## Supplemental Methods

Blood donor samples

We made use of 9,041 DNA samples from British and Dutch blood donors over the course of this study. The samples and data from all participants were obtained after informed consent, see specific panel descriptions for more detail.

***Test set: NIHR BioResource panel (n=507).*** The NIHR BioResource is a panel of over 100,000 volunteers (NIHR BioResource – Research Tissue Bank, Research Ethics Committee reference: 17/EE/0025)**,** with and without health problems. Data is made available to enable studies on the association between phenotype and genotype. Approximately 10,000 members of the NIHR BioResource were NHSBT blood donors and clinical antigen typing data was retrieved from NHSBT databases. We selected DNA from 507 participants with the most antigen typing data available as a test panel for validation of our first iteration of a donor typing array. This was manufactured using the Axiom 384HT platform and called the Applied Biosystems™ Axiom™ Blood Typing SNP Screen Array (Blood Typing 384HT array).

***Trial set: COMPARE study panel (n=4,795).*** The COMPARE study enrolled 29,066 British blood donors between February 2016 and March 2017 (Comparison of NHSBT's current approach with three alternative strategies to assess haemoglobin levels in whole blood donors (Research Ethics Committee reference: 11/EE/0335).[1] The study aim is to find the optimum technology for haemoglobin screening. All participants were active blood donors and clinical antigen typing data was retrieved from NHSBT databases. The 4,795 participants used in this study were selected based on also being participants in the NIHR BioResource.

***Trial set: Donor InSight III panel (n=2,682).*** The Donor InSight-III (DIS-III) enrolled 3,046 Dutch blood donors to form a research panel to allow scientific insight into donor characteristics, motivations and health (METC 2014/124, NL47865.018.14).[2] Additionally, 95 newly registered donors were enrolled between May 2017 and August 2017 using the DIS-III study protocol. All participants were active blood donors and clinical antigen typing data was retrieved from Sanquin databases. The 2,682 participants used in this study were selected based on availability of extracted DNA samples.

The COMPARE and DIS-III panels, totalling 7,477 British and Dutch blood donors, were used for validation of the final donor typing array presented in this study, named the Applied Biosystems UK Biobank – version 2 Axiom Array (UKBBv2 array).

Final modifications to the bloodTyper analysis algorithm

The custom interpretive blood typing software bloodTyper utilises a curated antigen allele database, to infer antigen status from genomic data.[3,4] Originally developed to infer RBC antigen typing from WGS data, several adaptations were required to enable array data processing. To evaluate file formats and data integrity, genotypes from 100 samples in the test-set (including 2 control samples: NA19315 and NA19318) were run through the analysis pipeline. Several issues with genotype file formats were identified, and genotype data was subsequently reformatted to comply with Variant Call Format (VCF) 4.2.[5] Furthermore, algorithms were introduced to handle specific technical events, for example, in the event a probeset fails and therefore no genotype

is called for an important ABO O phenotype variant such as c.261delG, no ABO antigen prediction is made for safety.

Following these changes, a final round of pipeline compatibility testing was performed on all samples in the test-set (including 17 repeat control samples, 9x NA19315 and 8x NA19318) to verify that all previous identified compatibility issues had been fixed and no new ones were identified.

Reformatting of clinical HLA typing data for the DIS-III cohort

Clinical HLA typing data for DIS-III participants was recorded using several different formats in the Sanquin database and required formatting before use. Where a multiple allele code was used (e.g. 07:GS), a lookup table developed and maintained by NHSBT was used to generate an allele string (e.g. *07:01*07:03*07:040). Where an allele name ended in G, the result was not changed (e.g. A*01:01:01G). A “G” code indicates groups of alleles which share the same sequence in the peptide binding groove, many of the member alleles in each group only differ in the third field which was not typed in this study. Where an allele had the XX suffix (e.g. *02:XX) the result was converted into a single field result (e.g. *02). The XX suffix indicates that the second field can be any allele within the primary group making any second field comparison unreliable.

**Supplemental results**

Allele frequency validation of the UKBBv2 array

A total of 7,984 DNA samples from COMPARE blood donors were genotyped using the UKBBv2 array for 789,550 DNA variants, 10,923 of which code for HLA, RBC and HPA antigens. We compared the GRCh37 alternative allele frequencies (AAF) of the genotyped variants with the corresponding AAFs measured by WGS in 8,510 European ancestry participants in the NIHR BioResource pilot phase of the 100,000 Genomes Project.[8] No significant difference in AAF was detected for 99·02% of the 716,102 variants with MAF>0·1% in both WGS and array data at a Bonferroni adjusted critical threshold ($\alpha$=0.05/716,102, **Figure S1a,b**). Frequency comparison data for HLA, and RBC/HPA antigen typing variants are shown in **Figure S1b-d**. For a small number (0·08%) of total variants measured AAFs differed significantly, which prompted a visual inspection of genotype call plots for all antigen typing variants (**Figure S1d**). Three categories of probe-set were identified in this final QC analysis: *good* (n=1,523), *requires improvement* (n=12) and *poor* (n=37). Data generated by poorly performing probe-sets were excluded from further analysis (**Table S2)**.

Call-plot based probe-set quality control

As a result of genotype call plot inspection, no adequately performing probe-set could be identified for 37 variants included specifically for antigen typing. All of these variants were filtered during genotype QC and did not require manual removal. Expert review deemed that none of the failed variants were critical for donor typing, however, seven were identified as highly desirable. All 37 variants were flagged for redesign on the next iteration of array content (**Table S2)**.

Reasons for excluding antigens and samples from concordance analysis

In total antibody-based typing data was available for 59 RBC and HPA antigens. Antigens for which fewer than 10 comparisons between clinical and genetically-inferred antigen types were possible due to lack of clinical typing data were excluded from concordance analysis. We also excluded $Le^a$ and $Le^b$ antigens of the Lewis system from concordance analysis because anti-Le antibodies are not clinically significant and the ISBT table required for variant interpretation is lacking.[9] P1 antigen typing was disabled in bloodTyper as the molecular basis of this antigen was not clearly defined at the time of platform design.[10]

Following genotype QC, four samples with gender mismatch and samples with more than four non-concordant antigen types per individual were also removed prior to final analysis (**Table S4**). Further investigation revealed that discordances in these samples were caused by erroneous handling of the research-grade blood samples and/or corresponding DNA samples. Furthermore 11 samples in the DIS-III cohort had gender mismatches and more than five mismatches between clinical and UKBBv2 array antigen typing results. Further investigation revealed that a clerical error was made in the clinical data export for these samples, this error was corrected, and samples re-included.

Further explanations of discordance between clinical and UKBBv2 array antigen typing results

In addition to the examples previously discussed in the main text, several sources of discordance between clinical and array antigen typing results are complex and require some further discussion.

*Additional algorithmic discordances*

For three samples with ABO group O clinical phenotype results we observed group A array genotyping results. In array genotyping data for these samples we observed the classical A haplotype (*ABO*A1.01*) and classical O phenotype (ABO*O.01.01). Upon sequencing of these samples using the targeted platform, we also identified heterozygous genotypes for the variants NM_020469.2:c.646T>A, NM_020469.2:c.681G>A, NM_020469.2:c.771C>T, and NM_020469.2:c.829G>A. Alone, these variants usually encode an $A_{weak}$ phenotype, however, when observed alongside another O phenotype (ABO*O.01.01) these variants underpin a second O phenotype (ABO*O.09.01/2) phenotype.[11] Inspection of raw array genotypes revealed that all above variants were all accurately called by the UKBBv2 array. However, at the time of analysis variant antigen expression (including $A_{weak}$) was disabled in the bloodTyper array analysis workflow and therefore an incorrect result was reported. Variant antigen typing has since been enabled and these discordances were resolved.

Additionally, bloodTyper reported $Jk^b$ positive typing results for two samples that were serologically $Jk^b$ negative. In both cases heterozygous genotypes were observed for the variant NM_015865.7:c.342-1g>a which underpins a $Jk_{null}$ phenotype (*JK*01N.06*). Currently the bloodTyper algorithm does not use this variant to infer phenotype due to lack of haplotype frequency data, instead it issues a warning when the variant is detected that antibody-based confirmation of typing results is required.

*Array issues*

18 discordant results were due to array issues which can be subdivided into three two categories:

*Incorrect genotype calls(n=5):* For three ABO and two RH (e) antigen discordances incorrect genotypes were reported by the UKBBv2 array. In these cases, the genotype call confidence was extremely low. Inspection of call plots revealed these calls sat between cluster boundaries. Increasing genotype call QC thresholds would eliminate these errors by producing no genotype call, for these antigens bloodTyper would then subsequently not infer antigen status for safety reasons. In practice, these samples would be flagged for re-typing or typing via alternative methodology.

*Probeset issues (n=5):* Although all genotype calls underpinning these discrepant results were of high quality, inspection of genotype call plots revealed that the probe sets, whilst performing adequately, require further improvement to increase cluster resolution. The M, N, S and s antigens are those affected by this type of error.

*Lack of optimum typing variants (n=8):* All cases here refer to discrepancies in typing results for the C antigen of the RH system. Currently, C antigen status is inferred using the variant NM_020485.5:c.307C>T, which directly encodes antigen expression. Many DNA-based technologies fail to accurately type the variant at this position due to extreme high homology between the *RHD* and *RHCE* genes, particularly in exon 2 of both genes where this variant is situated. A 109 base pair (bp) insertion in intron 2 of the *RHCE* gene, located at NC_000001.10:25732083-25732084 (GRCh37), which has been classically used for DNA-based C antigen typing has no working probeset on the UKBBv2-version of the array. Using two confirmatory variants for C typing is the best strategy to improve accuracy; in light of this, improved probesets for typing the 109bp insertion have been included in the next version of the array.

*Error in clinical typing data*

Discrepancies in this category are most likely explained by errors in clinical typing data, which can be divided into two categories:

*Individuals with variant antigen expression (n=20):* In all cases, samples were antigen negative according to classical antibody-based typing and antigen positive according to genotype. Inspection of bloodTyper reports revealed that in these cases variants that modify antigen expression were identified (**Table S5**). Presence of these variants means that antigen expression is greatly reduced. In these cases, antibody-based antigen typing accuracy is highly dependent on methods and reagents used and false-negative antigen typing results are possible.

*Error in clinical typing data (n=13):* Negative genotype results with positive antibody-based typing results were observed for seven antigens; $Co^b$, C(X), $Fy^a$, $Fy^b$, HPA-5b, N, and $Wr^a$. Inspection of genotype call plots indicated that the probesets were working correctly. There is evidence to suggest that typing results for these antigens are not always accurate at scale.[12,13]

A full breakdown of discordances observed is given in **Table S5**.

**Supplemental references**


1. Di Angelantonio, E. Comparison of alternative strategies to assess haemoglobin levels in whole blood donors (COMPARE study). *ISRCTN Regist.* (2017). doi:ISRCTN90871183

2. Timmer, T. C. *et al.* Donor InSight: characteristics and representativeness of a Dutch cohort study on blood and plasma donors. *Vox Sang.* (2018). doi:10.1111/vox.12731

3. Lane, W. J. *et al.* Comprehensive red blood cell and platelet antigen prediction from whole genome sequencing: Proof of principle. *Transfusion* **56**, 743–754 (2016).

4. Lane, W. J. *et al.* Automated typing of red blood cell and platelet antigens: a whole-genome sequencing study. *Lancet Haematol.* **5**, e241–e251 (2018).

5. Li, H. *The Variant Call Format (VCF) Version 4.2 Specification*. *Online Resource* (2015). doi:10.1016/j.ymeth.2012.07.021

6. Dilthey, A. T. *et al.* High-Accuracy HLA Type Inference from Whole-Genome Sequencing Data Using Population Reference Graphs. *PLoS Comput. Biol.* **12**, 1–16 (2016).

7. Dilthey, A. *et al.* Multi-Population Classical HLA Type Imputation. *PLoS Comput. Biol.* **9**, (2013).

8. Ouwehand, W. H. Whole-genome sequencing of rare disease patients in a national healthcare system. *bioRxiv* (2019). doi:10.1101/507244

9. ISBT, Red Cell Immunogenetics and Blood Group Terminology. *Acesso em 16 fevereiro 2019* 2 (2019). Available at: http://www.isbtweb.org/fileadmin/user_upload/Red_Cell_Terminology_and_Immunogenetics/Table_of_blood_group_systems_v6_180621.pdf.

10. Westman, J. S. *et al.* Allele-selective RUNX1 binding regulates P1 blood group status by transcriptional control of A4GALT. *Blood* (2018). doi:10.1182/blood-2017-08-803080

11. Yazer, M. H., Hosseini-Maaf, B. & Olsson, M. L. Blood grouping discrepancies between ABO genotype and phenotype caused by O alleles. *Current Opinion in Hematology* (2008). doi:10.1097/MOH.0b013e3283127062

12. Novaretti, M. C., Abreu, S. F., Medeiros, V. R., Dorlhiac-Llacer, P. E. & Chamone, D. A. Genotype duffy and alloimmunization against blood group antigens risk in patients with sickle cell disease. *Transfusion* **50**, 139A-140A (2010).

13. Parvizian, S. M., Heddle, N. M., Athale, U. & Goldman, M. R. Red cell antigen genotyping compared to serological phenotyping in sickle cell disease patients in Canada: Potential for reducing alloimmunization. *Transfusion* (2016). doi:http://dx.doi.org/10.1111/trf.13807

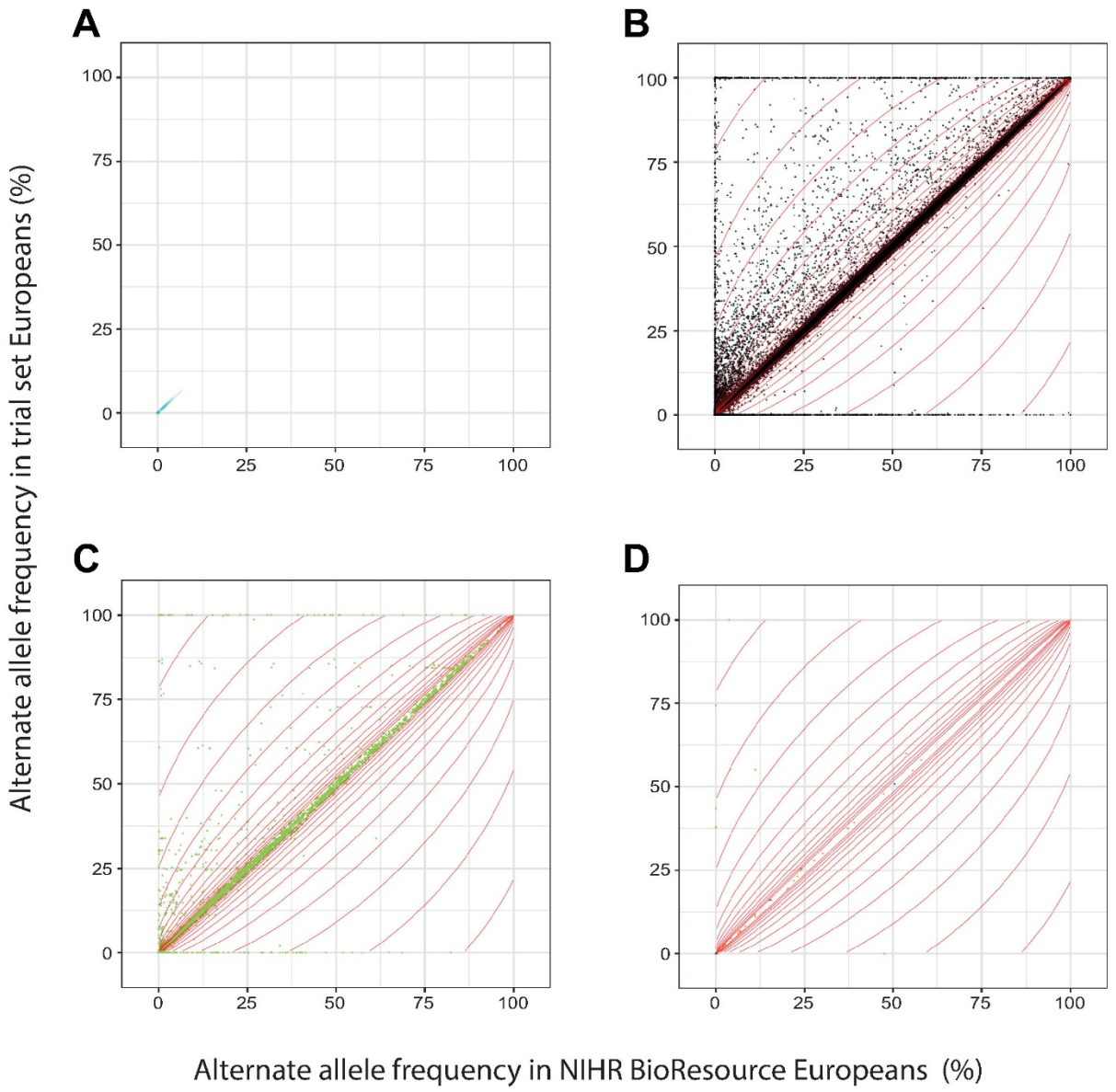


**Figure S1 - Comparison of European ancestry allele frequencies measured by the UKBBv2 array and Illumina short read whole genome sequencing (WGS).** (**A**) A kernel density plot showing excellent average concordance between alternative allele frequency in 4,795 UKBBv2 array typed European ancestry British donors and 8,510 WGS typed European ancestry NIHR BioResource volunteers, across the 709,713 variants with alternative allele frequency >0.1% in both datasets. (**B**) Bivariate scatter of the data used to generate the kernel density plot, emphasising outliers. Each red arc corresponds to the critical threshold for a Pearson test of a given size comparing allele frequencies in the two datasets, under the assumptions that genotypes were observed for all participants and that Hardy-Weinberg equilibrium holds in both datasets. The arcs are uniformly separated on a log-scale. (**C**) Variants in the HLA locus (**green**). (**D**) Variants encoding RBC (**red**) or HPA (**blue**) antigens.

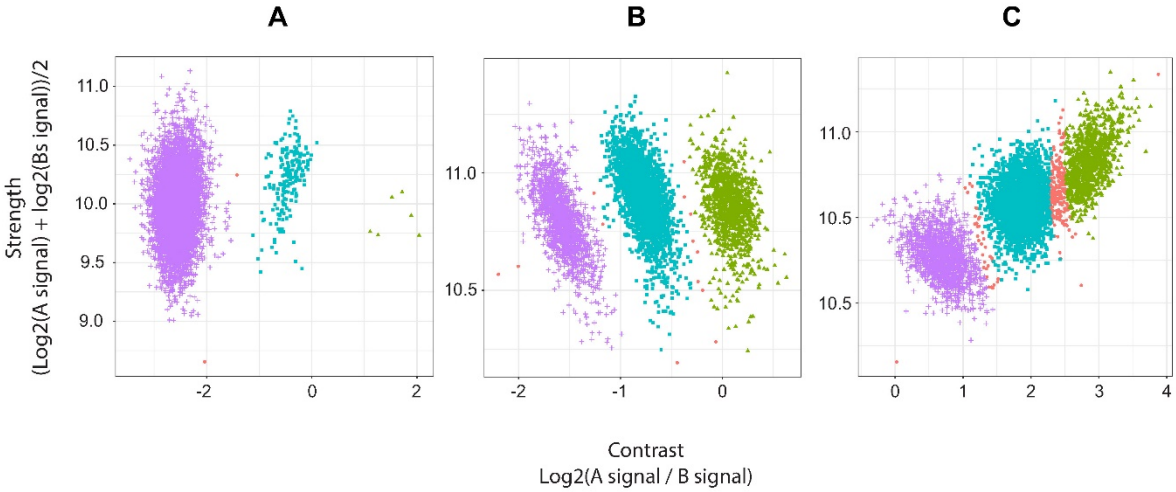


**Figure S2 - Genotype call-plots were visually inspected as part of DNA probe performance assessment**. Examples of (**A**) a “good quality” call-plot with clear separation between genotype clusters; (**B**) a “requires development” call-plot, three unique genotype clusters can be observed, however, boundaries are too close; (**C**) a “poor quality” call-plot with overlap between genotype clusters. Colour represents genotype call: homozygous reference (**purple**), heterozygous (**blue**), homozygous alternate (**green**); failed call (**red**).

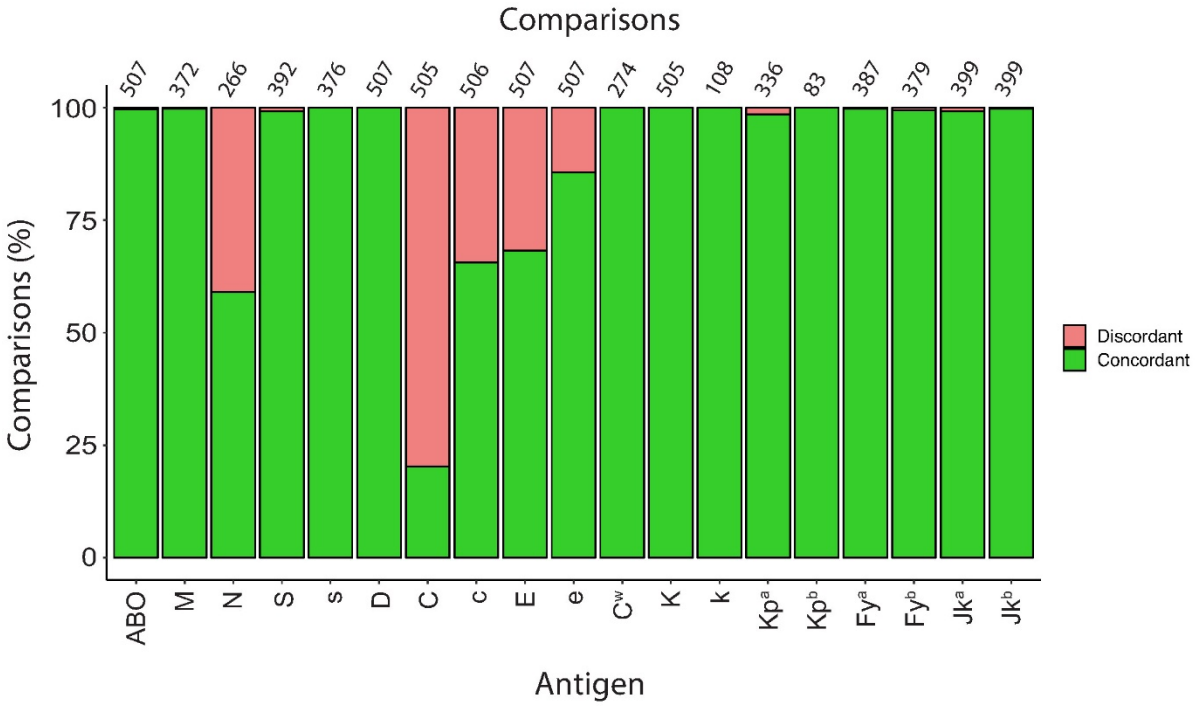


**Figure S3 - Antigen typing capabilities of the test ("donor only content") array.** Concordance per antigen is shown as a percentage of the total number of comparisons (given at the top of each bar) with concordant and discordant results in green and red, respectively.

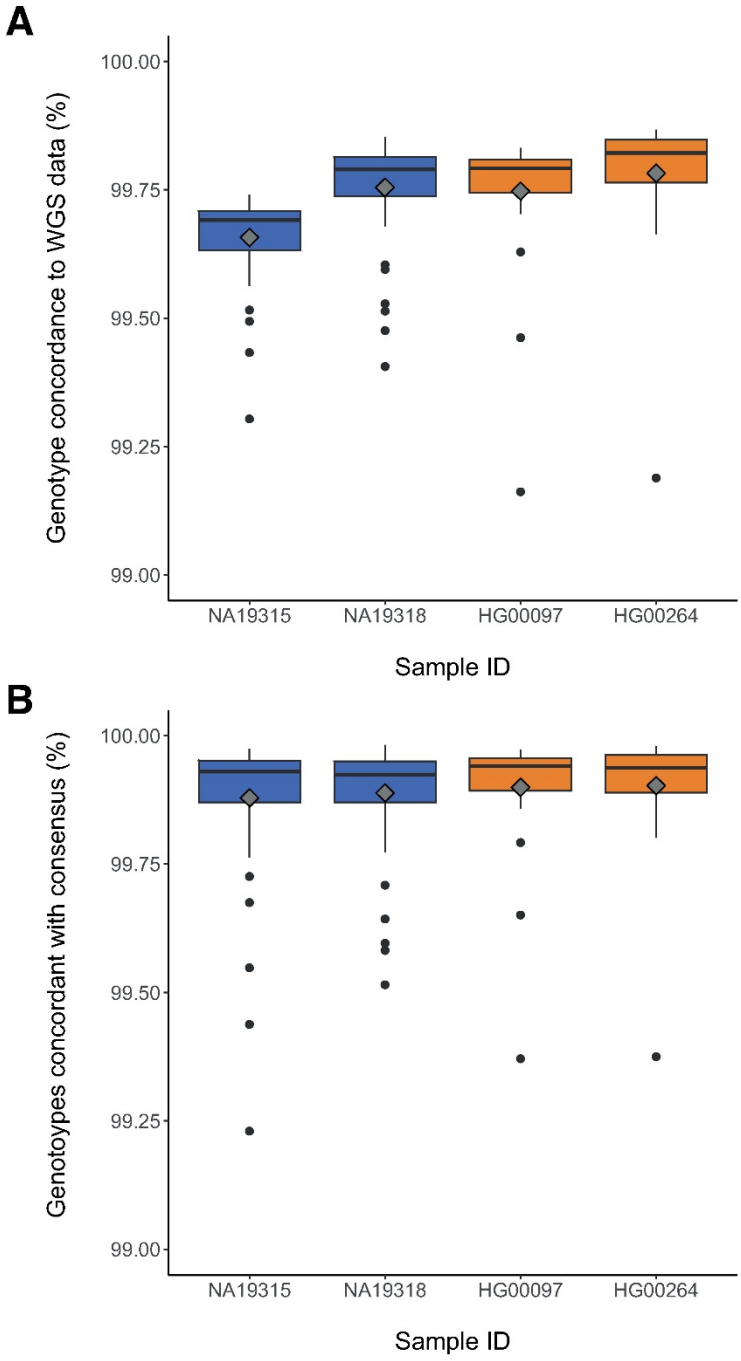


**Figure S4 - Quality control of genotyping using HapMap samples.** (**A**) Array genotype concordance to next generation sequencing data across multiple repeats of each sample for all typed variants. A drop in UKBBv2 array and WGS genotype concordance can be observed for NA19315. This sample is from an individual of African ancestry and it is therefore likely there is bias in the read alignment and genotype calling for the WGS data. (**B**) Genotype repeatability across all repeats of each sample for all typed variants. Samples NA19315 and NA19318 were used for the British donor samples (**blue**) and repeat-tested 52 times each, and HG00097 and HG00264 used for the Dutch donor samples (**orange**) and repeat-tested 32 times each, on separate genotyping runs. Grey diamonds display means, middle box lines, box bounds and whiskers represent median, upper and lower quartiles, and value spread beyond middle 50% of overall distribution and circles are outliers.

**Membership of the Blood Transfusion Genomics Consortium**

**James Daly**, Medical Director Pathology Services , Australian Red Cross Blood Service, Australia

**Gwen Clarke**, Hematopathologist, Canadian Blood Services, Canada

**Jukka Partanen**, Research Director, Finnish Red Cross Blood Service, Finland

**William Lane**, Assistant Professor of Haematology, Harvard Medical School, Brigham and Womans Hospital, United States of America

**Connie Westhoff**, Executive Scientific Director, New York Blood Centre, United States of America

**Andrea Harmer**, Interim Director Diagnostics, NHS Blood and Transplant, United Kingdom

**Martin Howell**, Head of H&I service development, NHS Blood and Transplant, United Kingdom

**John Ord**, Blood service data specialist, NHS Blood and Transplant, United Kingdom

**Nick Watkins**, Assistant Director of Research and Development, NHS Blood and Transplant, United Kingdom

**Mark Wheelan**, Lead Quality Specialist, Technology Assurance and Change Programmes, NHS Blood and Transplant, United Kingdom

**Ellen van der Schoot**, Head of Department - Experimental Immunohematology, Sanquin, the Netherlands

**Barbera Veldhuisen**, Senior Scientist - Molecular Biology, Sanquin, the Netherlands

**Anton van Weert**, Director of Operations - National Screening Laboratory, Sanquin, the Netherlands

**Lavendri Govender**, Senior Biomedical Scientist, South African National Blood Service, South Africa

**Ute Jentsch**, Lead Pathology Consultant, South African National Blood Service, South Africa

**Claire Bloor**, Genotyping Specialist, Thermo Fisher Scientific, United Kingdom

**Jeremy Gollub**, Associate Director Bioinformatics, Thermo Fisher Scientific, United States of America

**William Astle**, University Senior Lecturer, University of Cambridge, United Kingdom

**Willem H Ouwehand**, Professor of Experimental Haematology, University of Cambridge, United Kingdom

**Nicholas Gleadall**, PhD Student, University of Cambridge, United Kingdom