Mass-scale donor red cell genotyping using real-time array technology

G.A. Denomme and M.J. Schanen

Blood centers are in the unique position to evaluate large numbers of blood donations for antigen-negative blood types. The limitations with the use of hemagglutination, however, can be circumvented with red cell genotyping. The reagents used for genotyping are synthesized and can be designed for any of the known blood group antigen single nucleotide polymorphisms that are associated with blood group antigen expression. There is interest in the application of mass-scale red cell genotyping of blood donors to find rare phenotypes and rare combinations of antigens. When performed on donors who are predicted to donate again after testing, integrating the genotype information with existing donor data and demographics provides the blood center with real-time information to identify the common clinically relevant blood group antigens demanded by hospital transfusion services. This review outlines a red cell genotype methodology using TaqMan chemistry and existing algorithms and data handling to gain the full value of mass-scale red cell genotyping of blood donors. ***Immunohematology* 2015;31:69–74.**

Key Words: red cell genotyping, blood group genotyping, mass-scale, high-throughput TaqMan chemistry

Transfusion recipients who become alloimmunized to blood group antigens require crossmatch-compatible blood that is negative for the cognate antigens. Blood collection facilities have the responsibility of finding antigen-negative blood among donor units. Historically, manual and automated phenotyping methods are still used to screen blood donations routinely for antigen-negative types that are likely to be in demand.[1] The screening is done with a limited collection of licensed or unlicensed blood grouping reagents, and the results are recorded for later use should a blood donor donate again. Knowing the donor's history of donations provides efficiencies to antigen typing programs because repeat donors are predicted to donate in the future and do not need to be repeatedly typed.

Knowledge of the molecular basis of blood group expression[2] and the development of platforms capable of screening on a mass scale[3] have created the opportunity to screen blood donors for antigen-negative types using DNA rather than red blood cells. High-throughput genotyping platforms can enable mass-scale screening when the output files from these instruments are electronically interfaced with the laboratory information management system (LIMS), making red cell genotyping an attractive screening alternative to hemagglutination-based platforms. A desired goal for a blood center is to have enough antigen-negative blood units to meet approximately 95 percent of blood requests that it receives.[4] For blood centers that support the American Rare Donor Program, comprehensive genotyping of a large number of donors contributes to the desperately needed rare blood types at an incremental cost to screening for the common blood group antigens of the RH, MNS, Kell, Duffy, and Kidd systems. This article outlines a method used to genotype blood donors on a high-throughput mass scale using TaqMan chemistry. Several manufacturers have instruments that use TaqMan chemistry and can be designed for red cell genotyping. Systems that have been evaluated for red cell genotyping using TaqMan technology include the Wafergen SmartChip (Wafergen Biosystems, Fremont, CA), Fluidigm June 96.96 Genotyping IFC (Fluidigm, South San Francisco, CA), and QuantStudio 12K Flex Real-Time PCR System (Life Technologies, Grand Island, NY).

Red Cell Genotyping

Red cell genotyping platforms are based on the polymerase chain reaction (PCR) coupled with nucleic acid chemistry that can be performed with rapid turnaround time. A desirable turnaround time from setup to result output is in the 6- to 8-hour range. When used to screen a large number of donor samples for antigen-negative blood types (i.e., mass-scale), data parsing and comparative analysis of the results with other donor information optimally should follow within 24 hours. At a maximum, the process should be completed to coincide with the release of blood units after infectious disease testing.

The term "array" used in this review refers to platforms that simultaneously interrogate a multitude of single nucleotide polymorphisms (SNPs) in a single format, whether or not the array is attached to a support like a bead or chip. An "open" system refers to the fact that the make-up of the array can be changed with minimal to no impact on platform performance. This flexibility is based on the fact that the assays are not multiplex; each assay is a single reaction performed in a single well. The design layout can therefore be altered at the discretion

of the operator, simply by expanding the number of assays or excluding one assay and introducing another.

The term "high-throughput" applies to any platform regardless of the nucleic acid chemistry: PCR–enzyme-linked immunosorbent assay (ELISA), PCR–oligonucleotide extension, sequence-specific PCR–mass spectrometry, and PCR–oligonucleotide hybridization all have high-throughput capability. For red cell genotyping of donors, this term means that blood units are tested and the desired antigen-negative units are identified before the blood unit leaves the control of the blood center. Mass-scale genotyping is achieved when an array of SNPs is analyzed on many samples. The method for high-throughput mass-scale red cell genotyping includes DNA extraction, genotyping, and data transfer to a LIMS. Process control is achieved by collecting metrics at key steps that will identify process failures should they occur. Desirable metrics are:

- DNA concentration and the OD260/280 ratio
  - A concentration or ratio less than a given threshold = sample extraction failure
- Evaluation of the result values (fluorescence integrity)
  - Lack of passive dye (e.g., ROX™ Passive Reference Dye) detected during results imaging = DNA load failure
- Genotype results of known controls
  - Controls that fail to give the expected result = an assay or run failure
- Genotype results of unknown samples
  - Outlier, Do not Call, or No Call flags across multiple assays = sample failure (likely because of a DNA extraction or load failure)
  - Unexpected combinations of SNPs (e.g., homozygous FY GATA mutation with FY*A/FY*B)

OpenArray Genotyping System Workflow

The OpenArray Real-Time PCR System (Life Technologies) is summarized here. It uses real-time PCR–fluorogenic 5′ nuclease TaqMan chemistry and is available as the QuantStudio 12K Flex system. TaqMan chemistry can discriminate between SNPs by using specifically designed pairs of single-stranded DNA oligonucleotide probes and the 5′ to 3′ exonuclease activity of *Taq* polymerase.[5] Each TaqMan probe consists of a short single-stranded oligonucleotide containing a 5′ fluorophore (e.g., 6-carboxyfluorescein [FAM] or 4,7,2′-trichloro-7′-phenyl-6-carboxyfluorescein [VIC]) and a 3′ quencher (e.g., tetramethylrhodamine [TAMRA]). The principle was named TaqMan in reference to its similarity with the PacMan video game: the 5′ to 3′ exonuclease activity of *Taq* polymerase degrades a hybridized probe during PCR amplification. In brief, a TaqMan assay comprises two amplification primers and two probes of approximately 20 nucleotides that differ by one SNP. The *Taq* polymerase extends the amplification primers in the 3′ to 5′ direction during the elongation stage of PCR. The single-stranded DNA oligonucleotide probes anneal to the region containing the SNP of interest. With each successive round, as the polymerization encounters a hybridized probe, the 5′ to 3′ exonuclease activity degrades the probe and releases the fluorophore from the quencher. The observed fluorescence is an indication that a particular probe was annealed and the SNP is present (Fig. 1). The SNP discrimination is attributable to the poor hybridization that each probe has for the alternate allele containing the mismatched nucleotide. Mismatched probes are not hybridized efficiently and are not degraded. Because SNPs represent the molecular basis of many blood group antigens, a pair of TaqMan probes can determine the three possible genotypes. TaqMan technology can also be used to distinguish nucleotide insertions (or deletions) because one probe can be specific to the insertion sequence and the other probe can be designed to anneal across the deletion (e.g., the 17–base pair deletion of the common Vel– genotype[6]).

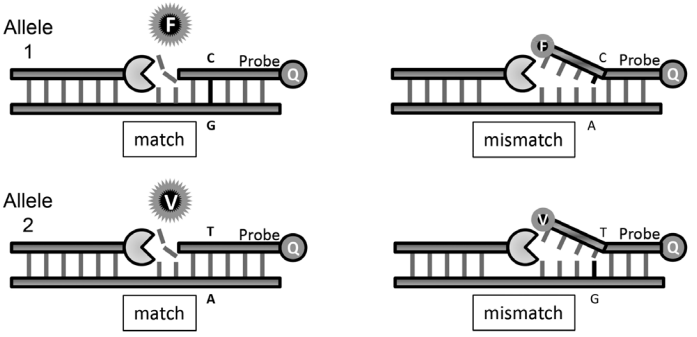

Fig. 1. The principle of the TaqMan assay involves prevention of fluorescence of a fluorophore by the close proximity with a quencher in a nucleotide probe. As the *Taq* polymerase reaches the probe, the 5′ to 3′ exonuclease activity of the polymerase degrades the probe, the fluorophore is released from the quencher, and fluorescence is detected (left panels). A typical TaqMan assay has two such probes that detect a single nucleotide difference at a specific location between two alleles of a given gene. Probes that are not 100 percent homologous with the DNA sequenced do not bind sufficiently and the *Taq* polymerase cannot degrade the probe to release the fluorophore from the quencher (right panels).

Materials and Methods

The red cell genotyping method described in the proof-of-principle[7] consists of three major steps: DNA extraction templated on a 96-well format; red cell genotyping (assay

setup, PCR amplification, and fluorescence measurement), and results analysis. Electronic uploading to a LIMS is included in the methodology because it is desirable to compare existing historical phenotype with genotype results, including D antigen. Without this electronic data step, high-throughput genotyping is reduced to mass-scale testing followed by perusal of instrument printouts for desired antigen-negative genotypes.

The OpenArray system requires the use of a robotic sample loader for the arrays (OpenArray Accufill System, Life Technologies) because an array consists of 3072 open wells. The OpenArray is a thin stainless steel wafer comprising holes capable of holding 33 nL of fluid that are precision-drilled in an arrangement of 48 matrices or subarrays. Each subarray is configured in an 8 × 8 format. The hydrophobic exterior of the wafer maintains the liquid inside the wells. Each well is preloaded with a TaqMan assay by the manufacturer. Genomic DNA concentrations can range from 10 to 50 ng/μL, with the higher concentration preferred because the nano-volume reduces the input DNA to approximately 80 genome equivalent, or 40 copies of an allele for a heterozygous sample. For red cell genotyping, 32 TaqMan assays custom-designed to identify 42 common and rare blood group antigens are preloaded into the wells.[3] Therefore, each 8 × 8 subarray consists of two sets of four rows containing the 32 assays (Fig. 2). The instrument can handle up to four arrays in a single run. The attractive feature of single-well testing is the flexibility to change one or more assays without compromising the validity of other assays, since each well represents a single TaqMan reaction. Arrays must be ordered as a minimum of a set of 10, however.

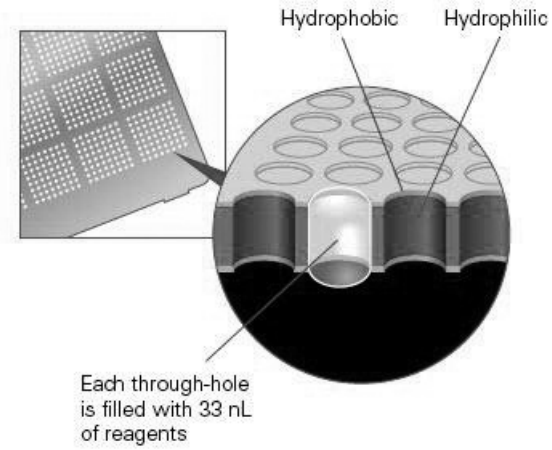

Fig. 2. A schematic diagram of an OpenArray (Life Technologies). Each wafer consists of 48 subarrays (bottom) arranged in matrices of 8 × 8 wells (magnified on the right) and capable of holding 33 nL of fluid. The negative hydrophobic nature of the outside maintains the fluid within the hydrophilic wells (© 2015 Thermo Fisher Scientific Inc. Used under permission).

DNA Extraction

Any DNA extraction method is acceptable provided the yield is greater than 10 ng/μL and of high purity (OD260/280 nm ratio: 1.7 ± 0.2). ISBT unit barcodes can be used as the sample ID electronically scanned into a 96-well template, which matches the number of samples interrogated in a single array. A PCR-grade water "sample" serves to evaluate cross-contamination during DNA extraction and can be used as the No Template Control (NTC) during genotyping. The DNA concentration and OD206/280 nm ratio is measured on a subset of samples from each 96-well plate using an eight-channel spectrophotometer by Nanodrop chemistry. If one of the eight samples falls outside the acceptable limits, the entire plate can be evaluated for repeat extraction. Alternatively, all samples can be evaluated and then the decision made whether individual samples should be subjected to DNA extraction or whether the extraction run failed. The extraction of DNA into removable eight-well strips facilitates the replacement of any one column with a set of DNA controls that are pre-pipetted into eight-well strips.

Red Cell Genotyping Steps

1) Array setup for 384 samples: Typically, a set of 16 controls and 4 NTCs are included with each batch of 4 arrays. In this way, 364 unknown samples can be interrogated in a single run. A batch of 4 arrays potentially produces 15,288 discrete results (91 unknown samples per array × 4 arrays × 42 blood group antigens). The initial sample and master mix setup has three steps that can be performed with an 8- or 12-multichannel micropipettor.
   a. Pipette the master mix (*Taq* polymerase, dNTPs, and PCR buffer) into a 384-well plate
   b. Pipette DNA samples, controls, or NTCs into the 384-well plate
   c. Use the robotic OpenArray AccuFiller to distribute the samples from the master plate to the subarrays
2) PCR amplification: The arrays are placed into cassettes and sealed with an oil-based product. The arrays are laid flat on a standard thermal cycler that has been fitted with an array holder. The assays are designed for standard TaqMan two-stage cycling as outlined in Table 1.
3) Fluorescence measurement: After PCR amplification, the arrays are evaluated for FAM and VIC fluorescence. The array can be evaluated for fluorescence integrity if the

Table 1. Thermal cycling conditions for real-time array polymerase chain reaction

Temperature (°C)	Time (min)	Cycles	Ramp rate (%)
93	10:00	1	100
95	00:45		84
94	00:13	40–50	100
53	2:14		44
25	2:00	1	100
4	Hold		

master mix contains a passive fluorescence dye like ROX. The passive dye fluorescence evaluation allows each well to be assessed for DNA content.

4) Results analysis: The platform software assigns samples with similar FAM/VIC fluorescence ratios into single groups, which represents the three genotype clusters. Any result that does not meet the stringency set by the software is defined as "Outlier" (excessively high fluorescence), a "Do Not Call" (no cluster designation), or a "No Call" (low fluorescence), each of which are excluded from assigning a result to the output file. Essentially, a sufficient number of DNA controls are tested such that the software correctly assigns the value to the expected homozygous genotypes and the heterozygous genotype. Assays that interrogate SNPs for rare blood group genotypes such as Yt(a–) may not have sufficient data for the software to assign a cluster. In these instances, the rare homozygous cluster can be defined by the operator or assumed on the basis of the two remaining genotype clusters. Controls that do not generate a result or are not assigned to the correct cluster indicate that the assay has failed. Unknown samples that reproducibly lie outside a cluster for a particular assay may be an indication of a possible variant allele and can be further evaluated by manual molecular methods.

Upload to a LIMS

The advantage of red cell genotyping is the ability to manufacture synthetic reagents for assays that interrogate the vast number of SNPs associated with blood group antigen expression. Thus, the output of high-throughput mass-scale testing creates "big data" that cannot be reasonably scanned by eye or reviewed on a sample-by-sample results basis. End-to-end connectivity from DNA extraction to predicted phenotype display is necessary and can be accomplished by parsing the instrument output file, composed of the A, C, G, or T nucleotides, and then translating the parsed data into predicted phenotypes using a defined electronic rules engine.[8] We chose to express predicted phenotypes using ISBT nomenclature (e.g., FY1) to distinguish them from serological phenotypes (e.g., Fy[a]). The predicted phenotype data set is then aligned with other phenotype and donation information. Figure 3 summarizes the steps necessary to upload the data to be compatible with a blood center LIMS. The parsed data and donor information can be subjected to logical arguments to create exception reports. The following exception reports provide important control and information throughout the entire red cell genotyping process:

- Donor did not meet repeat donation criteria
- Previous genotype does not match current genotype (repeat testing control)
- Donor genotyped twice (selection criteria not met)
- Phenotype does not match genotype (discrepancy)
- *RHD* detected in a D– donor (serologic weak, Del, or non-functional allele)
- Carrier of a rare recessive allele detected (e.g., RHCE*ce667C/T, 676G/G)

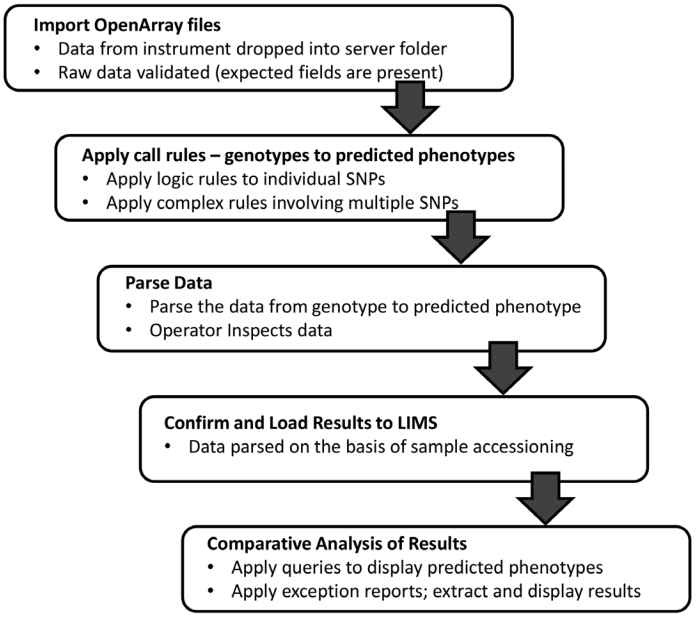

Fig. 3. Summary of the steps taken to import, translate, parse, and analyze red cell genotyping data in relation to donor information. Queries and reports can be designed to extract useful predicted phenotype (genotype) information. Predicted phenotypes are distinguished from antigen phenotypes by using a distinct nomenclature (not shown).

Discussion

Strengths and Limitations

TaqMan open array systems are very flexible. QuantStudio arrays can be formatted to run 16 assays on 144 samples, 32 assays on 96 samples, or 64 assays on 48 samples. Array

setup, PCR amplification, fluorescence evaluation, and results analysis for a run of four arrays in a single run of 32 assays on 96 samples plus controls (n = 384) consumes about 2.5 hours of hands-on time and can be accomplished comfortably in a single 8-hour shift. The equivalent hands-on time using a manual TaqMan and 96-well microplate format run on a fluorescent-based instrument (e.g., Lightcycler) would be three samples plus controls. The QuantStudio Accufiller provides the automation necessary for the critical sample loading step. Reagent costs include DNA extraction, TaqMan assays, master mix, nanofluidic arrays, and precision pipette tips for nanoliter dispensing. Fixed costs including capital equipment (depreciated over 7 years), reagents and consumable costs, along with hands-on time and data review places the overall cost at approximately $1.35 per SNP.

One advantage of an "open" system is the ease of assay design and validation. Each assay is designed and tested for robustness in separate validation experiments. The probe length and position across the SNP of interest can be adjusted so that the assay performs optimally using one TaqMan amplification profile. Software programs exist that help with the *in silico* design of amplification primers and TaqMan probes; typically, adjustments to optimize an assay are needed in about 10 percent of all assays developed. Once validated using the standard protocol, the assay can be added to the platform. For example, the gene responsible for the expression of Vel was reported in 2013, and a TaqMan assay was developed shortly thereafter. It could not be added to a multiplex assay without re-validating all the remaining assays in the mix. Nevertheless, using a single reaction well design, the Vel assay was validated and could be added to the suite, or substituted for another to maintain a preset format of, for example, 32 assays. In addition, as phenotype–genotype discrepancies identify low-frequency nucleotide variation that affects performance, assays can be redesigned to account for this variant nucleotide. This feature is particularly attractive when a founder allele is identified among an ethnic or racial population.

As with any nucleic chemistry, the TaqMan principle has some weaknesses. Genotyping methods that rely on PCR to amplify a region of interest before interrogation are error-prone to nucleotide variation in the region of the amplification primers. Nucleotide variations in these regions can appear as functional hemizygotes. In other words, only the single non-variant allele is amplified, and the sample appears homozygous for the associated SNP. A low-frequency nucleotide variation in the region of the oligonucleotide probes affects TaqMan assays. A probe will not have sufficient homology to an SNP if there is a mismatch elsewhere in the probe. The sample will appear homozygous for the alternate SNP that does not contain the variant nucleotide. For regions with known SNPs in close proximity, Taqman assays can be designed with inosine incorporated into the probes to compensate for the destabilizing effect of a nearby SNP.[9]

Validation of Custom Assays

We prefer to evaluate 10 samples of each genotype (10 homozygous, 10 heterozygous, 10 low frequency) for SNPs that have relatively equal allele frequencies. For rare blood group antigens, we test 10 of the common antigens, and as many heterozygous samples as can be reasonably found, and one rare homozygous allele. Without at least one rare homozygous allele, the three genotype clusters cannot be defined and with that shortcoming, the heterozygous cluster may not be clearly defined.

Implementation

Mass-scale testing can be implemented for one of two purposes. Red cell genotyping can be intended to be used as a screening tool, with the results confirmed by serology with licensed or unlicensed antisera, or alternatively can be confirmed by a validated molecular test. As of April 2014, Immucor's PreciseType is the only U.S. Food and Drug Administration–licensed molecular test for RBC genotyping. Lab developed tests that have been validated in compliance with national lab standards present an alternative at the present time.

When used as a screening tool, assay robustness may be relaxed so that the number of samples tested is maximized. This approach means that assays with high guanine/cytosine content or that show a limited variation are still acceptable for use because (1) clustering will for the most part exclude these samples from assigning a wrong genotype and (2) the confirmation of antigen expression performed to label blood units will identify discrepancies. When implemented as a screening tool, our preference is to optimize assays such that the number of samples that are not assigned a genotype (the No Call rate) is less than 1 percent for 90 percent of the assays, with no assays having a No Call rate exceeding 5 percent.[3] Otherwise, too much data are irretrievably lost.

We have evaluated over 50,000 samples using TaqMan assays in an open platform design. Whether the approach will be widely adopted in transfusion medicine is uncertain. The flexible design and development costs per assay, along with multiple platform choices, make it attractive to become the preferred screening tool for rare antigens among donors.

The flexibility does have the disadvantage that reagent quality control is the responsibility of the purchaser of the platform. The lack of commercially available reagent kits relies on collaborative agreements between institutions to share technical information. Otherwise, the utility of this approach is limited to labs with the technical expertise to design their own assays and the availability of rare DNA genotyped samples from repositories such as the Serum Cells and Rare Fluids exchange.

References

1. Mohamed S, Muna I. Characterisation of Rh and other blood group systems amongst the Maldivian blood donors. Med J Malaysia 2013;68:393–6.
2. Denomme GA. Molecular basis of blood group expression. Transfus Apher Sci 2011;44:53–63.
3. Flegel WA, Gottschall JL, Denomme GA. Implementing mass-scale red cell genotyping at a blood center. Transfusion 2015 June 20, doi: 10.1111/trf.13168
4. Flegel WA, Gottschall JL, Denomme GA. Integrating red cell genotyping into the blood supply chain: a population-based study. Lancet Haematol 2015;2:e282–8
5. Holland PM, Abramson RD, Watson R, et al. Detection of specific polymerase chain reaction product by utilizing the 5′–3′ exonuclease activity of Thermus aquaticus DNA polymerase. Proc Natl Acad Sci U S A 1991;88:7276–80.
6. Storry JR, Jöud M, Christophersen MK, et al. Homozygosity for a null allele of SMIM1 defines the Vel-negative blood group phenotype. Nat Genet 2013;45:537–41.
7. Hopp K, Weber K, Bellissimo D, et al. High-throughput red blood cell antigen genotyping using a nanofluidic real-time polymerase chain reaction platform. Transfusion 2010;50: 40–6.
8. Bhagat A, Briggs L, Denomme G, et al. Genotyping: creating a rosetta stone. AABB News 2011;13:4–8.
9. Ohtsuka E, Matsuki S, Ikehara M, et al. An alternative approach to deoxynucleotides as hybridization probes by insertion of deoxyinosine at ambiguous codon positions. J Biol Chem 1985;260:2605–8.

Gregory A. Denomme, PhD, FCSMLS(D) (corresponding author), Director, Immunohematology and Transfusion Services, Diagnostic Laboratories, Blood Center of Wisconsin, and Senior Investigator, Blood Research Institute, greg.denomme@bcw.edu; and Michael J. Schanen, Manager, Immunohematology and Transfusion Services, Blood Center of Wisconsin, 638 North 18th Street, P.O. Box 2178, Milwaukee, WI 53201-2178.