (30)
6 BG genes, 5 HPA antigens	Detection of natural chimerism	Orzinska *et al.* (35)
RHD	6 novel *RHD* variants	El Wafi *et al.* (39)
28 BG genes, *KLF1*, *GATA1*	Revealing unknown variants	Schoeman *et al.* (36)
15 rare BG antigens	Revealing rare variants	Jakobsen *et al.* (44)
RHD, *RHAG*	Identification of intronic SNV pattern link to DcE or DCe, DCE, Dce haplotypes	Tounsi *et al.* (40)
ABO	Identifying 3 novel variants and detection of 5% chimerism	Wu *et al.* (32)
36 BG genes, *KLF1*, *GATA1*	152 novel variants in *ABO*, *RH*, *KN*, *IN*, *MNS*, *CD59* genes	Schoeman *et al.* (47)
RH	*RHCE* and *RHD* alleles	Wheeler *et al.* (41)
42 BG genes	Identification of 12 novel variants	Montemayor *et al.* (52)
P1/XG	Revealing rs311103 associated with Xg(a-)phenotype	Lane *et al.* (48)
	Revealing 4 SNVs related to P1- negative phenotype	
HPA	Novel ITGB3 and GP1BA variants, noncoding variants	Vorholt *et al.* (46)
MNS	*GYP*B* variants	Lane *et al.* (64)
Rh	Novel *RH* variants	Chang *et al.* (66)
9 BG antigens	Novel *DO*, *IN*, *JK* variants	Paganini *et al.* (62)
Duffy	*ACKR1* haplotypes	Fichou *et al.* (67)
Kell, Duffy, Kidd	Novel *SLC14A1* variant	Dinardo *et al.* (43)
Kidd	Novel *SLC14A1* variant	Montemayor *et al.* (52)

NGS, next generation sequencing; HPA, human platelet antigens.

patients with discrepant serological results (32). Boccoz *et al.* demonstrated fast and complete antigen screening using a customized panel for 9 BGs in order to identify rare blood units by testing 95 donors in a single run of the Illumina platform (45). Recently Roulis *et al.* have presented a comprehensive panel for targeted exome sequencing of BGA, HPA and HNA genes correctly predicting genotypes for 99.5% of BGA and 100% HPA and HNA alleles (37). Also Vorholt *et al.* have described a HPA panel for high-throughput screening to build up HPA donor registry and provide matched blood products (46).

The above studies show the wide range of applications of this panel method in the field of blood grouping and how NGS technology has complemented basic knowledge in the context of immunohematology (see *Table 2*). The panels were precisely designed to SNV regions, exons and whole genes and also detected non-coding or flanking fragments known to interfere with the expression of some BGAs. Targeted NGS was focused on clinically significant immunogenic antigens such as: ABO, Rh but also on the rare antigens Xg(a-), P1. Various groups were tested: donors with defined phenotypes/genotypes, individuals with discrepant serological results and transfusion-dependent patients. The Ion Torrent platform with ion chemistry was the one most often used in this NGS strategy but sequencing-by-synthesis on the Illumina system was also employed successfully. Depending on the aim of the study, modified initial PCR protocols were used with long-product amplification, separating libraries or sequence specific primers for highly homological genes as well as different

Table 3 Non-invasive prenatal testing of fetal BGA/HPA genotypes using NGS technology from plasma DNA of pregnant women

BGS/HPA panel	Tested group, NGS platform	Authors (Ref)
Kell	3 patients, Illumina	Rieneck *et al.* (70)
6 BGS/4HPA/8 markers	4 patients, Ion Torrent	Wienzek-Lischka *et al.* (71)
HPA-1	2 patients, Ion Torrent	Orzińska *et al.* (72)
6 BGS/5HPA/4 markers	13 patients, Ion Torrent	Orzińska *et al.* (73)
6 BGS/4 HPA/15 markers	>200 patients, Ion Torrent, Illumina	Wienzek-Lischka *et al.* (74)
21 BGA+HPA/17 markers	Illumina	O'Brien *et al.* (75)
HPA-1	26 patients, Illumina	Nogués *et al.* (76)
RhD	8 patients, Illumina	Takahashi *et al.* (77)
ABO	26 patients, Illumina	Rieneck *et al.* (78)

NGS, next generation sequencing; HPA, human platelet antigens.

amplicon sizes, read depths and numbers of tested samples per NGS run.

The main advantage of the panel strategy is simplifying the bioinformatic analysis of limited NGS data focused only on SNVs and regions associated with the expressing BGAs. The researchers applied different bioinformatic tools for manual data processing, alignment, variant calling or copy number analysis and additional optimizing calculations for the interpretation of the genotype.

The BGA genotypes predicted using NGS panels had high correlation to serology and conventional genetic tests. However, often antigens encoded by homologous genes (*RHD/CE*, *GYPA/B*, Rg-Ch and *ABO*) were excluded from the final analysis. Targeted NGS of even a small number of cases revealed the presence of unknown SNVs and structural variants in coding or noncoding ROIs (see *Table 2*). Also, due to its high read depth, targeted NGS circumvented some misalignments in the structural variation analysis of highly homological regions, facilitated CNV analysis and was reported to determine a natural microchimerism. On the other hand, the strategy proved to have a huge capacity for high-throughput screening of a number of regions encoding clinically significant antigens in a number of individuals. The best example is Lang *et al.*'s report describing NGS of *ABO* exons 6 and 7 together with HLA that revealed novel 287 *ABO* variants in 1,693,287 bone marrow donors (30).

The targeted NGS strategy is also applied in non-invasive prenatal testing (NIPT) of fetal antigen genotype (see *Table 3*). A deep quantitative analysis of both maternal and fetal—paternally inherited sequences from cell-free plasma DNA of pregnant women enables an estimation of feto-maternal chimerism and the risk of fetal/neonatal disease in alloimmunized cases. There are two NIPT protocols—the first involves sequencing only ROIs encoding the incompatible antigen for testing several pregnant women with antibodies of the same specificity; the second—uses a wide panel of ROIs encoding the most clinically significant antigens together with additional fetal markers for testing pregnant women with different antibodies and confirming target-negative results in one experiment.

The second NGS strategy is based on previously generated whole genome sequencing (WGS) data filtered in the context of BGAs/HPAs. Möller *et al.* showed the possibilities of retrospective analysis of 36 blood group systems from a 1000 Genome database of over 2,500 individuals, obtained from classical sequencing, revealing 1,241 non-synonymous variants and creating an *Erythrogene* database of the complete coding regions of 43 blood group related genes and finding a novel *ABO* variant frequent in individuals of African origin (79,80). Lane *et al.* presented firstly an analysis of 45 blood group and 6 HPA related genes for one individual from the MedSeqProject WGS database proving the NGS utility in the field of transfusion medicine (60). Then, using automated WGS-based antigen typing software (bloodTyper), they filtered 38 BGAs and 22 HPA antigens for a few hundred cases from the INTERVAL database with 99% accuracy and later revealed SNVs correlating with Xg(a-) and P1 negative phenotypes in over 100 individuals from the MedSeqProject database (61,63). All the above published reports proved the utility of WGS-based data for BGA prediction. Untargeted sequencing of the whole genome produces the largest data, but short reads with low read depth may lead to misalignment of

homologous fragments. On the other hand, WGS data gives a full landscape of coding, noncoding and regulatory regions that may have an impact on the final BGA phenotype and is useful for solving complex cases (62,64-66). In this NGS strategy an automated tools for WGS analysis is necessary for extracting a number of BGA genotypes in a number of individuals. And a novel algorithm, called RHtyper, for *RH* variation characteristics from WGS data was proposed by Chang *et al.* (66).

The third WES-based NGS strategy makes the prediction of BGAs/HPAs from coding regions of existing individual data possible. Fox *et al.* presented the usefulness of this sort of NGS for ABO genotyping in over 6,000 WES-typed cases revealing novel structural variants (31). Schoeman *et al.* employed a commercial WES-panel for predicting 28 BGAs in 28 blood donors using manual analysis of CNVs of exons to identify homological *RH* and *GYP* genes (36). Later, from WES-data of indigenous Australians, they determined the complete BG profile of this ethnic cohort revealing 152 novel variants (47). Chou *et al.* proved the utility of WES analysis for *RH* genotyping to improve red blood cell matching in SCD patients (50). Lane *et al.* achieved 100% accuracy in BG analysis using the automated WES-based bloodTyper software (48). The WES strategy is very useful for resolving the genetic background of orphan and novel antigens or even new blood group systems such as SARA, Mi(a), k, In(Lu), CORS, DEL, AnWj, At(a), Vel, AUG, KANNO, Xg, CTL2, MAM, PEL, EMM (49,53-59,81-88). WES is cheaper and faster than WGS and there are commercial exome panels available. WES gives data on exons and intronic regions near exon boundaries and in the context of BGA genotyping only P1, Xg(a) and ABO intron 1 are now known to have SNVs not captured by WES but interfering with the presence of BGAs (20,63). The read depth and throughput in the WES strategy are higher than in WGS, but short reads may be misaligned to homological genes and also CNV analysis is more complex, however, WES produces small data files facilitating bioinformatic work and data storage. Recently, Lane *et al.* have proposed a flexible software called bloodTyper, adapted to analyze data from WES data as well as BGA panels (48). Also, Montemayor *et al.* have just provided DTM-Tools—an open-source software for Kell, Duffy and Kidd blood group interpretation from an existing 1,018 ES data files generated for the ClinSeq cohort (52). The newest long-read single molecule sequencing technique, called third generation, helps to overcome the limitation of short reads where phasing of associated polymorphisms generates complicated bioinformatic analysis. The solution is very useful for the determination of reference BGA haplotypes and the confirmation of novel findings (see *Tables 1*,*2*). The technology does not require preparation of any library but works directly from genomic DNA. Here, there are two technical solutions for the detection of sequencing based on fluorescence or electric current. In Pacific Bioscience the dye-labelled nucleotides emit signals in real time of incorporation allowing up to 45 kb of single read (89). The second solution—Oxford Nanopore sequencing is based on passing of DNA strands through protein nanopores embedded in a membrane where special electric sensors connected to the pores measure the change in electric current made by altering the conformation of the pores (90). The technique is prone to concentration of DNA strands and homopolymer fragments. However, the main limitation of single molecule sequencing is the high percentage of indel errors that requires correcting through deeper sequencing and also sophisticated bioinformatic analysis.

Genetic background of BGAs and NGS

The utility of short-read systems for BGA genotyping depends on the kind of mutation responsible for a wide landscape of antigen variety. Firstly, genetic variants encoding BGAs mainly belong to SNVs exchanging one nucleotide and converting an amino acid (1,91). The change is crucial for expressing an antithetical epitope and triggering an antibody response in individuals that lack the opposite structure in their red cell membranes. Also, one or several single changes are met in genes encoding enzymes relevant for the synthesis of polysaccharide antigens; altering their reactivity and substrate specificity, leading to e.g., a different pattern of ABO forward/reverse typing. SNVs may occur in coding as well as splice site regions, regulatory elements and promotors (92). Such mutations lead to premature stop codons; disturbing splicing or turning off the transcription and finally causing serologically null phenotypes directly in peptide chains or indirectly through inactive enzymes (e.g., *ABO*O.01*). Also, there are deletions of short or long fragments (i.e., Vel and RhD negativity) and conversions of highly homological genes (i.e., *RHD/RHCE*, *GYPA/B/E*) leading to hybrid variants that express new antigens or a lack of some (partial D). Diverse molecular backgrounds of BGAs lead to different strategies to analyze NGS results. All the reports of NGS of SNV-based antigens have a high accuracy independent of

the NGS platforms and designed panels used (see *Table 1*). But, as mentioned above, the phasing of several SNVs that are located far away in the presence of the second allele and the mapping of short reads to the correct reference in hybrid variants, formed by *RHD/RHCE*, *GYPA/GYPB* and *ABO*, are challenging and demand sophisticated bioinformatic analysis (31,62,65,66). Also, the correct primer design to the most conservative regions is crucial for omitting allele dropouts and falsely antigen-negative results (22,23).

Applications of BGA NGS

In the field of immunohematology NGS technology is used for several applications but mainly for blood donor screening, to puzzle out the background of complicated serological cases and for NIPT of fetal antigens.

Short-read NGS is useful mainly for BGA panel genotyping for donor screening in order to extend the donor registry for immunized and transfusion dependent patients (see *Table 1*). High throughput capacity makes this sort of typing cost-effective. Collecting such big data from panels for hundreds of multiplexed donors with automatic transfer of genotype to serological phenotypes will facilitate widely-matched blood product provision to recipients (93). This application for screening is analogous to donor HLA sequencing which has been successfully implemented in the field of transplantation (24). But HLA NGS diagnostics also covers recipient screening. Nowadays, NGS methodology is too expensive for the routine BGA genotyping of a single patient. However, bioinformatic extraction of donors, genetically fitting a patient's wide BGA genotype, leads to perfect donor/recipient blood matching, undoubtedly eliminating alloimmunization events. For now, patient's sequencing is narrowed down to a group of transfusion-dependent recipients and the most complicated cases with discrepant serological results (42,43). But extracting BGA data from individuals with previously performed WES/WGS for other reasons is becoming an important tool in the era of personalized medicine (93). For blood recipients in whom there are difficulties in the serological typing of a single antigen NGS, it also gives the possibility of low-cost testing through joining the library for only a single target gene to experiments performed with a similar depth of sequencing.

All applications generate basic knowledge on BGA genetics in association with a determined serological pattern of reactions and often with the discovery of a new allele. So now we are still revealing unknown sequences widening the database of alleles encoding BGAs in particular populations and ethnic groups. Nearly all publications on applying NGS for BGA genotyping contain reports of a new polymorphism found unexpectedly in serologically characterized individuals (listed in *Table 2*).

Conclusions

It should be emphasized that at present BGA NGS is a technology that is being set up, implemented and requires validation. BGA NGS is at the stage of collecting and completing genetic data with non-ideal tools (phasing of short reads, reading errors, allele drop-outs, homological allele variations) but it is also at the stage of revealing the wide world of BGA alleles and building reference databases, designing software for correct analysis that omits any technical limitations of the technology and creating programs similar to a Google translator for the automatic conversion of genetic data to transfusion language (48,62,94). Currently, the correlation between NGS and phenotyping/genotyping results varies from 92% to 99%. So far, there are no commercially available BGA NGS panels for CE-IVD diagnostics combined with dedicated software for translating genotype to phenotype. Undoubtedly, it is a matter of time until such tests, analogically to HLA, will be commercially available on the immunohematological market as well as dedicated external quality control programs.

Summarizing, the recently published works on NGS-based genotyping of BGAs have made a huge contribution to understanding the allelic variants responsible for the presence of antigenic specificities. Applying NGS protocols is useful for donor BGA screening as well as the scientific research of novel SNVs and structural variations and also in building the basis of reference sequences for BGA interpretation. The current state of BGA NGS requires a deep standardization and validation of all the steps including bioinformatic tools as well as ethical guidelines before it can be implemented in routine BGA diagnostics. Also, setting a new regulation for BG NGS validation is a separate issue, because the number of tested samples in the classical serological validation does not correspond to massive parallel sequencing technology.

Acknowledgments

Funding: The work was supported from the IHTM funds (No. B2-T4-2021).

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editor (Frederik Banch Clausen) for the series "Blood Group Genotyping" published in *Annals of Blood*. The article has undergone external peer review.

Reporting Checklist: The author has completed the Narrative Review reporting checklist. Available at https://aob.amegroups.com/article/view/10.21037/aob-21-39/rc

Peer Review File: Available at https://aob.amegroups.com/article/view/10.21037/aob-21-39/prf

Conflicts of Interest: The author has completed the ICMJE uniform disclosure form (available at https://aob.amegroups.com/article/view/10.21037/aob-21-39/coif). The series "Blood Group Genotyping" was commissioned by the editorial office without any funding or sponsorship. The author has no other conflicts of interest to declare.

Ethical Statement: The author is accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Westhoff CM. Blood group genotyping. Blood 2019;133:1814-20.
2. Denomme GA, Anani WQ, Avent ND, et al. Red cell genotyping precision medicine: a conference summary. Ther Adv Hematol 2017;8:277-91.
3. Flegel WA, Gottschall JL, Denomme GA. Implementing mass-scale red cell genotyping at a blood center. Transfusion 2015;55:2610-5; quiz 2609.
4. Belsito A, Costa D, Napoli C. Blood group genotyping for patients with autoimmune hemolytic anemia. Transl Res 2014;164:177-8.
5. Svensson AM, Delaney M. Considerations of red blood cell molecular testing in transfusion medicine. Expert Rev Mol Diagn 2015;15:1455-64.
6. Shin KH, Lee HJ, Park KH, et al. Extended Red Blood Cell Genotyping to Investigate Immunohematology Problems. Ann Lab Med 2018;38:387-8.
7. Khan J, Delaney M. Transfusion Support of Minority Patients: Extended Antigen Donor Typing and Recruitment of Minority Blood Donors. Transfus Med Hemother 2018;45:271-6.
8. Kulkarni S, Choudhary B, Gogri H, et al. Molecular genotyping of clinically important blood group antigens in patients with thalassaemia. Indian J Med Res 2018;148:713-20.
9. Hyland CA, O'Brien H, Flower RL, et al. Non-invasive prenatal testing for management of haemolytic disease of the fetus and newborn induced by maternal alloimmunisation. Transfus Apher Sci 2020;59:102947.
10. Matteocci A, Pierelli L. Red blood cell alloimmunization in sickle cell disease and in thalassaemia: current status, future perspectives and potential role of molecular typing. Vox Sang 2014;106:197-208.
11. Wheeler MM, Johnsen JM. The role of genomics in transfusion medicine. Curr Opin Hematol 2018;25:509-15.
12. McBean RS, Hyland CA, Flower RL. Blood group genotyping: the power and limitations of the Hemo ID Panel and MassARRAY platform. Immunohematology 2015;31:75-80.
13. Molano A, Apraiz I, España P, et al. Performance evaluation study of ID RHD XT, a new genotyping assay for the detection of high-prevalence RhD negative and weak D types. Vox Sang 2018;113:694-700.
14. Jongruamklang P, Gassner C, Meyer S, et al. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry analysis of 36 blood group alleles among 396 Thai samples reveals region-specific variants. Transfusion 2018;58:1752-62.
15. McBean RS, Hyland CA, Davis AC, et al. Blood group genotype analysis of Australian reagent red blood cell donors across three genotyping platforms: consistent detection of 70% phenotype genotype non concordance. ISBT Sci Ser 2014;9:309-14.</jrn>
16. Storry JR, Jöud M, Olsson ML. Automatic for the people: a rapidly evolving movement for the future of genotyping. Transfusion 2019;59:3545-7.
17. Polin H, Pelc-Klopotowska M, Danzer M, et al.

Compound heterozygosity of two novel RHAG alleles leads to a considerable disruption of the Rh complex. Transfusion 2016;56:950-5.
18. Stenfelt L, Hellberg Å, Möller M, et al. Missense mutations in the C-terminal portion of the B4GALNT2-encoded glycosyltransferase underlying the Sd(a-) phenotype. Biochem Biophys Rep 2019;19:100659.
19. Guz K, Pelc-Kłopotowska M, Purchla-Szepioła S, et al. A novel ABO*A allele with 21 base pair duplication in Polish individuals. Transfusion 2020;60:E48-50.
20. Hyland CA, Roulis EV, Schoeman EM. Developments beyond blood group serology in the genomics era. Br J Haematol 2019;184:897-911.
21. Fichou Y, Férec C. NGS and blood group systems: State of the art and perspectives. Transfus Clin Biol 2017;24:240-4.
22. Muzzey D, Evans EA, Lieber C. Understanding the Basics of NGS: From Mechanism to Variant Calling. Curr Genet Med Rep 2015;3:158-65.
23. McBean RS, Hyland CA, Flower RL. Approaches to determination of a full profile of blood group genotypes: single nucleotide variant mapping and massively parallel sequencing. Comput Struct Biotechnol J 2014;11:147-51.
24. Schöfl G, Lang K, Quenzel P, et al. 2.7 million samples genotyped for HLA by next generation sequencing: lessons learned. BMC Genomics 2017;18:161.
25. Yamamoto F, Clausen H, White T, et al. Molecular genetic basis of the histo-blood group ABO system. Nature 1990;345:229-33.
26. McGowan EC, Lopez GH, Knauth CM, et al. Diverse and novel RHD variants in Australian blood donors with a weak D phenotype: implication for transfusion management. Vox Sang 2017;112:279-87.
27. Fürst D, Tsamadou C, Neuchel C, et al. Next-Generation Sequencing Technologies in Blood Group Typing. Transfus Med Hemother 2020;47:4-13.
28. Slatko BE, Gardner AF, Ausubel FM. Overview of Next-Generation Sequencing Technologies. Curr Protoc Mol Biol 2018;122:e59.
29. Merriman B; Ion Torrent R&D Team; Rothberg JM. Progress in ion torrent semiconductor chip based sequencing. Electrophoresis 2012;33:3397-417.
30. Lang K, Wagner I, Schöne B, et al. ABO allele-level frequency estimation based on population-scale genotyping by next generation sequencing. BMC Genomics 2016;17:374.
31. Fox K, Johnsen JM, Coe BP, et al. Analysis of exome sequencing data sets reveals structural variation in the coding region of ABO in individuals of African ancestry. Transfusion 2016;56:2744-9.
32. Wu PC, Lin YH, Tsai LF, et al. ABO genotyping with next-generation sequencing to resolve heterogeneity in donors with serology discrepancies. Transfusion 2018;58:2232-42.
33. Fichou Y, Audrézet MP, Guéguen P, et al. Next-generation sequencing is a credible strategy for blood group genotyping. Br J Haematol 2014;167:554-62.
34. Fichou Y, Mariez M, Le Maréchal C, et al. The experience of extended blood group genotyping by next-generation sequencing (NGS): investigation of patients with sickle-cell disease. Vox Sang 2016;111:418-24.
35. Orzińska A, Guz K, Mikula M, et al. A preliminary evaluation of next-generation sequencing as a screening tool for targeted genotyping of erythrocyte and platelet antigens in blood donors. Blood Transfus 2018;16:285-92.
36. Schoeman EM, Lopez GH, McGowan EC, et al. Evaluation of targeted exome sequencing for 28 protein-based blood group systems, including the homologous gene systems, for blood group genotyping. Transfusion 2017;57:1078-88.
37. Roulis E, Schoeman E, Hobbs M, et al. Targeted exome sequencing designed for blood group, platelet, and neutrophil antigen investigations: Proof-of-principle study for a customized single-test system. Transfusion 2020;60:2108-20.
38. Dezan MR, Ribeiro IH, Oliveira VB, et al. RHD and RHCE genotyping by next-generation sequencing is an effective strategy to identify molecular variants within sickle cell disease patients. Blood Cells Mol Dis 2017;65:8-15.
39. El Wafi M, El Housse H, Zaid N, et al. Novel intronic RHD variants identified in serologically D-negative blood donors. Vox Sang 2017;112:796-802.
40. Tounsi WA, Madgett TE, Avent ND. Complete RHD next-generation sequencing: establishment of reference RHD alleles. Blood Adv 2018;2:2713-23.
41. Wheeler MM, Lannert KW, Huston H, et al. Genomic characterization of the RH locus detects complex and novel structural variation in multi-ethnic cohorts. Genet Med 2019;21:477-86.
42. Dinardo CL, Kelly S, Dezan MR, et al. Diversity of RH and transfusion support in Brazilian sickle cell disease patients with unexplained Rh antibodies. Transfusion 2019;59:3228-35.
43. Dinardo CL, Oliveira TGM, Kelly S, et al. Diversity of variant alleles encoding Kidd, Duffy, and Kell antigens in individuals with sickle cell disease using whole genome

sequencing data from the NHLBI TOPMed Program. Transfusion 2021;61:603-16.
44. Jakobsen MA, Dellgren C, Sheppard C, et al. The use of next-generation sequencing for the determination of rare blood group genotypes. Transfus Med 2019;29:162-8.
45. Boccoz SA, Fouret J, Roche M, et al. Massively parallel and multiplex blood group genotyping using next-generation-sequencing. Clin Biochem 2018;60:71-6.
46. Vorholt SM, Hamker N, Sparka H, et al. High-Throughput Screening of Blood Donors for Twelve Human Platelet Antigen Systems Using Next-Generation Sequencing Reveals Detection of Rare Polymorphisms and Two Novel Protein-Changing Variants. Transfus Med Hemother 2020;47:33-44.
47. Schoeman EM, Roulis EV, Perry MA, et al. Comprehensive blood group antigen profile predictions for Western Desert Indigenous Australians from whole exome sequence data. Transfusion 2019;59:768-78.
48. Lane WJ, Vege S, Mah HH, et al. Automated typing of red blood cell and platelet antigens from whole exome sequences. Transfusion 2019;59:3253-63.
49. Lopez GH, Wilson B, Turner RM, et al. Frequency of Mia (MNS7) and Classification of Mia-Positive Hybrid Glycophorins in an Australian Blood Donor Population. Transfus Med Hemother 2020;47:279-86.
50. Chou ST, Flanagan JM, Vege S, et al. Whole-exome sequencing for RH genotyping and alloimmunization risk in children with sickle cell anemia. Blood Adv 2017;1:1414-22.
51. Stef M, Fennell K, Apraiz I, et al. RH genotyping by nonspecific quantitative next-generation sequencing. Transfusion 2020;60:2691-701.
52. Montemayor C, Simone A, Long J, et al. An open-source python library for detection of known and novel Kell, Duffy and Kidd variants from exome sequencing. Vox Sang 2021;116:451-63.
53. Millard GM, Lopez GH, Turner EM, et al. Modified expression of the KEL2 (k) blood group antigen attributed to p.Leu196Val amino acid change three residues from the K/k antigen polymorphism site: implications for donor screening. Transfusion 2019;59:1156-8.
54. Millard GM, McGowan EC, Wilson B, et al. A proposed new low-frequency antigen in the Augustine blood group system associated with a severe case of hemolytic disease of the fetus and newborn. Transfusion 2018;58:1320-2.
55. Fraser NS, Knauth CM, Moussa A, et al. Genetic Variants Within the Erythroid Transcription Factor, KLF1, and Reduction of the Expression of Lutheran and Other Blood Group Antigens: Review of the In(Lu) Phenotype. Transfus Med Rev 2019;33:111-7.
56. Vrignaud C, Chiaroni J, Durieux-Roussel E, et al. CORS (CROM20): A new high-prevalence antigen in the Cromer blood group system. Transfusion 2020;60:E40-2.
57. Thornton N, Karamatic Crew V, Tilley L, et al. Disruption of the tumour-associated EMP3 enhances erythroid proliferation and causes the MAM-negative phenotype. Nat Commun 2020;11:3569.
58. Azouzi S, Mikdar M, Hermand P, et al. Lack of the multidrug transporter MRP4/ABCC4 defines the PEL-negative blood group and impairs platelet aggregation. Blood 2020;135:441-8.
59. Vrignaud C, Mikdar M, Koehl B, et al. Alloantibodies directed to the SLC44A2/CTL2 transporter define two new red cell antigens and a novel human blood group system. Transfusion 2019;59:18A.
60. Lane WJ, Westhoff CM, Uy JM, et al. Comprehensive red blood cell and platelet antigen prediction from whole genome sequencing: proof of principle. Transfusion 2016;56:743-54.
61. Lane WJ, Westhoff CM, Gleadall NS, et al. Automated typing of red blood cell and platelet antigens: a whole-genome sequencing study. Lancet Haematol 2018;5:e241-51.
62. Paganini J, Nagy PL, Rouse N, et al. Blood group typing from whole-genome sequencing data. PLoS One 2020;15:e0242168.
63. Lane WJ, Aguad M, Smeland-Wagman R, et al. A whole genome approach for discovering the genetic basis of blood group antigens: independent confirmation for P1 and Xga. Transfusion 2019;59:908-15.
64. Lane WJ, Gleadall NS, Aeschlimann J, et al. Multiple GYPB gene deletions associated with the U- phenotype in those of African ancestry. Transfusion 2020;60:1294-307.
65. Halls JBL, Vege S, Simmons DP, et al. Overcoming the challenges of interpreting complex and uncommon RH alleles from whole genomes. Vox Sang 2020;115:790-801.
66. Chang TC, Haupfear KM, Yu J, et al. A novel algorithm comprehensively characterizes human RH genes using whole-genome sequencing data. Blood Adv 2020;4:4347-57.
67. Fichou Y, Berlivet I, Richard G, et al. Defining Blood

Group Gene Reference Alleles by Long-Read Sequencing: Proof of Concept in the ACKR1 Gene Encoding the Duffy Antigens. Transfus Med Hemother 2020;47:23-32.
68. Srivastava K, Khil PP, Sippert E, 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;22:1272-9.
69. Stabentheiner S, Danzer M, Niklas N, et al. Overcoming methodical limits of standard RHD genotyping by next-generation sequencing. Vox Sang 2011;100:381-8.
70. Rieneck K, Clausen FB, Dziegiel MH. Next-Generation Sequencing for Antenatal Prediction of KEL1 Blood Group Status. Methods Mol Biol 2015;1310:115-21.
71. Wienzek-Lischka S, Krautwurst A, Fröhner V, et al. Noninvasive fetal genotyping of human platelet antigen-1a using targeted massively parallel sequencing. Transfusion 2015;55:1538-44.
72. Orzińska A, Guz K, Uhrynowska M, et al. Noninvasive prenatal HPA-1 typing in HPA-1a negative pregnancies selected in the Polish PREVFNAIT screening program. Transfusion 2018;58:2705-11.
73. Orzińska A, Guz K, Mikula M, et al. Prediction of fetal blood group and platelet antigens from maternal plasma using next-generation sequencing. Transfusion 2019;59:1102-7.
74. Wienzek-Lischka S, Bachmann S, Froehner V, et al. Potential of Next-Generation Sequencing in Noninvasive Fetal Molecular Blood Group Genotyping. Transfus Med Hemother 2020;47:14-22.
75. O'Brien H, Gardener G, Millard G, et al. An overview on droplet digital PCR for precise fetal typing and the use of MSP-based red blood cell typing. Vox Sang 2020;115:3-387.
76. Nogués N. Recent advances in non-invasive fetal HPA-1a typing. Transfus Apher Sci 2020;59:102708.
77. Takahashi K, Migita O, Sasaki A, et al. Amplicon Sequencing-Based Noninvasive Fetal Genotyping for RHD-Positive D Antigen-Negative Alleles. Clin Chem 2019;65:1307-16.
78. Rieneck K, Egeberg Hother C, Clausen FB, et al. Next Generation Sequencing-Based Fetal ABO Blood Group Prediction by Analysis of Cell-Free DNA from Maternal Plasma. Transfus Med Hemother 2020;47:45-53.
79. Möller M, Jöud M, Storry JR, et al. Erythrogene: a database for in-depth analysis of the extensive variation in 36 blood group systems in the 1000 Genomes Project. Blood Adv 2016;1:240-9.
80. Möller M, Hellberg Å, Olsson ML. Thorough analysis of unorthodox ABO deletions called by the 1000 Genomes project. Vox Sang 2018;113:185-97.
81. McBean RS, Hyland CA, Hendry JL, et al. SARA: a "new" low-frequency MNS antigen (MNS47) provides further evidence of the extreme diversity of the MNS blood group system. Transfusion 2015;55:1451-6.
82. Lopez GH, Turner RM, McGowan EC, et al. A DEL phenotype attributed to RHD Exon 9 sequence deletion: slipped-strand mispairing and blood group polymorphisms. Transfusion 2018;58:685-91.
83. Yahalom V, Pillar N, Zhao Y, et al. SMYD1 is the underlying gene for the AnWj-negative blood group phenotype. Eur J Haematol 2018;101:496-501.
84. McBean R, Liew YW, Wilson B, et al. Genotyping confirms inheritance of the rare At(a-) type in a case of haemolytic disease of the newborn. J Pathol Clin Res 2015;2:53-5.
85. Cvejic A, Haer-Wigman L, Stephens JC, et al. SMIM1 underlies the Vel blood group and influences red blood cell traits. Nat Genet 2013;45:542-5.
86. Omae Y, Ito S, Takeuchi M, et al. Integrative genome analysis identified the KANNO blood group antigen as prion protein. Transfusion 2019;59:2429-35.
87. Yeh CC, Chang CJ, Twu YC, et al. The molecular genetic background leading to the formation of the human erythroid-specific Xga/CD99 blood groups. Blood Adv 2018;2:1854-64.
88. Duval R, Nicolas G, Willemetz A, et al. Inherited glycosylphosphatidylinositol defects cause the rare Emm-negative blood phenotype and developmental disorders. Blood 2021. doi: blood.2020009810.
89. Ardui S, Ameur A, Vermeesch JR, et al. Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. Nucleic Acids Res 2018;46:2159-68.
90. Branton D, Deamer DW, Marziali A, et al. The potential and challenges of nanopore sequencing. Nat Biotechnol 2008;26:1146-53.
91. Available online: http://www.isbtweb.org/working-parties/red-cell-immunogenetics-and-blood-group-terminology/
92. Fennell K, Hoffman R, Yoshida K, et al. Effect on gene expression of three allelic variants in GATA motifs of ABO, RHD, and RHCE regulatory elements. Transfusion

2017;57:2804-8.
93. Montemayor C, Brunker PAR, Keller MA. Banking with precision: transfusion medicine as a potential universal application in clinical genomics. Curr Opin Hematol 2019;26:480-7.
94. Giollo M, Minervini G, Scalzotto M, et al. BOOGIE: Predicting Blood Groups from High Throughput Sequencing Data. PLoS One 2015;10:e0124579.

doi: 10.21037/aob-21-39
Cite this article as: Orzińska A. Next generation sequencing and blood group genotyping: a narrative review. Ann Blood 2023;8:4.