Rapid *RHD* zygosity determination using digital PCR

Kelly A. Sillence, Amr J. Halawani, Wajnat A. Tounsi, Kirsty A. Clarke, Michele Kiernan, Tracey E. Madgett*, Neil D. Avent

School of Biomedical and Healthcare Sciences, Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, PL4 8AA, UK

*Address Correspondence to this author at: School of Biomedical and Healthcare Sciences, Plymouth University, Plymouth, UK, PL4 8AA. Telephone +44 1752 584894; e-mail tracey.madgett@plymouth.ac.uk

Keywords: Zygosity, digital PCR (dPCR), long-range PCR (LR-PCR), next generation sequencing (NGS), genotype, serology and single nucleotide polymorphisms (SNPs).

List of Abbreviations: Hemolytic disease of the newborn and fetus (HDFN), real-time PCR (qPCR), digital PCR (dPCR), minutes (mins), double-stranded DNA (dsDNA), droplet digital PCR (ddPCR), *RHD* exon 5 (*RHD*5), *RHD* exon 7 (*RHD*7), High Performance Liquid Chromatography (HPLC), seconds (s), long-range PCR (LR-PCR), next generation sequencing (NGS), Ion sphere particles (ISPs), Ion Personal Genome Machine™ (Ion PGM™), Variant Caller Files (VCF), human

genome 19 (hg19), Integrative Genomics Viewer (IGV), the Single Nucleotide Polymorphism Database (dbSNP), single nucleotide polymorphism (SNP), Browser Extensible Data (BED), cell-free DNA (cfDNA), National Center for Biotechnology (NCBI).

Human Genes: Reference genes: *AGO1* (argonaute RISC catalytic component 1, HGNC: 3262) (or *AGO1* (eukaryotic translation initiation factor 2C, 1)). Target genes: *RHD* (Rh blood group, D antigen, HGNC: 10009) (or Rhesus blood group, D antigen).

Abstract

Background: Paternal zygosity testing is used to determine the hemi- or homozygosity of *RHD* in pregnancies at risk of hemolytic disease of the newborn and fetus (HDFN). Currently, this is achieved using real-time PCR or the RH box PCR, which can be difficult to interpret and unreliable, particularly for black African populations.

Method: DNA samples extracted from 58 blood donors were analysed using two multiplex reactions for *RHD* specific targets against a reference (*AGO1*) to determine gene dosage using digital PCR. Results were compared to serological data and the correct genotype for two discordant results was determined by long range-PCR, next generation sequencing and conventional Sanger sequencing.

Results: The results show clear and reliable determination of *RHD* zygosity using digital PCR and revealed that four samples did not match the serologically predicted genotype. Sanger sequencing and long range-PCR followed by next generation sequencing revealed that the correct genotypes for samples 729M and 351D, which were serologically typed as R_1R_2 (DCe/DcE), were R_2r' (DcE/dCe) for 729M and R_1r'' (DCe/dcE), R_0r^y (Dce/dCE) or R_Zr (DCE/dce) for 351D, in concordance with the digital PCR data.

Conclusion: Digital PCR provides a highly accurate method to rapidly define blood group zygosity, and has clinical application in the analysis of Rh phenotyped or genotyped samples. The vast majority of current blood group genotyping platforms are not designed to define zygosity, and thus this technique maybe used to define paternal *RH* zygosity in pregnancies at risk of HDFN, and distinguish between homo- and hemizygous *RHD* positive individuals.

Introduction

Of the 36 blood group systems Rh is the most complex at the genetic level and is the major cause of hemolytic disease of the fetus and newborn (HDFN) and a major cause of transfusion reactions during alloimmunization events. The *RH* genes, *RHD* and *RHCE*, are well characterized at the genetic level, with a combination of SNPs in *RHCE* being responsible for the C/c and E/e polymorphisms, whilst gene deletion and hybrid *RHD-RHCE* genes are responsible for D-negative phenotypes, plus SNPs and hybrid genes being responsible for partial and weak D phenotypes (1, 2). All known mutations have been well catalogued and best described in the RhesusBase resource (3).

Paternal *RHD* zygosity testing is important for prenatal management of alloimmunized women. Where fathers are homozygous D- there is no risk of HDFN for the current pregnancy or subsequent pregnancies that may follow. Pregnancies to homozygous D+ fathers (with the assumption of paternity) will by definition carry RhD positive fetuses, and can be considered for more focused clinical management. For hemizygous D+ fathers non-invasive prenatal testing (NIPT) is required for a definitive diagnosis.

Previously published methods for *RHD* zygosity testing have included real-time PCR (qPCR) assessment of *RHD* gene dosage, assessment of the hybrid *Rhesus box* found in D-negative individuals with the *RHD* gene deletion genotype and allele-specific PCR methods, as well as mass spectrometry-based methods. (2, 4-10). Zygosity testing targeting the hybrid *Rhesus box* found in *RHD*-deletion type cde haplotypes is complicated because of differences in the hybrid box amongst individuals of African descent (5, 11).

The incidence of common *RH* haplotypes in Caucasian, African black and Asian populations has been defined serologically. In *RHD* positive individuals the DCe haplotype is prevalent in Asian (73%) and Caucasian (42%) populations, but in African black populations the Dce haplotype has a higher incidence (59%) (12). *RHD* negative individuals are rarely found in Asian populations (<4%), but the dce haplotype is frequently found in Caucasian (39%) and African black (20%) populations (12). Rare haplotypes such DCE, dCe, dcE and Dce are considerably less prevalent with frequencies of 0.24%, 0.98%, 1.19% and 2.57% respectively, in Caucasian populations (12). However, it has been difficult to define the precise population frequencies of the various *RH* haplotypes due to the inability to differentiate between hemi- or homozygous individuals. For example, an individual with the phenotype DCe would be designated as the most common presumed genotype DCe/DCe rather than DCe/dCe. Thus presumed genotype, based on probability, is the manner in which donor and patient red cells are labelled. Zygosity determination of the above would define which presumed genotype (DCe/DCe or DCe/dCe) (two copies of the *RHD* gene versus one copy of the *RHD* gene) is carried by a particular individual.

Previously we have applied digital PCR (dPCR) to the analysis of free fetal DNA derived from maternal plasma (13). In this study we have utilized dPCR as a more accurate quantitative PCR method than conventional qPCR to define *RHD* zygosity. We found rare haplotypes in a relatively small cohort of samples and identified that for three samples (plus one weak D sample) their predefined and labelled presumed genotype was indeed incorrect.

Materials and Methods

Study Participants

Human whole blood samples (*n*= 79) were supplied by the National Health Service Blood and Transplant (NHSBT) (Bristol, UK) (donated with informed consent) and transported to NHS Plymouth Hospitals Trust, Plymouth, UK for collection.

Sample Processing

Samples were processed in two ways. Human whole blood samples (*n = 25*) were collected in EDTA tubes (5-10 mL total blood volume) and centrifuged at 1 600xg for 10 minutes (min) at room temperature. The plasma was carefully removed and transferred to a 15 mL tube. The plasma was then re-centrifuged at 16 000xg for 10 min. All samples were processed within 48 to 96 hours of collection and plasma aliquots (1 mL) were stored at -80°C.

Human whole blood samples (*n = 54*) were collected in EDTA tubes (5-10 mL total blood volume) and centrifuged at 2 500xg for 10 min at room temperature. The buffy coat layer was carefully removed and transferred to a 1.5 mL tube for immediate processing to genomic DNA (gDNA). All blood samples were processed within 48 to 96 hours of blood collection.

DNA extraction from plasma

Plasma extractions were performed as non-pregnant controls from maternal plasma experiments (13) and were further utilised in this study. DNA was extracted from two 1 mL aliquots of plasma using the QIAamp Circulating Nucleic Acid (CNA) kit (Qiagen, West Sussex, UK) using the QIAvac 24 Plus (Qiagen). The extraction process was as the manufacturer's protocol and each sample was eluted in 60 μL Buffer AVE (RNase free water containing 0.04% (*w/v*) sodium azide). No DNase or

RNase treatment was used. Following DNA extraction, samples were quantified on the Qubit® 2.0 Fluorometer (Life Technologies, Paisley, UK) using the Qubit® dsDNA HS assay kit (Life Technologies). Samples were stored at -20°C as 60 µL aliquots for up to four weeks.

DNA extraction from buffy coat

For *RHD* intronic SNP sequencing, gDNA was extracted from buffy coats using the QIAamp DNA Blood Mini kit (Qiagen) according to the manufacturer's instructions. DNA was eluted in 200 µL Buffer AE and incubated at room temperature for 5 min before centrifugation at 11 865xg for 1 min. For the *RHD* long-range PCR (LR-PCR), gDNA was extracted from buffy coats using the Gentra® Puregene® Blood kit (Qiagen) according to the manufacturer's instructions for RNA-free DNA. As the buffy coat contained red blood cells, RBC Lysis Solution was used. Each sample was eluted by adding 300 µL of DNA hydration solution and mixed vigorously for 5 seconds, followed by incubation at 65°C for 1 hour. The tube was then incubated at room temperature overnight with gentle shaking in order to mix the gDNA with the DNA hydration solution. Finally, the pure gDNA was transferred into a new 1.5 mL tube and stored at −20°C. Following DNA extraction, samples were quantified on the Qubit® 2.0 Fluorometer (Life Technologies) using the Qubit® double-stranded DNA (dsDNA) High Sensitivity assay kit (Life Technologies).

PCR Primers and Probes for dPCR

Two multiplex reactions were tested on the QX100™ droplet digital PCR (ddPCR) platform (Bio-Rad Laboratories, Herfordshire, UK) for *RH* zygosity testing (Table 1), as previously described in Sillence *et al.* (13). The oligonucleotide sequences (High Performance Liquid Chromatography (HPLC) purified, Eurofins Genomics, Germany)

and amplicon sizes for all target (FAM-labelled) and reference (HEX-labelled) regions are shown in Table 1. Prior to zygosity testing, primer annealing temperatures (56°C to 60°C) were optimised for both multiplex reactions (see Supplemental Figure 1). The results in Supplemental Figure 1a showed successful droplet separation of the *RHD5* (FAM) target at all annealing temperatures, but the *AGO1* (HEX) reference showed sub-optimal separation at 60°C. Droplet separation for the *RHD7* (FAM)/ *AGO1* (HEX) multiplex reaction (see Supplemental Figure 1b) demonstrated the same pattern as previously discussed for the *RHD5* (FAM)/*AGO1* (HEX) multiplex reaction. However, the optimal ratio was visible at 58.4°C (0.995). Therefore 58°C was determined to be the optimum annealing temperature for both multiplex reactions.

dPCR

The dPCR reactions were conducted in duplicate and run on the QX100™ Droplet Generator (Bio-Rad) following manufacturer's instructions (see Sillence *et al.* (13)). Plasma extracted samples were not diluted and a standard volume of template DNA (5 µL) was added. Samples extracted from buffy coat were diluted and 50ng of DNA was added to each 20 µL reaction and a non-template control (NTC) was included in each assay.

Data Analysis for dPCR

The raw fluorescent data from the ddPCR platform was analysed using the Bio-Rad QuantaSoft v1.2 software. Once thresholds for each sample had been set manually using the 1D amplification plot, positive and negative droplets were determined (see Supplemental Figure 1). The concentration was then determined by the software using Poisson statistics (95% confidence interval) for each sample. The ratio of the

target (*RHD5*-FAM and *RHD7*-FAM) over the reference (*AGO1*-HEX) for each sample was calculated as follows: FAM (copies/ µL)/HEX (copies/ µL). All statistical analysis was performed using Mann Whitney U Test (SigmaPlot Version 12.5) and significance was accepted at $p<0.05$.

RHD **LR-PCR and Next generation sequencing (NGS)**

gDNA samples from blood donors of different phenotypes were tested using LR-PCR. Three PCR products were designed to cover the entire *RHD* gene (Table 2). The HPLC-purified primers were from Eurofins MWG Operon (London, United Kingdom). The PCR reaction contained a final 1x concentration of PrimeSTAR GXL Buffer (Takara, Japan), 200 µM dNTP mixture, 0.2 µM of each primer and 1.25 unit PrimeSTAR GXL Polymerase per 50 µL and 500ng DNA per reaction. A two-step protocol was performed as 25 cycles of 98°C for 10 s and 68°C for 24 min, final hold at 4°C.The amplicons were purified on 0.5% *w/v* agarose gel in 1X TAE buffer. The long amplicons were purified by Agencourt® AMPure® XP beads (Beckman Coulter, High Wycombe, UK) to ensure removal of primer dimers, polymerase and free nucleotides. The samples were eluted in 50 µL nuclease-free water. Purified amplicons were quantified by Qubit® dsDNA Broad-Range assay kit (Life Technologies) to allow the starting concentration of the sequencing libraries to be 100 ng. Following quantification, enzymatic fragmentation was completed using the Ion Xpress™ Plus Fragment Library Kit (Life Technologies) resulting in fragments of ~200bp. Next, the fragments were ligated with barcoded adapters, which add about 70bp to the fragments. P1 and Ion Xpress™ Barcode X adapters from the Ion Xpress™ Barcode Adapters Kit (Life Technologies) were used to distinguish the samples when pooled prior to sequencing. The adapter-ligated library was size selected by SPRIselect® reagent kit (Beckman Coulter, High Wycombe, UK). After

each step (fragmentation, ligation and size selection), purification was conducted using magnetic beads and the integrity, size distribution, concentration and quality of the library in those steps was checked using the Agilent® 2100 Bioanalyzer® instrument and Agilent High Sensitivity DNA Kit (Agilent Technologies UK Limited, Stockport, UK).

Template-positive ion sphere particles (ISPs) containing clonally amplified DNA were prepared by the Ion Personal Genome Machine™ (PGM™) Template OT2 200 Kit (for 200 base-read libraries) (Life Technologies) with the Ion OneTouch™ 2 System. Then the percentage of template-positive ISPs was checked by the Ion Sphere™ Quality Control assay (Life Technologies, Paisley, UK) on the Qubit® 2.0 Fluorometer (Life Technologies) and then enriched by the Ion OneTouch™ ES Instrument before loading onto a 316™ chip. Sequencing was carried out using the Ion PGM™ Sequencing 200 Kit v2 (Life Technologies) and the Ion Torrent PGM™.

Bioinformatics for *RHD* LR-PCR

Torrent Suite™ Software Version 4.4 was utilised in order to generate a summary sequencing report indicating the number of reads generated by the sequencer, the percentage of chip loading and the sequencing files. The FastQC software was run to assess the quality control across the reads generated (17). The sequencing samples were aligned to the human genome reference sequence (hg19) using the Binary Alignment/Map (BAM) and were visualised using Integrative Genome Viewer (IGV) Version 2.3.46.

The samples were annotated using the Variant Call Format (VCF) files to obtain the SNPs and indels to analyse the genotype and predict the phenotype. Antigens were determined by choosing the right transcript according to the Blood Group Antigen

Factsbook (18). Each antigen was determined by its chromosomal location, the type of variant (SNP or indel), gene, the reference nucleotide, the changing nucleotide, depth of coverage, the transcript used in analysis based on the NCBI database, the location of the variant (intronic or exonic), codon, an exon number of that variant, an amino acid substitution and the position of the nucleotide change. The SeattleSeq Annotation tool 141 site was used to annotate the sequencing data of the LR-PCR approach (19). By using Browser Extensible Data (BED) files, the bedtools website was used to mask the *RHCE* gene in order to analyse the *RHD* gene (20). The *RHCE* gene was annotated by 'Ns' on its sequencing nucleotides.

RHD **Intronic SNP sequencing**

gDNA samples from blood donors of different phenotypes were tested. *RHD*-specific primers amplified the regions around the intronic SNPs (Table 2). Two different enzymes were used, BioMix™ 2X master mix (Bioline Reagents Limited, United Kingdom) or Q5® Hot Start High-Fidelity 2X Master Mix, (New England Biolabs, United Kingdom). A 50 µL PCR reaction was prepared containing 1X master mix, 200ng of DNA template, 1 µM of each of the primers. Cycling was carried out on a Veriti Thermal Cycler (Life Technologies) following optimised conditions; 95°C for 10 min, 35 cycles of 95°C for 30 s and optimised annealing temperature for 1 min, 72°C for 30 s, followed by a final 72°C step for 10 min. To validate PCR amplification, PCR products were run on a 1% *w/v* agarose gel in 1x TAE buffer. PCR products were purified using the QIAquick Gel Extraction Kit, (Qiagen Ltd, West Sussex, United Kingdom) according to the manufacturer's instructions. PCR amplicons were subjected to Sanger sequencing by Eurofins Genomics. Results were aligned with the human genome reference sequence (hg19). CodonCode Aligner 6.0 software was used to analyse the data.

Results

Determination of *RHD* Zygosity

For zygosity testing, the presence or absence of *RHD* amplification on the ddPCR platform was used to determine whether the samples were *RHD* negative or *RHD* positive, respectively. The mean number of copies per droplet for all molecules was 0.15 (0.03 – 0.57) for plasma DNA samples and 0.39 (0.05 – 0.69) for buffy coat DNA samples. The ratio of *RHD5* (FAM)*/ AGO1 (*HEX) and *RHD7 (*FAM)*/ AGO1 (*HEX) generated by the QuantaSoft v1.2 Software was then used to determine whether the D-positive samples were hemizygous or homozygous for the *RHD* gene. Samples showing ratios close to 1 were determined to be homozygous *RHD* positive and samples with ratios closer to 0.5 were classified as hemizygous *RHD* positive (Figure 2).

The results demonstrated that the assay worked equally well on cell free DNA and gDNA for zygosity determination (Table 3) (Figure 2). Three rr control samples were tested (147J, 1660, 7807) and results demonstrated amplification of only the reference (*AGO1*), giving a ratio of zero (Figure 2). The hemizygous D+ R_0r (Dce/dce) (*n* = 8), R_1r (DCe/dce) (*n* = 12) and R_2r (DcE/dce) (*n* = 1) samples demonstrated ratios close to 0.5 as expected (Table 3) (Figure 2), except for sample 1777. Sample 1777, previously classified by serology as being phenotypically R_1r (DCe/dce), expressed ratios of 0.97 and 1.04 for the *RHD5* and *RHD7* multiplex reactions, respectively (Table 3). This result contradicted previous serological classification and indicated that the sample expressed two copies of the *RHD* gene. Therefore, it is more feasible that this sample actual expresses the R_1R_0 (DCe/Dce) phenotype. The homozygous D+ R_1R_1 (DCe/DCe) (*n* = 13), R_2R_2 (DcE/DcE) (*n* = 5),

R_1R_2 (DCe/DcE) (*n* = 10) and R_2R_Z (DcE/DCE) (*n* = 1) samples were expected to generate a ratio close to 1, and this was achieved in 90% of samples. Sample 087W was serologically typed as expressing the R_2R_2 (DcE/DcE) phenotype. However, the dPCR results demonstrate that this sample is hemizygous for the *RHD* gene, since both assays illustrated a ratio close to 0.5 (Figure 2). Therefore, it is likely that sample 087W has the R_2r'' (DcE/dcE) genotype as opposed to the R_2R_2 (DcE/DcE) serologically predicted genotype. Further sequencing analysis was required to determine the actual genotype of the incorrectly labelled R_1R_2 samples (729M and 351D) (Figure 2).

RHD **intronic polymorphisms**

We sequenced the complete *RHD* gene from individuals with defined *RH* genotypes using LR-PCR (Table 2) and we identified several intronic polymorphisms that closely correlated with the individuals *DCE* status. On further analysis using Sanger sequencing, five SNPs showed complete concordance when scrutinised using primers flanking these regions (Tables 2 and 4).

Comparison of ***RHD*** **intronic polymorphisms and zygosity**

Two of the R_1R_2 (DCe/DcE) presumed genotype samples tested (729M and 351D) expressed ratios close to 0.5 for both assays (Figure 2b). Since sample 729M has also been typed as weak D, it is highly unlikely that this sample is homozygous *RHD* positive. Therefore, it is clear this sample has been misclassified as R_1R_2, but we could not ascertain whether the true genotype for sample 729M was R_2r' (DcE/dCe), R_Zr (DCE/dce), R_0r^y (Dce/dCE) or R_1r'' (DCe/dcE). Consequently, LR-PCR coupled with NGS revealed that sample 729M displayed the exon 9 Gly385Ala 1154G>C SNP, and thus was classified as weak D type 2. In addition, the sample illustrated

multiple *RHD* intronic SNPs which appear to be associated with the R_2 (DcE) haplotype, which demonstrates that sample 729M is likely to be R_2r' (DcE/dCe) (Table 4). Sample 351D was not typed serologically as weak D but the dPCR data shows that only one copy of *RHD* is present (Figure 2b) and thus the genotype must either be R_2r' (DcE/dCe), R_Zr (DCE/dce), R_0r^y (Dce/dCE) or R_1r'' (DCe/dcE). This sample did not show the R_2 associated *RHD* intronic SNPs and hence is likely to have a genotype of R_1r'' (DCe/dcE), R_0r^y (Dce/dCE) or R_Zr (DCE/dce).

Discussion

RHD zygosity assignment has proved to be a useful diagnostic tool in the clinical management of HDFN. Here, determination of homozygous (*RHD/RHD*) fathers would give confidence (assuming paternity) of prenatal prediction of D-positive fetuses, and signal where further monitoring or administration of prophylactic anti-D maybe required. Without doubt, the most appropriate technique would be the assessment of D-positive infants directly by analysis of free fetal DNA in maternal plasma. However, in repeat pregnancies fathered by *RHD/RHD* homozygotes maternal plasma testing would not be necessary as the fetus would invariably be D-positive. This is of course with the caveat that paternity can be assured during the maternal consenting process. Previous methods have utilised qPCR (4, 7-9), MLPA (6), mass spectrometry (10) and analysis of the *Rhesus box* (2, 5, 11). However, as we have previously mentioned, individuals have been described that confound zygosity testing when relying on analysis of the *Rhesus box* repeat sequences (11). Here we describe a rapid and accurate further method for defining *RHD* zygosity. We have used this on a small cohort of phenotyped blood samples and demonstrated that this method could be used effectively to define paternal zygosity,

and in addition, to correct presumed phenotype in blood donors which is presently dependent on phenotype prediction.

In three samples we have analysed, and a weak D sample, we have clearly demonstrated homo and hemizygosity for *RHD*, which was not in concordance with predicted phenotype. The vast majority of current genotyping methods (22-27) are not able to define zygosity (except the study by Gassner *et al.* (10) or unless an assessment of intronic *RHD*-specific SNPs is performed, some of which are described in this paper). Our description of candidate SNPs that define the *RHD* gene within the DcE haplotype will also provide a method to differentiate homo or hemizygosity, and we have candidate *RHD* intronic SNPs that define the DCe and Dce *RHD* genes (in preparation). However, much more work on a larger number of donors (including the testing of rare RH haplotypes) has to be done before these candidate *RHD* intronic SNPs can be confirmed as being truly DCe and Dce specific. Nevertheless, these *RHD* intronic SNPs may not be able to differentiate between DcE/DcE and DcE/dcE; DCe/DCe and DCe/dCe; and Dce/Dce and Dce/dce genotypes, however, the dPCR method described here is able to facilitate this (differentiating homo and hemizygosity). Clearly, for these candidate SNPs to have clinical utility, a larger cohort of phenotyped samples will require sequencing. We have subsequently performed such an analysis on 37 Rh phenotyped individuals, and have found complete concordance with the five DcE-associated candidate SNPs described in this study. We have identified a further 11 such candidate SNPs that also are in concordance with DcE genotype. (WAT, KAS, AJH, MK, TEM and NDA, manuscript in preparation). We are currently investigating a number of Rh variants and rare phenotypes (e.g. R_Z) to assist in their identification.

This method provides a quick and accurate platform for rapid determination of *RHD* zygosity. In this small cohort of samples, we would be unlikely to see rare haplotypes such as DCE, dCe, dcE and Dce. However, both dCe and dcE haplotypes were identified. Further zygosity-based studies are clearly necessary to reassess the population frequencies of these D-negative haplotypes. It is important also to consider that fathers that are *RHD* hemizygous DCe/dCe or DcE/dcE may pass the dCe or dcE haplotypes to their children, and these fetuses may be at risk of HDFN due to anti- C or G (28) or anti-E (29). Fetal genotyping for inheritance of both Rh C and Rh E has been routinely performed using maternal plasma and should therefore be used in such cases where hemizygosity has been defined. We believe that the method we describe here is a useful addition to the diagnostic repertoire available to the clinician in the management of HDFN.

References

1. Avent ND, Reid ME. The Rh blood group system: a review. Blood 2000;95(2):375-87.

2. Wagner FF, Flegel WA RHD gene deletion occurred in the Rhesus box. Blood 2000;95: 3662-8.

3. The Rhesusbase. The RHD Mutation Database [Online]. Available at: http://www.rhesusbase.info/ (Accessed October 2016).

4. Chiu RW, Murphy MF, Fidler C, Zee BC, Wainscoat JS, Lo YD. Determination of RhD zygosity: comparison of a double amplification refractory mutation system approach and a multiplex real-time quantitative PCR approach. Clin Chem. 2001;47(4):667-72.

5. Grootkerk‐Tax MG, Maaskant‐van Wijk PA, Van Drunen J, Der Schoot V, Ellen C. The highly variable RH locus in nonwhite persons hampers RHD zygosity determination but yields more insight into RH‐related evolutionary events. Transfusion 2005;45(3):327-37.

6. Haer‐Wigman L, Veldhuisen B, Jonkers R, Lodén M, Madgett TE, Avent ND, et al. RHD and RHCE variant and zygosity genotyping via multiplex ligation‐dependent probe amplification. Transfusion 2013;53(7):1559-74.

7. Perco P, Shao CP, Mayr WR, Panzer S, Legler TJ. Testing for the D zygosity with three different methods revealed altered Rhesus boxes and a new weak D type. Transfusion 2003;43(3):335-9.

8. Pirelli KJ, Pietz BC, Johnson ST, Pinder HL, Bellissimo DB. Molecular determination of RHD zygosity: predicting risk of hemolytic disease of the fetus and newborn related to anti‐D. Prenatal Diag 2010;30(12‐13):1207-12.

9. Krog GR, Clausen FB, Dziegiel MH. Quantitation of RHD by real-time polymerase chain reaction for determination of RHD zygosity and RHD mosaicism/chimerism: an evaluation of four quantitative methods. Transfusion 2007;47: 715-22.

10. Gassner C, Meyer S, Frey BM, Vollmert C Matrix-Assisted Laser Desorption/Ionisation, Time-of-Flight Mass Spectrometry–Based Blood Group Genotyping—The Alternative Approach. Transfusion Medicine Reviews 2013;27(1): 2–9.

11. Matheson KA, Denomme GA. Novel 3′ Rhesus box sequences confound RHD zygosity assignment. Transfusion 2002;42(5):645-50.

12. Daniels G. Rh and RHAG Blood Group Systems. In: Daniels G, editor. Human Blood Groups. 3rd Ed. Oxford (UK): Wiley-Blackwell; 2013 p.185. 13. Sillence KA, Roberts LA, Hollands HJ, Ross Welch C, Thompson HP, Kiernan M, et al. Fetal sex and *RHD* genotyping with digital PCR demonstrates greater sensitivity than real-time PCR. Clin Chem 2015;239137.

14. Fan HC, Blumenfeld YJ, Chitkara U, Hudgins L, Quake SR. Analysis of the size distributions of fetal and maternal cell-free DNA by paired-end sequencing. Clin Chem 2010;56(8):1279-86.

15. Finning K, Martin P, Summers J, Massey E, Poole G, Daniels G. Effect of high throughput RHD typing of fetal DNA in maternal plasma on use of anti-RhD

immunoglobulin in RhD negative pregnant women: prospective feasibility study. Brit Med J 2008;336(7648):816-8.

16. Legler TJ, Maas JH, Köhler M, Wagner T, Daniels GL, Perco P, Panzer S. RHD sequencing: a new tool for decision making on transfusion therapy and provision of Rh prophylaxis. Transfusion Med 2001;11:383-388.

17. Andrews, S. *FastQC: a quality control tool for high throughput sequence data* (2010) [Online]. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc [Accessed October 2016].

18. Reid ME, Lomas-Francis C, Olsson ML The Blood Group Antigen FactsBook, 3rd Ed. Boston, Academic Press; 2012.

19. Seattleseq Annotation Tool 141. *SeattleSeq Variation Annotation* [Online]. Available at: http://snp.gs.washington.edu/ (Accessed October 2016).

20. Bedtools. *bedtools: a powerful toolset for genome arithmetic* [Online]. Available at: http://bedtools.readthedocs.org/ (Accessed October 2016).

21. National Center for Biotechnology Information NCBI). dbSNP Short Genetic Variations [Online]. Available at https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs (Accessed October 2016).

22. Hashmi G, Shariff T, Zhang Y, Cristobal J, Chau C, Seul M, et al. Determination of 24 minor red blood cell antigens for more than 2000 blood donors by high-throughput DNA analysis. Transfusion 2007;47(4):736-47.

421 23. Avent ND, Martinez A, Flegel WA, Olsson ML, Scott ML, Nogues N, et al. The
422 Bloodgen project of the European Union, 2003–2009. Transfus Med and Hemoth
423 2009;36(3):162-7.

424 24. Avent ND. Large-scale blood group genotyping–clinical implications. Brit J
425 Haematol 2009;144(1):3-13.

426 25. Flegel WA, Gottschall JL, Denomme GA. Implementing mass-scale red cell
427 genotyping at a blood center. Transfusion 2015;55(11):2610-5.

428 26. Boccoz SA, Le Goff GC, Mandon CA, Corgier BP, Blum LJ, Marquette CA
429 Development and Validation of a Fully Automated Platform for Extended Blood
430 Group Genotyping. The Journal of Molecular Diagnostics 2016;18(1)

431 27. López M, Apraiz I, Rubia M, Piedrabuena M, Azkarate M, Veldhuisen B, Vesga
432 MA, van Der Schoot E, Puente F, Tejedor D Performance evaluation study of ID
433 CORE XT, a high throughput blood group genotyping platform. Blood Transfus. 2016:
434 Nov 25:1-7. doi: 10.2450/2016.0146-16. [Epub ahead of print]

435 28. Hadley A, Poole G, Poole J, Anderson N, Robson M. Haemolytic Disease of the
436 Newborn Due to Anti-G. Vox Sang 1996;71(2):108-12.

437 29. Rouillac C, Kim CLV, Blancher A, Roubinet F, Cartron JP, Colin Y. Lack of G
438 blood group antigen in DIIIb erythrocytes is associated with segmental DNA
439 exchange between RH genes. Brit J Haematol 1995;89(2):424-6.

440

441

442

443

444

445

446 **Table 1:** *RHD5, RHD7* and *AGO1* oligonucleotide sequences, product size and gene

447 location.

Amplicon location	Multiplex Reaction	Primer	Sequence (5' - 3')	Dual-Labelled Hydrolysis Probe (5' - 3')	Length (bp)
1p36.11 *RHD* Exon 5	1	*RHD5* Forward*	CGCCCTCTTCTTG TGGATG	FAM-TCTGGCCAAGTTTCA ACTCTGCTCTGCT-BHQ1	82bp
		RHD5 Reverse*	GAACACGGCATTC TTCCTTTC		
1p36.11 *RHD* Exon 7	2	*RHD7* Forward*	CAGCTCCATCATG GGCTACAA	FAM-AGCTTGCTGGGTCTG CTTGGAGAGATC-BHQ1	75bp
		RHD7 Reverse*	AGCACCAGCAGCA CAATGTAGA		
1p34.3	1 and 2	*AGO1* Forward**	GTTCGGCTTTCAC CAGTCT	HEX-CTGCCATGTGGAAGATG ATG -BHQ1	81bp
		AGO1 Reverse**	CTCCATAGCTCTC CCCACTC		

448 *Taken from Finning *et al.* (15)

449 **Taken from Fan, *et al.* (14)

450

Table 2: *RHD* intronic SNP and *RHD* long-range PCR oligonucleotide sequences, product sizes and corresponding SNP in the *RHD* gene (hg19 human genome reference sequence, for intronic SNPs).

Intronic SNPs	**Intron**	**Forward Primer Sequence 5'-3'**	**Reverse Primer Sequence 5'-3'**	**Length (bp)**
25,611,580 G>A	2	TTTTACTGGACAGCCCTACTCC	CATGGCTATTTATTGTCTAGCAGCA	558
25,614,400 C>G	2	GCTACCATGCCCTGCTAAT	TCCAGTACTTTTCAGAGCC	417
25,625,471 T>C	3	GGGGCAGCTTCATCTTATCAAGAG	CTCACTGCAACCTCCACCCGTT	419
25,627,066 C>G	3	TGGGATTACAGGCAAAATTAG	AGGTGTGACTTGAAGCCAT	834
25,648,349 T>C	8	TCCAGGAATGACAGGGCT	TGAGGACTGCAGATAGGG	525
***RHD* exons covered**				
1-3	1,2	GATTGGGTCCGTGATTGGCATT	GGCCGCGGGAATTCGATTGTTGTCTTTATTTTTCAAAACCCT	22,829
2-7	2-6	GCCGCGAATTCACTAGTGTGACGAGTGAAACTCTATCTCGAT (Ds2-s*)	GGCCGCGGGAATTCGATTGAGGCTGAGAAAGGTTAAGCCA	23,610
7-10	7-9	GCCGCGAATTCACTAGTGACAAACTCCCCGATGATGTGAGTG	GGCCGCGGGAATTCGATTGTGGTACATGGCTGTATTTTATTG	22,731

*Adapted from Legler *et al.* (16)

458 **Table 3:** Zygosity testing results determined by ratio analysis for DNA samples
459 extracted from both the plasma (cfDNA) and buffy coat (gDNA) of human whole
460 blood samples.

Sample	*RH* [1]	Ratio (*RHD5* (FAM) / *AGO1* (HEX))	Ratio (*RHD7* (FAM) / *AGO1* (HEX))	Hemizygous or homozygous	Genotype determined by dPCR**
147J*	rr (dce/ dce)	0	0	Homozygous *RHD* negative	rr (dce/ dce)
1660*		0	0	Homozygous *RHD* negative	rr (dce/ dce)
7807*		0	0	Homozygous *RHD* negative	rr (dce/ dce)
9763*	R_0r (Dce/ dce)	0.45	0.43	Hemizygous	R_0r (Dce/ dce)
069F*		0.5	0.49	Hemizygous	R_0r (Dce/ dce)
740B*		0.47	0.46	Hemizygous	R_0r (Dce/ dce)
258D*		0.51	0.51	Hemizygous	R_0r (Dce/ dce)
(079*)*		0.51	0.50	Hemizygous	R_0r (Dce/ dce)
649B*		0.5	0.5	Hemizygous	R_0r (Dce/ dce)
8931*		0.49	0.49	Hemizygous	R_0r (Dce/ dce)
5784*		0.49	0.50	Hemizygous	R_0r (Dce/ dce)
065S*	R_1r (DCe/ dce)	0.49	0.49	Hemizygous	R_1r (DCe/ dce)
118Z*		0.5	0.49	Hemizygous	R_1r (DCe/ dce)
1226*		0.52	0.51	Hemizygous	R_1r (DCe/ dce)
1306*		0.51	0.53	Hemizygous	R_1r (DCe/ dce)
1777*		0.97	1.04	Homozygous *RHD* positive	R_1R_0 (DCe/ Dce)
180H*		0.52	0.52	Hemizygous	R_1r (DCe/ dce)
181F*		0.52	0.49	Hemizygous	R_1r (DCe/ dce)
148R [2]		0.50	0.50	Hemizygous	R_1r (DCe/ dce)
6418 [2]		0.51	0.49	Hemizygous	R_1r (DCe/ dce)
3093		0.51	0.51	Hemizygous	R_1r (DCe/ dce)
572R [2]		0.50	0.50	Hemizygous	R_1r (DCe/ dce)
7687 [2]		0.50	0.51	Hemizygous	R_1r (DCe/ dce)
5481 [2]	R_2r (DcE/ dce)	0.50	0.51	Hemizygous	R_2r (DcE/ dce)
1220*	R_1R_1 (DCe/ DCe)	0.98	1.01	Homozygous *RHD* positive	R_1R_1 (DCe/ DCe)
131Z*		0.99	1.04	Homozygous *RHD* positive	R_1R_1 (DCe/ DCe)
165F*		0.94	0.9	Homozygous *RHD* positive	R_1R_1 (DCe/ DCe)
1793*		0.99	1	Homozygous *RHD* positive	R_1R_1 (DCe/ DCe)
0670*		0.91	0.85	Homozygous *RHD* positive	R_1R_1 (DCe/ DCe)
1347*		0.99	1.03	Homozygous *RHD* positive	R_1R_1 (DCe/ DCe)
138R*		0.95	0.98	Homozygous *RHD* positive	R_1R_1 (DCe/ DCe)
052M		0.99	1.03	Homozygous *RHD* positive	R_1R_1 (DCe/ DCe)
247X		1.02	1.01	Homozygous *RHD* positive	R_1R_1 (DCe/ DCe)
078U		0.99	1.01	Homozygous *RHD* positive	R_1R_1 (DCe/ DCe)
103N		1.01	1.03	Homozygous *RHD* positive	R_1R_1 (DCe/ DCe)
1461		0.99	1.01	Homozygous *RHD* positive	R_1R_1 (DCe/ DCe)
877L		1.01	0.98	Homozygous *RHD* positive	R_1R_1 (DCe/ DCe)
658G	R_2R_2 (DcE/ DcE)	1.02	1.03	Homozygous *RHD* positive	R_2R_2 (DcE/ DcE)
738W		1.02	1.04	Homozygous *RHD* positive	R_2R_2 (DcE/ DcE)
087W		0.51	0.49	Hemizygous	R_2r'' (DcE/ dcE)

132H		1.01	1.03	Homozygous *RHD* positive	R_2R_2 (DcE/ DcE)
689U		0.99	1.01	Homozygous *RHD* positive	R_2R_2 (DcE/ DcE)
729M [2]		0.50	0.49	Hemizygous	R_1r'' (DCe/ dcE) or R_2r' (DcE/ dCe)
896H		0.98	1.03	Homozygous *RHD* positive	R_1R_2 (DCe/ DcE)
898D		0.99	0.97	Homozygous *RHD* positive	R_1R_2 (DCe/ DcE)
351D		0.51	0.51	Hemizygous	R_1r'' (DCe/ dcE) or R_2r' (DcE/ dCe)
9316	R_1R_2 (DCe/ DcE)	1.02	1.01	Homozygous *RHD* positive	R_1R_2 (DCe/ DcE)
911E		1.02	1.03	Homozygous *RHD* positive	R_1R_2 (DCe/ DcE)
4195		1.02	1.01	Homozygous *RHD* positive	R_1R_2 (DCe/ DcE)
645C		1.06	1.03	Homozygous *RHD* positive	R_1R_2 (DCe/ DcE)
3627		0.99	1.01	Homozygous *RHD* positive	R_1R_2 (DCe/ DcE)
8873		1.02	1.03	Homozygous *RHD* positive	R_1R_2 (DCe/ DcE)
746P	R_2R_Z (DcE/ DCE)	1.02	0.99	Homozygous *RHD* positive	R_1R_2 (DCe/ DcE)

[1] Serologically predicted phenotype provided by National Health Service Blood and Transplant (NHS BT) (Bristol, UK).

[2] Sample is Weak D.

*DNA samples tested from plasma.

** The C/c and E/e status based on serological information. Only the D/d genotype was corrected by dPCR.

Table 4: *RHD* intronic SNP sequencing and *RHD* LR-PCR NGS results for a range of DNA samples.

Intronic SNPs in *RHD* (R_1/R_1 to R_2)	***RHD* Intron**	**R_1R_1 (DCe/ DCe) (n=4)**	**R_1r (DCe/ dce) (n=1)**	**R_0r (Dce/ dce) (n=8)**	**R_2R_2 (DcE/ DcE) (n=6)**	**R_2r (DcE/ dce) (n=1)**	**R_1R_2 (DCe/ DcE) (n=5)**	**Sample 729M (n=1)**	**Sample 351D (n=1)**
25,611,580 G>A	2	G/G	G	G	A/A	A	G/A	A	G
25,614,400 C>G rs28718098*	2	C/C	C	C	G/G	G	C/G	G	C
25,625,471 T>C rs2904843*	3	T/T	T	T	C/C	C	T/C	C	T
25,627,066 C>G rs2986167*	3	C/C	C	C	G/G	G	C/G	G	C
25,648,349 T>C rs28669938*	8	T/T	T	T	C/C	C	T/C	C	T

* Taken from the National Center for Biotechnology Information (NCBI) (16).

The table indicates the serologically inferred genotype of the samples provided by the National Health Service Blood and Transplant (NHS BT) (Bristol, UK).

Figure 1: LR-PCR products for the Rh blood group system. Three long-range amplicons (1, 2, and 3) were designed to amplify the entire *RHD* gene. (a) An RhD-negative sample shows no bands for the *RHD* LR-PCR in lanes 1, 2 and 3, which represent the three amplicons. (b) An RhD-positive sample gives amplification of all three products, with each product being about 22 kb.

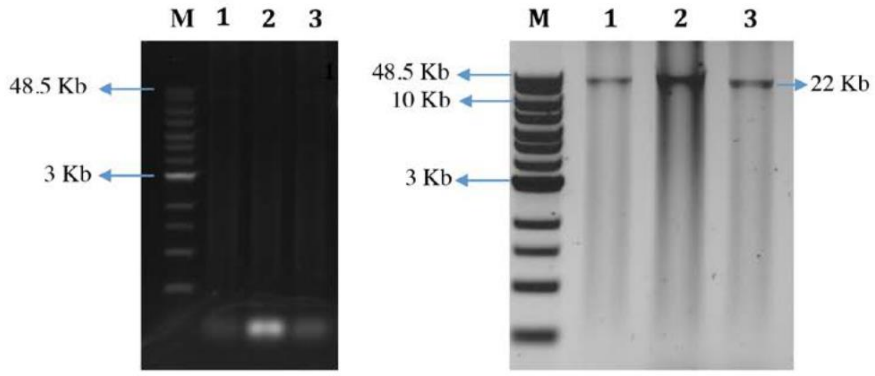
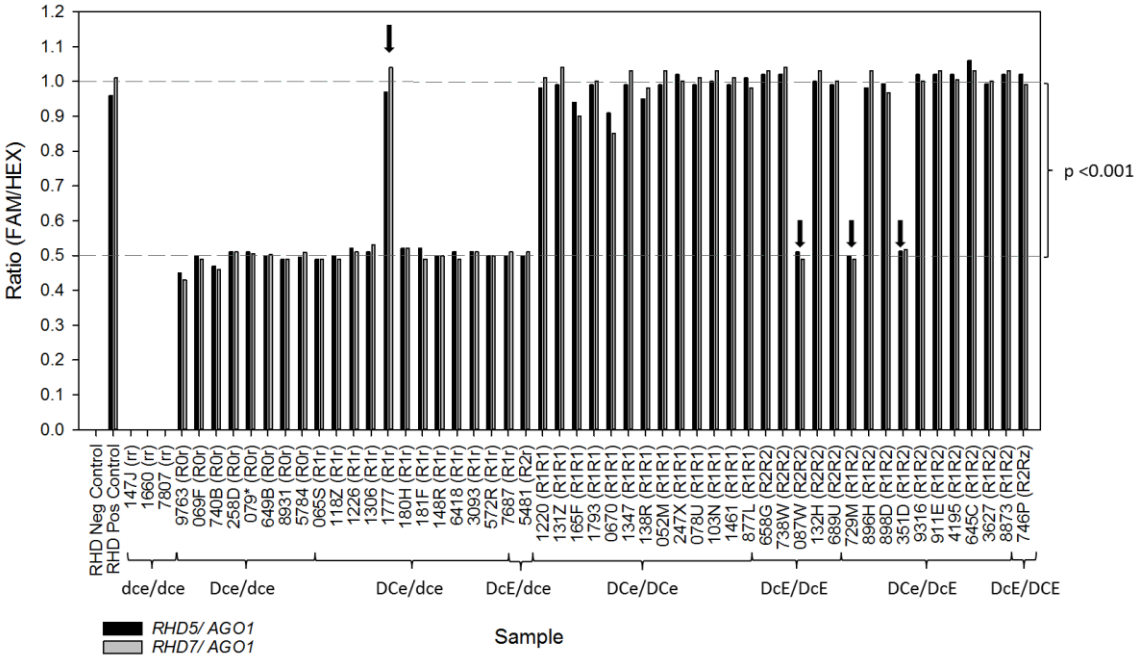
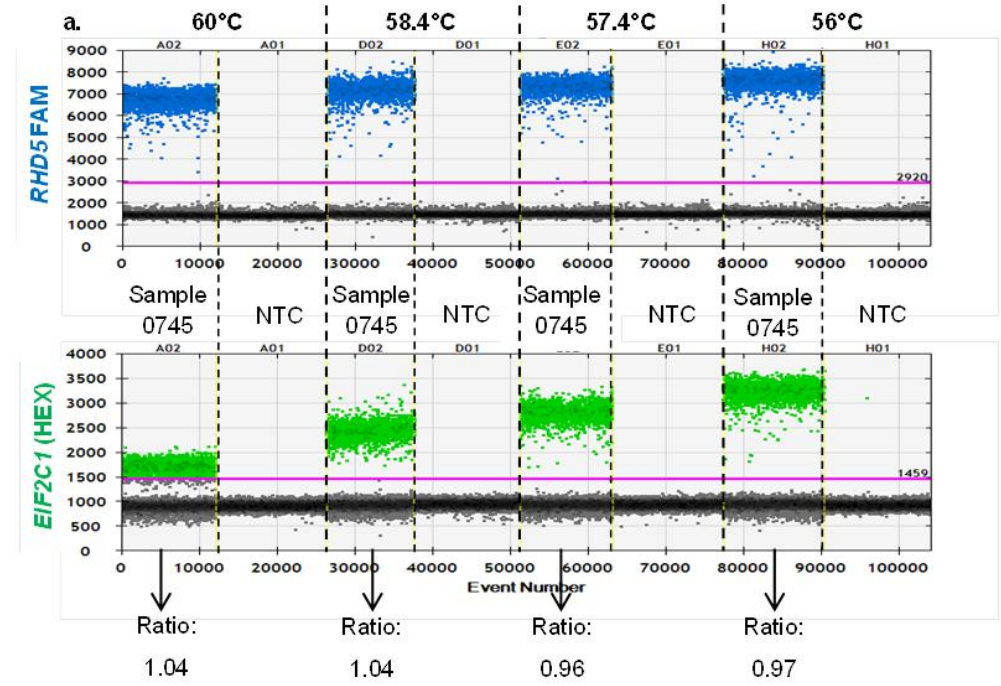
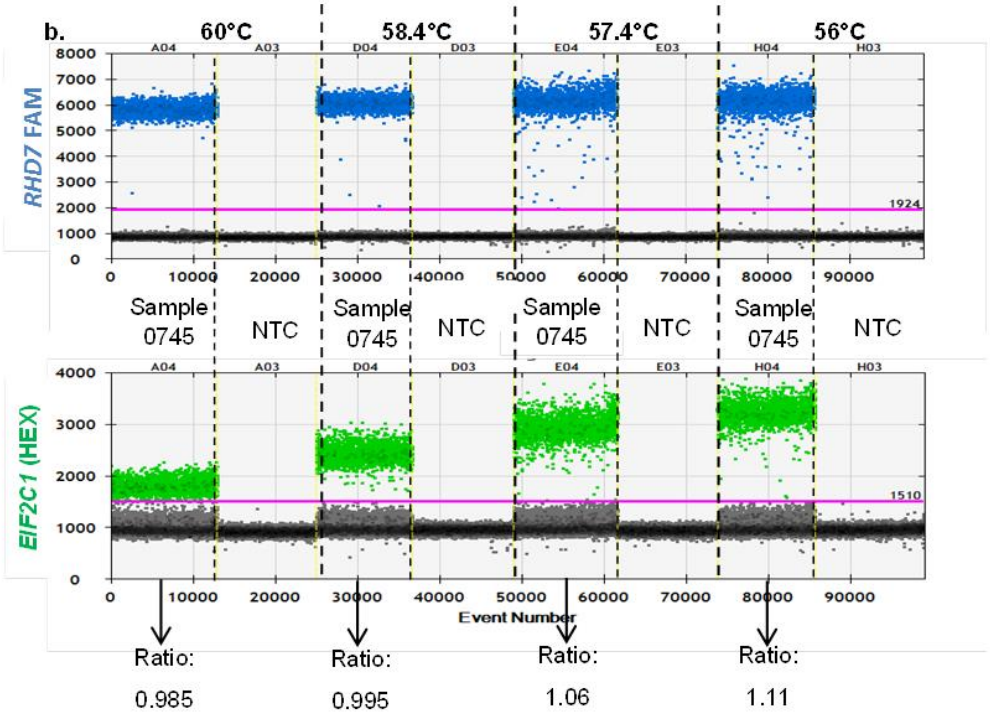


Figure 2: Ratio analysis to determine zygosity using two multiplex reactions (*RHD5* (FAM)/*AGO1* (HEX) and *RHD7* (FAM)/*AGO1* (HEX) for samples with varying Rh phenotypes. The grey dotted lines at 0.5 and 1 on the y axis represent the ratio generated by hemizygous D+ samples and homozygous D+ samples, respectively. The mean ratio for hemizygous and homozygous D+ positive samples for both plasma and buffy coat extracted samples (Table 3) illustrated significant difference ($p<0.001$). The arrows indicate the samples that illustrated discordant results compared with the serologically predicted genotype.



Supplemental Figure 1: Annealing temperature gradient (60°C, 58.4°C, 57.4°C and 56°C) of dPCR for both multiplex reactions (*RHD5* (FAM)/*AGO1* (HEX) and *RHD7* (FAM)/*AGO1* (HEX)) using sample 0745 (homozygous for *RHD*) extracted from human whole blood and a NTC. a) Optimisation of the *RHD5* (FAM) and *AGO1* (HEX) multiplex reaction. The results illustrate that separation is visible for both targets at all annealing temperatures. Marginally greater separation is visible at 56°C, but all annealing temperatures are no more than 0.03-0.04 away from a ratio of 1. b) Optimisation of the *RHD7* (FAM) and *AGO1* (HEX) multiplex reaction. The results illustrate that separation is visible for both targets at all annealing temperatures. However, optimal separation was determined to be 58.4°C, since this temperature expressed a ratio closer to 1 (0.995).


a.
60°C
58.4°C
57.4°C
56°C
RHD5FAM
EIF2C1 (HEX)
Sample 0745
NTC
Event Number
Ratio: 1.04
Ratio: 1.04
Ratio: 0.96
Ratio: 0.97


b.
60°C
58.4°C
57.4°C
56°C
RHD7 FAM
EIF2C1 (HEX)
Sample 0745
NTC
Event Number
Ratio: 0.985
Ratio: 0.995
Ratio: 1.06
Ratio: 1.11

503