High-throughput genotyping assays for identification of glycophorin B deletion variants in population studies

Dominic SY Amuzu1,2,3, Kirk A Rockett3,4, Ellen M Leffler4,5, Felix Ansah,1,2 Nicholas Amoako1,2, Collins M Morang’a1,2, Christina Hubbart3, Kate Rowlands3, Anna E Jeffreys3, Lucas N Amenga-Etego1,2, Dominic P Kwiatkowski3,4,6 and Gordon A Awandare1,2

1West African Centre for Cell Biology of Infectious Pathogens, University of Ghana, Accra, GH 0233, Ghana; 2Department of Biochemistry, Cell and Molecular Biology, University of Ghana, Accra, GH 0233, Ghana; 3Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7BN, UK; 4Wellcome Sanger Institute, Hinxton CB10 1SA, UK; 5Department of Human Genetics, Eccles Institute of Human Genetics, University of Utah, Salt Lake City, UT 84112-5330, USA; 6Big Data Institute, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7FZ, UK

Corresponding authors: Kirk A Rockett. Email: krockett@well.ox.ac.uk; Gordon A Awandare. Email: gawandare@ug.edu.gh

Impact statement
Glycophorins are of interest because they serve as receptors for pathogens including Plasmodium sp., Babesia sp., Influenza virus, Vibrio cholerae El Tor Hemolysin, and Escherichia coli. Variation in the genes encoding these receptors may be important in influencing disease susceptibility. Due to high sequence homology (~96) between glycophorins A, B, and E genes, it is challenging designing assays to genotype specific variations in the three genes. This work reports the development of two separate high-throughput assays for reliably genotyping glycophorin B deletions (DEL1 and DEL2) on large-scale. This is important for population prevalence studies and identification of affected individuals for investigating the functional effects of the gene variation. Furthermore, the work identified the location of the breakpoint for GYPB DEL2, which is important for understanding the gene cluster and the deletion mechanism.

Abstract
Glycophorins are the most abundant sialoglycoproteins on the surface of human erythrocyte membranes. Genetic variation in glycophorin region of human chromosome 4 (containing GYPB, GYPB, and GYPE genes) is of interest because the gene products serve as receptors for pathogens of major public health interest, including Plasmodium sp., Babesia sp., Influenza virus, Vibrio cholerae El Tor Hemolysin, and Escherichia coli. A large structural rearrangement and hybrid glycophorin variant, known as Dantu, which was identified in East African populations, has been linked with a 40% reduction in risk for severe malaria. Apart from Dantu, other large structural variants exist, with the most common being deletion of the whole GYPB gene and its surrounding region, resulting in multiple different deletion forms. In West Africa particularly, these deletions are estimated to account for between 5 and 15% of the variation in different populations, mostly attributed to the forms known as DEL1 and DEL2. Due to the lack of specific variant assays, little is known of the distribution of these variants. Here, we report a modification of a previous GYPB DEL1 assay and the development of a novel GYPB DEL2 assay as high-throughput PCR-RFLP assays, as well as the identification of the crossover/breakpoint for GYPB DEL2. Using 393 samples from three study sites in Ghana as well as samples from HapMap and 1000 G projects for validation, we show that our assays are sensitive and reliable for genotyping GYPB DEL1 and DEL2. To the best of our knowledge, this is the first report of such high-throughput genotyping assays by PCR-RFLP for identifying specific GYPB deletion types in populations. These assays will enable better identification of GYPB deletions for large genetic association studies and functional experiments to understand the role of this gene cluster region in susceptibility to malaria and other diseases.

Keywords: Glycophorins, malaria, Plasmodium, red blood cell, invasion, GYPB deletion

Introduction
Malaria is still an important public health issue worldwide and the leading cause of death among children in sub-Saharan Africa (sSA). It is estimated that every year about 216 million cases of malaria and 445,000 deaths occur globally, with sSA being the most affected.1
Plasmodium falciparum, which is responsible for most of these deaths, has evolved complex machinery for invading erythrocytes. The mechanism is mediated by multiple redundant parasite ligands and specific human host receptors on the surface of erythrocytes to facilitate invasion.2–4 Many studies have shown that host-pathogen factors influence malaria outcomes and most certainly the parasite has affected the evolution of the human genome over the years.5,6 Several genetic host factors from single nucleotide polymorphisms to large structural variants are known to influence an individual’s susceptibility or resistance to malaria.7–10

Glycophorins (GYP) are glycosylated sialoglycoproteins found on the surface of human and animal erythrocytes.11 Human GYP A and GYP B are determinants of the major MNS blood group system while GYP C and GYP D are determinant of the Gerbich Blood Group. Some of these glycophorins have also been shown to be receptors on erythrocytes used by the \textit{P. falciparum} to invade these cells. In addition, these glycophorins serve as receptors for other pathogens such as Babesia sp., Influenza virus, encephalomyocarditis virus, Vibrio cholerae El Tor Hemolysin, and \textit{E. coli}.11–15 These include GYP A which interacts with the \textit{P. falciparum} protein erythrocyte binding antigen (EBA)-175,16–18 GYP B which interacts with erythrocyte binding ligand 1 (EBL-1),19,20 and GYP C which interacts with EBA-140, all in a sialic acid-dependent manner.21–23 The GYPE gene is not known to be expressed as a protein on the erythrocyte surface.24

The genes GYPE, GYPB, and GYP A are located in a gene cluster on the long arm of chromosome 4 (4q28-q31) approximately 360 kb long with each gene segmental duplication unit (SDU) spanning ~120 kb comprised of a gene region of ~30 kb and an intergenic region of ~90 kb (Figure 1). GYP C and GYP D are located on chromosome 2 and are not discussed here. The GYP A, GYP B, and GYPE genes are evolutionarily related, with at least 95% sequence homology between them resulting from duplication events, whereby GYP B evolved from GYP A, and GYPE evolved from GYP B.9,25,26

Recently, a large structural variant in the chromosome 4 GYP region that gives rise to the Dantu glycophorin (DUP4) has been associated with about a 40% reduction in risk for severe malaria.9,27 Interestingly, this Dantu variant was found predominantly in East African populations, but there are many other common structural variants across all the West and East African populations that have been studied.9 The most common variants identified were deletions of the whole GYP B gene and the surrounding region known as GYP B DEL1 and GYP B DEL2 (Figure 1). However, little research has been conducted on these variants and their population distributions due to the lack of high throughput methods for genotyping these structural variants. With the difficulties in screening for these deletions and other variants, there is also lack of functional data on the effect of GYP B deletions on erythrocyte invasion, the growth of \textit{P. falciparum}, and the changes that occur on the surface of the erythrocytes with respect to protein expression. Here, we show the development of two separate high-throughput assays for reliably detecting and genotyping GYP B deletions DEL1 and DEL2 that can be used to determine their distribution in populations and identify phenotypes functional investigations of these deletions on \textit{P. falciparum} erythrocyte invasion and growth. Using data from 393 samples from different ethnic populations in southern Ghana as well as DNA samples from the HapMap and 1000 G projects, we show the development of high throughput assays for GYP B DEL1. We also DEL2 and show that these are sensitive and reliable for screening population samples with little interference from other glycophorin structural variants.

**Materials and methods**

**Location of putative breakpoints for the GYP B whole-gene deletions DEL1 and DEL2**

The breakpoints for DEL1 have previously been located on GRCh37 at chr4:144835160–144835280 (4:143914007–
14391427 in GRCh38) in the 3' region of the GYPE unit of the SDU, and chr4:144945398–144945517 (4:144024245–144024364 in GRCh38) in the 3' region of the GYPB unit of the SDU,\(^9\) while the predicted location of the DEL2 breakpoint was given as 206,000 bases from the 5' end of the GYP region (GRCh37:4:144706830, GRCh38:4:143785677).\(^8\) We downloaded 4 kb of sequence surrounding Figure 2.

Experimental Biology and Medicine Volume 246 April 2021

of the DEL1 or DEL2 putative breakpoint coordinates for each of the GYPE, GYPB, and GYP A SDUs, from Ensembl (http://grch37.ensembl.org/Homo_sapiens/). The sequences were aligned using Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo/) and manually finished where required (Figure 1 and Supplementary Files 1 and 2). While initial designs were made using GRCh37, we have since compared our analysis with GRCh38 and provided coordinates with respect to GRCh38 or both where appropriate.

**Assay design for DEL1 and DEL2 putative breakpoint regions**

We developed a PCR-RFLP version of the published DEL1 assay\(^9\) using a similar primer strategy. A forward primer was positioned in the unique sequence 3' to the GYPE gene with a common reverse primer positioned in the GYPE-GYPB and GYPB-GYP A regions (Figure 2(a) and 2(b)) but placed to generate a shorter PCR amplicon (~2 kb) than that published. From the human genome reference sequence alignments (Supplementary File 1) of the equivalent SDUs, a restriction enzyme (AciI) site was identified that distinguished between the wild-type and DEL1 sequences (Figure 2(a) and (b)).

The assay for GYPB DEL2 used a strategy similar to that of the GYPB DEL1 assay, but with a unique primer such that at the GYPA end of the sequence (reverse primer) and a common primer placed at the 5' end (forward primer) (Figure 2(c) and (d)). From the human genome reference sequence alignments (Supplementary File 2) of the equivalent SDUs, a restriction enzyme (BsrBI) site was identified that distinguished between the wild-type and GYPB DEL2 sequences (Figure 2(c) and (d)). For both GYPB DEL1 and DEL2, the restriction enzymes used were non-palindromic and therefore strand oriented (Table 1, Supplementary Files 2 to 5).

**Generation of GYPB variant control DNA**

Cell lines from the 1000 G and associated projects with known GYPB deletions or wild-type (identified from the Leffler et al.\(^9\) study and Table 2) were identified and these

---

**Figure 2.** Schematic representation of strategies for amplifying and testing for the GYPB DEL1 and DEL2 structural variants. (a) Schematic representation of the alignment for the GYP SDUs showing the location of PCR primers (blue rectangles), putative breakpoint (gold rectangle), and AcI restriction site (yellow rectangle). The forward primer GYPB_DEL1_F10 is specific to upstream of GYPE in the GYPE-GYPB region. The reverse primer GYPB_DEL1_R2B5 binds to the upstream of the GYPB gene in the GYPB-GYP A region. In a normal or wild type individual, the GYPB_DEL1_R2B5 in the GYPE-GYPB region and the GYPB_DEL1_F10 forward primer forms a PCR product made of sequences in the GYPE_GYPB region. In the GYPB DEL1 state, the PCR product formed is made of sequences in the GYPE and GYPB region because the GYPB is deleted. (b) Alternate schematic representation of the GYPB DEL1 RFLP assay showing a normal chromosome and the GYPB DEL1 chromosomes aligned. Genes (green, orange, and purple rectangles) and primers (blue rectangle) are indicated as well as the AcI restriction site and PCR-digestion fragment lengths. (c) Schematic representation of the alignment for the GYP SDUs showing the location of PCR primers, putative breakpoint (gold rectangle), and BsrBI restriction site (yellow rectangle). The forward primer GYPB_DEL2_F3 is common to the GYP A downstream of the GYPB-GYP A region. The reverse primer GYPB_DEL2_R3 specifically binds to the upstream of the GYPE gene in the GYPE-GYPB region. In a normal or wild type individual, the GYPB_DEL2_F3 in the GYPB-GYP A region and the GYPB_DEL2_R3 primer forms a PCR product made of sequences in the GYPB-GYP A region. In the GYPB DEL2 state, the PCR amplicon formed is made of sequences in the GYPE and GYPB region because the GYPB is deleted. (d) Alternate schematic of the GYPB DEL2 RFLP assay showing a normal chromosome and the GYPB DEL2 chromosomes aligned. Genes (green, orange, and purple rectangles) and primers (blue rectangle) are indicated as well as the BsrBI restriction site (red dotted rectangular area) and PCR-digestion fragment lengths. Coordinates of sequences are given with respect to GRCh38.
Table 1. Samples selected for testing and sequencing to identify the breakpoints for DEL1 and DEL2.

| GYP Target Primers | Sequence (5'-3') | Dir | GC (%) | Tm (°C) |
|--------------------|-----------------|-----|--------|---------|
| PCR/sequencing DEL1 | GYP_DEL1_F10   | Fwd | 61     | 53      |
| PCR/sequencing DEL1 | GYP_DEL1_R2B5  | Rev  | 60     | 56      |
| PCR/sequencing DEL2 | DEL2_GYPEBAc_F3 | Fwd | 37.9   | 59      |
| PCR/sequencing DEL2 | DEL2_GYPBAs_R3 | Rev  | 45     | 56      |
| Sequencing DEL1    | DEL2_BP_seq_FWD | Fwd  | 5'     | 37.9    |
| Sequencing DEL2    | DEL2_BP_seq_REV1 | Rev | 5'     | 45      |

Note: Primers were designed using Primer3 and purchased from IDT (see methods).

Amuzu et al. Genotyping assays for identification of glycophorin B deletion variants

Cell lines or their genomic DNA were sourced from the NHGRI repository (Coriell Institute, NY, USA [https://www.coriell.org/1/Browse/Biobanks]).

**GYPB DEL1 and DEL2 PCR conditions and restriction digest**

QIAGEN Fast Cycling PCR Kit (Qiajen, UK) was used for both the GYPB DEL1 and DEL2 PCR reactions as detailed in Table 3a and 3b, with primers purchased from IDT (Leuven, Belgium). Restriction enzymes AciI (Catalog number: R0551L, NEB, UK) and BsrBI (Catalog number: R0102L, NEB, UK) were purchased (Table 4a). Reactions were prepared in 96-well plates (#AB-800, ThermoFisher Scientific, UK) and cycled on an MJ Tetrad (BioRad, UK) as described in Table 3a and 3b. The PCR products were digested with the relevant restriction enzymes (NEB, UK) for 2 hours at 37°C and then the digestion fragments were separated on 1% agarose gel electrophoresis containing ethidium bromide (3 ng/μL) for 2-2½ hours. Products were visualized under UV light and photographed to allow genotype assignment. All plates contained control samples obtained from the NHGRI repository (Table 2).

**Genotyping cell lines**

We tested the DEL1 and DEL2 assays on several cell lines (Table 2) that were identified from whole genome sequence analysis as having different GYP variants9 to check for cross-reactions or aberrant products.

**Screening for GYP Dantu (DUP4)**

Samples were screened for the glycophorin variant Dantu (DUP4) using the assay described by Leffler et al.9

**Ethical approval and screening population in Ghana for GYPB DEL1 and DEL2**

Ethical approval for two ongoing studies on glycophorins and malaria was granted by the Ethics Committee for Basic and Applied Sciences, College of Basic and Applied Sciences, University of Ghana (CPN: ECBAS 037/18-19), and the Noguchi Memorial Institute for Medical Research IRB, University of Ghana (CPN: 004/11-12). Written informed consent was obtained from all the study participants or their parents/guardians in the case of the children. The assays developed were used to genotype DNA samples obtained from volunteers who had been enrolled in various ongoing studies in three areas of Ghana namely: Accra, Kintampo, and Hohoe.

Venous blood samples were collected and following curation of self-reported ethnicity to only include individuals of Ghanaian origin, comprised; Kintampo (n = 147), Hohoe (n = 43), and Accra (n = 203) (Figure 6, Table 5). Genomic DNA was extracted using the Qiagen QiAmp Blood Mini Kit or Chelex-100 as described for different batches of samples.28 The DNA samples were quantified using Picogreen as described above. The gDNA samples were diluted to 20 ng/μL and stored in 96-well PCR plates at −20°C until ready for genotyping. GYPB DEL1
Table 2. Samples selected testing and sequencing to identify DEL1 and DEL2 Breakpoints

| Coriell Identifier* | Population Code | Population | Country | Genotype | DEL1 | DEL2 |
|----------------------|-----------------|------------|---------|----------|------|------|
| HG00097              | GBR             | British    | UK      | N/N      | I    | I    |
| GM06985              | CEU             | CEPH       | USA     | N/N      | I    | I    |
| GM06986              | CEU             | CEPH       | USA     | N/N      | I    | I    |
| GM06994              | CEU             | CEPH       | USA     | N/N      | I    | I    |
| GM18522              | YRI             | Yoruba     | Nigeria | N/N      | I    | I    |
| GM19140              | YRI             | Yoruba     | Nigeria | N/N      | I    | I    |
| GM19141              | YRI             | Yoruba     | Nigeria | N/N      | I    | I    |
| GM19152              | YRI             | Yoruba     | Nigeria | N/N      | I    | I    |
| GM18523              | YRI             | Yoruba     | Nigeria | DEL1/N   | DI   | I    |
| GM19207              | YRI             | Yoruba     | Nigeria | DEL1/N   | DI   | I    |
| GM19223              | YRI             | Yoruba     | Nigeria | DEL1/N   | DI   | I    |
| HG02464              | GWD             | Mandinka   | Gambia  | DEL1/DEL1| DD   | I    |
| HG02545              | ACB             | Afro-Caribbean | Barbados | DEL1/DEL1| DD   | I    |
| HG03072              | MSL             | Mende      | Sierra Leone | DEL1/DEL1| DD   | I    |
| HG03139              | ESN             | Essan      | Nigeria  | DEL1/DEL1| DD   | I    |
| GM18519              | YRI             | Yoruba     | Nigeria  | DEL1/DEL1| DD   | I    |
| GM18856              | YRI             | Yoruba     | Nigeria  | DEL2/N   | I    | Di   |
| GM19144              | YRI             | Yoruba     | Nigeria  | DEL2/N   | I    | Di   |
| GM17125              | ASW             | African-American | USA | DEL2/N   | I    | Di   |
| HG03385              | MSL             | Mende      | Sierra Leone | DEL2/DEL2| II   | DD   |
| GM20867              | GIH             | Gujarati Indian | USA | DEL14/N  | XX   | I    |
| GM18858              | YRI             | Yoruba     | Nigeria  | DEL17/N  | XX   | XX   |
| HG02586              | GWD             | Mandinka   | Gambia  | DUP1/N   | XX   | I    |
| HG02588              | GWD             | Mandinka   | Gambia  | DUP1/N   | I    | I    |
| GM18502              | YRI             | Yoruba     | Nigeria  | DUP1/N   | I    | XX   |
| GM18870              | YRI             | Yoruba     | Nigeria  | DUP1/N   | I    | I    |
| HG02250              | CDX             | Dai        | China   | DUP2/N   | I    | I    |
| HG02798              | GWD             | Mandinka   | Gambia  | DUP2/N   | I    | I    |
| GM18552              | CHB             | Han        | China   | DUP2/N   | I    | I    |
| GM18593              | CHB             | Han        | China   | DUP2/N   | I    | XX   |
| GM18605              | CHB             | Han        | China   | DUP2/N   | I    | I    |
| GM12829              | CEU             | CEPH       | USA     | DUP23/N  | XX   | XX   |
| GM12249              | CEU             | CEPH       | USA     | DUP28/N  | I    | I    |
| HG02554              | ACB             | Afro-Caribbean | Barbados | DUP4/N   | I    | I    |
| HG02585              | GWD             | Mandinka   | Gambia  | DUP6/N   | I    | I    |
| GM18545              | CHB             | Han        | China   | TRP1/N   | XX   | I    |
| GM18620              | CHB             | Han        | China   | TRP1/N   | XX   | I    |
| GM12341              | CEU             | CEPH       | USA     | TRP13/N  | I    | I    |
| GM11894              | CEU             | CEPH       | USA     | TRP5/N   | XX   | I    |
| GM18852              | YRI             | Yoruba     | Nigeria  | UNK      | XX   | I    |
| GM19221              | YRI             | Yoruba     | Nigeria  | UNK      | XX   | I    |

| Ghana Identifier§  | Population | Country | Genotype | DEL1 | DEL2 |
|---------------------|------------|---------|----------|------|------|
| GX0387              | Ghana      | Ghana   | N/N      | I    | I    |
| GX0540              | Ghana      | Ghana   | N/N      | I    | I    |
| GX0600              | Ghana      | Ghana   | N/N      | I    | I    |
| GX0610              | Ghana      | Ghana   | N/N      | I    | I    |
| GX0531              | Ghana      | Ghana   | N/N      | I    | I    |
| GX0258              | Ghana      | Ghana   | DEL1/DEL1| DD   | I    |
| GX0458              | Ghana      | Ghana   | DEL1/DEL1| DD   | I    |
| GX0537              | Ghana      | Ghana   | DEL1/DEL1| DD   | I    |
| GX0403              | Ghana      | Ghana   | DEL2/DEL2| II   | DD   |
| GX0440              | Ghana      | Ghana   | DEL2/DEL2| II   | DD   |
| GX0300+             | Nigeria    | Nigeria | DEL1/DEL2| DD   | I    |

Note: A set of 1000G samples obtained from the NHGRI repository at the Coriell Institute, NY, USA, and with known GYP variation status of GYPB DEL1 homozygote or DEL2 homozygote (Leffler et al., 2017) were selected for sequencing. A second set of samples, selected from the Ghana sample collection described in this manuscript, were also sequenced following GYPB DEL1 or DEL2 status identification from the RFLP assays described in this manuscript.

N/N: wild-type or at least not GYPB DEL1/DEL2 positive; GYPB DEL1/DEL2: DEL1 homozygous; GYPB DEL2/DEL2: DEL2 homozygous; GYPB DEL1/DEL2: Heterozygous

*Coriell identifier for 1000G and HapMap samples (https://www.coriell.org/1/Browse/Biobanks)

§Ghana identifier is an anonymised identifier with no relationship to local identifiers.

This individual was collected at the Accra study site and was later found to be of Nigerian origin.
and DEL2 genotyping assays were undertaken as given in Tables 3 and 4.

**Sequence analysis of GYPB DEL1 and DEL2 PCR products**

For Sanger sequencing, PCR amplicons were separated on the agarose gels and extracted from the gel using the Qiagen PCR gel-extraction kit (Qiagen QIAquick Gel Extraction) as described by the manufacturer. The concentration of the DNA recovered was determined by Quant-i Picogreen assay (Invitrogen, UK). Samples were prepared following instructions described by the sequencing company and sent for Sanger sequencing by Eurofins Genomics (Ebersberg, Germany [https://www.eurofinsgenomics.eu/en/custom-dna-sequencing/eurofins-services]). The sequence data were inspected and curated using Chromas (https://technelysium.com.au/wp/chromas/) to generate FASTA files for the different sequencing reactions. These data were aligned using the multiple-sequence-alignment tool Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/), after which pile-ups were manually curated and residues annotated according to the consensus sequence with respect to the PCR amplicon primers. Paralogous sequence
differences between the three genes SDUs were used to confirm the PCR products for both DEL1 and DEL2 and also to identify the putative breakpoint regions for GYPB DEL1 and DEL2.

Results
Development of novel GYPB DEL1 and DEL2 PCR-RFLP assays
We have successfully designed a PCR-RFLP assay for GYPB DEL1 using an AciI restriction enzyme digest to differentiate GYPB DEL1 from the reference (“normal”/wild type) or non-DEL1 forms. A schematic representation of the strategy for the restriction digest is presented in Figure 2(a) and 2(b) and Supplementary Figure 2. The DEL2 deletion assay was designed in a similar way as the DEL1 assay using the BsrBI restriction enzyme (Figure 2(c) and (d) and Supplementary Figure 3). Several cell lines identified as DEL1 or DEL2 homozygous or heterozygous,9 were used to test both assays and the expected banding patterns were observed (Figure 3). The non-DEL1 homozygous reference samples gave two visible PCR amplicons on agarose gel-electrophoresis after AciI digest (1.9 kb and 0.3 kb (V and W, respectively), five wells (11–15) contain homozygous cell-lines which give a single uncut band at 2.2 kb (U), while three cell-lines heterozygous for DEL1 (lanes 8–10) show two upper bands and one small band (2.3 kb [U], 1.9 kb [V], and 0.3 kb [W]). Lanes 16–19 are GYPB DEL2 positive cell lines that are all cut by the AciI enzyme (1.9 kb [V] and 0.3 kb [W]) indicating “normal” or non-DEL1. (b) GYPB DEL2 assay: The first seven wells (1–7) contain wild-type cell lines giving a single uncut band at 2.1 kb (X); lane (19) contains a GYPB DEL2 homozygous cell line giving two bands (1.3 kb [Y] and 0.8 kb [Z]); three wells (16–18) contain heterozygous cell-lines which give three bands (2.1 kb [X], 1.3 kb [Y], and 0.8 kb [Z]). Lanes 8–15 are DEL1 positive cell lines that are not cut by BsrBI indicating “normal” or non-DEL2.

Table 5. Details of Sample Collection from 3 Study Sites in Ghana

| Location | Gender | Age (years) | Ethnicity | Total |
|----------|--------|-------------|-----------|-------|
|          |        | Mean | range | Akan | Ewe | Ga | Gurunsi | Konkomba | Mo | Other |     |
| Accra    | Female | 5.44 | 1-15 | 78 | 48 | 63 | 5 | 0 | 0 | 9 | 203 |
|          | Male   | 5.15 | 1-15 | 48 | 22 | 30 | 4 | 0 | 0 | 3 | 107 |
|          |        | 5.67 | 1-15 | 30 | 26 | 33 | 1 | 0 | 0 | 8 | 98  |
| Hohoe    | Adults | >18  |       | 3  | 35 | 2  | 1 | 0 | 0 | 2 | 43  |
|          | Female | >18  |       | 2  | 21 | 0  | 1 | 0 | 0 | 2 | 26  |
|          | Male   | >18  |       | 1  | 14 | 2  | 0 | 0 | 0 | 0 | 17  |
| Kintampo | Female | 3.25 | 1-15 | 17 | 2  | 0  | 24 | 20 | 21 | 63 | 147 |
|          | Male   | 3.39 | 1-12 | 5  | 2  | 0  | 9  | 8  | 13 | 24 | 61  |
| Overall  | Female | 4.52 | 1.5 | 98 | 85 | 65 | 30 | 20 | 21 | 76 | 393 |
|          | Male   | 4.25 | 1.5 | 98 | 85 | 65 | 30 | 20 | 21 | 76 | 393 |

Note: The six named ethnic groups were represented by more than 20 samples in the full sample set.

Figure 3. GYPB DEL1 (a) and DEL2 (b) assays on cell lines with known GYP States. PCR using the GYPB DEL1 (a) or DEL2 (b) primers was carried out on the samples followed by AciI or BsrBI restriction enzyme digestion, respectively. (a) DEL1 assay: The first seven wells (1–7) contain wild-type cell lines giving bands at 1.9 kb and 0.3 kb (V and W, respectively), five wells (11–15) contain homozygous cell-lines which give a single uncut band at 2.2 kb (U), while three cell-lines heterozygous for DEL1 (lanes 8–10) show two upper bands and one small band (2.3 kb [U], 1.9 kb [V], and 0.3 kb [W]). Lanes 16–19 are GYPB DEL2 positive cell lines that are all cut by the AciI enzyme (1.9 kb [V] and 0.3 kb [W]) indicating “normal” or non-DEL1. (b) GYPB DEL2 assay: The first seven wells (1–7) contain wild-type cell lines giving a single uncut band at 2.1 kb (X); lane (19) contains a GYPB DEL2 homozygous cell line giving two bands (1.3 kb [Y] and 0.8 kb [Z]); three wells (16–18) contain heterozygous cell-lines which give three bands (2.1 kb [X], 1.3 kb [Y], and 0.8 kb [Z]). Lanes 8–15 are DEL1 positive cell lines that are not cut by BsrBI indicating “normal” or non-DEL2.
produced in this AciI digestion but are present in both the reference and alternate PCR amplicons (Table 4) and are generally not visible or resolvable. Similarly, the GYPB_DEL2 assay showed clear discrimination of genotypes using the BsrBI restriction enzyme (Figures 2(c), 2(d) and 3(b)). Reference non-DEL2 samples gave a single uncut band at 2.1 kb, while the alternate DEL2 variant gave visible bands at 1.3 kb and 0.8 kb showing that it was cut by this enzyme. Similar to the DEL1 assay, the DEL2 heterozygous samples showed three distinct bands. It is worth noting that samples that show negative result for any of the deletions should be classified as non-DEL1 or non-DEL2, respectively, since the assays only detect variants positive as homozygous or heterozygous. Also, the PCR-RFLP assays that we developed were used on non-DEL1 and non-DEL2 cell-lines that carried other GYP structural variants to confirm specificity of reactions and assays (Figure 4). Across the other cell lines tested, only non-DEL1 and non-DEL2 banding patterns were observed after restriction digest and gel electrophoresis, thus confirming the specificity of the assays.

Sanger sequencing of PCR amplicons for GYPB_DEL1 and DEL2

Several HapMap/1000G cell-lines (“normal,” DEL1 homozygote and DEL2 homozygote) as well as Ghanaian samples which comprised of three “normal” (non-DEL1, non-DEL2, and Dantu negative), three GYPB DEL1 homozygotes, one DEL1 homozygote, and one DEL1-DEL2 heterozygote were selected for Sanger sequencing of the PCR-RFLP amplicons. Pile-up sequences were aligned with the genome reference sequences (GRCh38) (Supplementary Files 2 to 5) showing that the “normal” amplicons were derived from the expected GYP regions (GYPE-GYPB for the DEL1 assay and GYPB-GYPA for the DEL2 assay [Supplementary Figure 1]). Amplicons from samples identified as homozygous for either DEL1 or DEL2 showed hybrid sequences (GYPE-GYPB/GYPB-GYPA).

For DEL1, the sequence changed from GYPE-GYPB to the GYPB-GYPA sequence between bases 1672 and 1783 in the amplicon (corresponding to 4:143914016–143914126 [GYPE-GYPB] and 4:144024254–144024364 [GYPB-GYPA] in GRCh38). The 5’ boundary was identified by a tandem repeat motif made up of CA and AT repeats (XXXXXX in Figure 5(a)), while the 3’ end was marked by a single paralogous base difference (A/G, marked Y in Figure 5(a)) between the reference sequences. A further 62 bases upstream from the 3’ boundary, there is also a 2-base paralogous difference between the reference sequences (ZZ in Figure 5(a)). The region bounded by these distinguishing motifs identifies a 111 base sequence within which the putative breakpoint occurs and is ~8.5 kb from the GYPE ATG start site and ~4.9 kb from the GYPB ATG start site, and deletes 110 kb to form DEL1 (Supplementary Files 4 and 6).

For DEL2, the sequence changed from GYPE-GYPB to the GYPB-GYP A sequence between bases 1043 and 1172 in the amplicon (corresponding to 4:143991718–143991848 [GYPE-GYPB] and 4:144094974–144095103 [GYPB-GYPA] in GRCh38). The 5’ boundary is identified by 4 different paralogous differences between the reference sequences (A/G, marked Y in Figure 5(a)) between the reference sequences. A further 62 bases upstream from the 3’ boundary, there is also a 2-base paralogous difference between the reference sequences (ZZ in Figure 5(a)). The region bounded by these distinguishing motifs identifies a 111 base sequence within which the putative breakpoint occurs and is ~8.5 kb from the GYPE ATG start site and ~4.9 kb from the GYPB ATG start site, and deletes 110 kb to form DEL1 (Supplementary Files 4 and 6).
marked by 3 separate single paralogous base differences between the reference sequences, all within 23 bases of each other (marked as Y in Figure 5(b)). This identifies a 129 base sequence within which the putative breakpoint occurs and is located /C24 86 kb from the GYPE ATG start site and /C24 76 kb from the GYPB ATG start site, and deletes 103 kb to form DEL2 (Supplementary File 5).

Distribution of GYPB DEL1 and DEL2 genotypes in Ghana

The assays developed were used to genotype DNA samples obtained from volunteers who had been enrolled into various ongoing studies in three areas of Ghana, namely Accra, Kintampo, and Hohoe. Individuals from Accra and Kintampo were all children between the ages of 1 to 15 years (mean 5.4 years from Accra [47:53% male:female ratio] and 3.25 years from Kintampo [41:59% male:female ratio]), while individuals from Hohoe were all adults (>18 years) with 39:61% male:female ratio (Table 5 and Figure 6). In total, 393 individuals were included for genotyping after curating for self-reported ethnicity to only include those of Ghanaian origin. There were 21 distinct self-reported ethnicities (Supplementary Table 1) of which 6 had N >20 individuals (No. chromosomes >40 providing a minimum detection of 2.5% allele frequency; Table 5).

The two new GYPB deletion assays as well as the assay for detecting the Dantu variant were used to measure the frequency of the DEL1, DEL2, and Dantu variants among the selected individuals at the three locations in Ghana (Table 6 and Supplementary Table 2). GYPB DEL1 was present at all three sites, with an overall GYPB DEL1 allele frequency of between 4.1% and 6.2%, while GYPB DEL2 was estimated at an overall allele frequency of between 1.3% and 5.0%. In Accra, the frequency of DEL1 was 4-fold higher than DEL2, while in Hohoe and Kintampo they were similar (4.1% vs. 5.0%, and 5.7% vs. 4.2%, respectively). As expected, we did not identify any individuals carrying the Dantu variant (Supplementary Table 2). The number of individuals from different ethnic backgrounds

Figure 5. Breakpoint sequences identified for GYPB DEL1 and DEL2 from Sanger sequencing. (a) The GYPB DEL1 highlighted regions correspond to 4:143914016-143914126 (GYPE-GYPB) and 4:144024254-144024364 (GYPB-GYPA) in GRCh38; XXXXXXX; identifies the 5’ boundary for the GYPB DEL1 breakpoint, Y identifies the 3’ end, and ZZ identifies 2 further distinguishing bases 62 bases downstream. (b) The GYPB DEL2 highlighted regions correspond to 4:143870925-143871054 (5’-GYPE), 4:143991718-143999188 (GYPE-GYPB) and 4:144094974-144095103 (GYPB-GYPA) in GRCh38. Bases marked with X’s identify the group of distinguishing bases at the 5’ end of the GYPB DEL2 breakpoint, while the Y’s identify the 3’ end of the GYPB DEL2 breakpoint. Supplementary Files 1–5 provide further detail and Sanger Sequencing pile-ups. The GYPE sequence was included to confirm that the amplicons were not amplifying this gene region.
influenced the overall allele frequencies which skewed the overall estimates. The data were therefore analyzed using the main ethnic groups represented in the dataset which were at least 20 individuals for any given group.

Of the 393 samples across the 3 study sites in Ghana, 325 came from 6 ethnic groups (Akan, Ewe, Ga, Konkomba, Mo, and Dagarti; Tables 1 and 2). GYPB DEL1 varied between 1.8% and 9.0% overall, while GYPB DEL2 varied between 0.7% and 11.9% overall. When analyzed by the study site, the DEL1/2 frequency estimates became more unreliable due to small sample sizes but where sites-ethnic groups had N ≥ 20, the estimates ranged from 2.17% to 7.4% (DEL1) and 0.7% to 11.9% (DEL2) (Table 6). In total, there were 21 ethnic groups represented of which 11/21 possess DEL1 (of the other 10 groups where DEL1 was not detected, all had 4 or fewer individuals). For DEL2, 7/21 groups showed the presence of DEL2 (but not necessarily the same groups as DEL1, Supplementary Table 2). GYPB DEL2 was not detected in 14/21 groups (of which 10 had 4 or fewer individuals and 4 groups between 8 and 19 individuals).

Discussion

Glycophorins on the surface of erythrocyte are used by malaria parasites to mediate invasion, as such, variants of the gene may protect against malaria parasite infection through mechanisms such as slowing parasite growth or reducing chances of developing severe malaria. To better assess the effects of these structural variants on resistance to malaria, there is a need for population surveys in malaria-endemic regions across sSA to generate prevalence data and identify phenotypes for conducting functional studies. However, surveys have been limited by the lack of reliable high-throughput assays for the identification of such phenotypes. The challenge of designing high-throughput PCR-based screening assays is due to the
high sequence homology (>96%) between the GYPE, GYPB, and GYP A genes. In this study, we overcame this challenge and successfully designed two high-throughput PCR-RFLP protocols that detect the presence of the two most common GYPB deletion variants in West African populations, GYPB DEL1 and DEL2. The assays were confirmed by Sanger sequencing and analyzing the PCR products of the assays. Furthermore, these assays can be performed in a 96-well plate format, followed by agarose-gel electrophoresis, making it possible to run over 90 samples in a single experiment. The assays were validated in the field by comparing their performance with an existing assay for the detection of the GYP Dantu variant (DUP4) in screening 393 individuals from three study sites in Ghana. The specificity and sensitivity of the assays in the field demonstrate their applicability as less time-consuming and less expensive options to long-read sequencing for conducting large-scale studies on the population distribution of these variants.

Details for the putative location of the breakpoints for GYPB DEL1 and DEL2 originally came from a previous study that used whole-genome sequencing from the 1000 G samples, plus additional African samples. This initial information was used to align the reference sequences for GYPA, GYPB, and GYPE across the putative breakpoints and identify the paralogous differences. Due to the high homology between the three GYP regions, identifying primers that will specifically amplify each region was challenging. It was however easy to design a single primer that could anchor the three genes and their surrounding regions. These features were used to overcome the challenge of developing an assay to differentiate between the three genes and their surrounding regions. For GYPB DEL1, a forward primer was designed to bind to the unique sequences near the GYPE transcription start site, while the counterpart (reverse primer) was common to the three regions. In the case of DEL2, designing a specific primer was more challenging because the breakpoint was not close to any entirely unique region. To overcome this, the DEL2 specific primer was placed at a location on the gene cluster close to the GYPB DEL2 breakpoint with sufficient sequence variation to allow the PCR conditions to discriminate between the sequences. In view of the fact that both the resulting reference and deletion amplicons for each assay would be of the same length, restriction enzymes were used to distinguish between them. The restriction enzymes AcI for the digestion of GYPB DEL1 PCR product and BsrBI for GYPB DEL2 PCR products were identified and selected. Other restriction enzymes may work well but have not been explored or tested in this study. One problem with using restriction enzymes is the possibility that the recognition sites themselves may contain population variation that could complicate the interpretation of the assays; however, from current information in genome-browsers and variation databases (dbSNP153), this variation appears to be uncommon for the restriction sites used here.

The two assays were used to analyze DNA from HapMap and 1000G cell-lines that had known GYPB types identified in the Leffler et al. study. This allowed further optimization of the PCR conditions and also the use of Sanger sequencing to validate the amplicons produced by the PCRs. When the homozygous GYPB DEL1 or DEL2 samples were amplified, the sequences obtained could be seen to change from one reference sequence to another and the paralogous sequence differences were used to identify the region where the switch occurred from GYP A to GYPE. Further sequence data would be required to identify whether these breakpoint boundaries are the same for all GYPB DEL1 and DEL2 chromosomes and begin to understand whether the flanking sequences were important for the mismatches during chromosome replication. In both GYPB DEL1 and DEL2 deletions, the equivalent of a whole SDU was removed amounting to ~100 kb each.

The performance of the two assay systems we have developed was evaluated by screening nearly 400 individuals from three different sites in Central to Southern Ghana. Considering the ethnic diversity at each study site and the ethnic group sample numbers, the overall allele frequency of GYPB DEL1 varied between 4.1% and 6.2%, while that of GYPB DEL2 varied between 1.3% and 5.0%. Work done by Gassner et al. reported that allele frequencies of the other three distinct deletions within African ethnicities varied greatly, to the extent that among the Congolese Mbuti Pygmy populations, cumulative allele frequencies were as high as 23.3%. The frequencies reported in our current study are similar to frequencies reported in other West African populations. In this study, analysis of the main ethnic groups with at least 20 individuals, showed the allele frequencies of DEL1 and DEL2 varied between ~1% and ~12%, with GYPB DEL2 mostly lower than DEL1. It is worth noting that the frequency estimation within ethnic groups where the sample sizes are below 100 (n = 200 chromosomes) will require confirmation by screening larger sample sizes to have a high power of study and confidence. In general, larger sample sizes would be required for DEL2 surveys as the allele frequency was less than that of GYPB DEL1. In all the ethnic groups with more than 20 individuals, we detected GYPB DEL1 and DEL2; however, non-DEL1 and non-DEL2 individuals could only be identified by increasing the number of study participants across all the ethnic groups. These two new assays will thus allow surveys on a larger scale to determine the distribution of these two main variants in other West African populations and also identify phenotypes for functional assays to investigate the effects of GYPB DEL1 and DEL2 on the susceptibility of erythrocytes to being invaded by P. falciparum and the resulting impact on disease pathogenesis.

The ability to identify these genotypes of interest from large populations accurately and rapidly has been an important goal in high-throughput genetic screening assay development. Therefore, the current DEL1 and DEL2 assays offer an opportunity for rapid screening of populations for these GYPB polymorphisms that are common especially in West Africa. Furthermore, for any malaria-related studies, it is important to collect alongside the GYP variants, other key genetic information such as sickle (rs334) which is present throughout Africa, HbC (rs33930165; present in Ghana and other West African
Countries), and G6PD as these may act as confounders in studies examining associations with susceptibility to malaria or effects on malaria parasite invasion and growth.

Developing less expensive high throughput assays targeting the less common GYP variants will provide a better understanding of the distribution and functional effects of these variants on susceptibility and pathogenesis of malaria and other disease causing pathogens that also use GYPB as a receptor. Such understanding may prove useful in guiding the design of vaccines or other therapeutic interventions targeting the pathogen interactions with these GYP proteins on the erythrocyte surface. These assays are also important for identifying individuals with the various genotypes of GYPB (homozygous, heterozygous and wild type), which is necessary for investigating the functional significance of these gene variations.

AUTHORS’ CONTRIBUTIONS: The research was conceptualized by GAA, DPK, KAR, DSYA, and EML; while AEJ, CH, KR, KAR, DSYA, and CMM curated the data; the formal analysis was done by DSYA, KAR, and EML; the funding for the research was acquired by GAA and DPK; the experiment were carried out by DSYA, KAR, AEJ, CH, and KR; GAA and DPK administered the project; AEJ, CH, KR, KAR, DSYA, CMM, NA, FA provided resources for the project; GAA, DPK, LNA, and KAR supervised the project; validation and visualization was done by DSYA and KAR; writing of the original manuscript draft was done by GAA, LNA, DSYA and KAR; the manuscript was reviewed and edited by all the authors, DSYA, KAR, EML, CMM, FA, NA, CH, KR, AEJ, LNA, DPK, and GAA.

ACKNOWLEDGEMENTS

We thank all the study participants for consenting to be part of this study.

DATA AVAILABILITY

The authors declare that the data supporting the findings of this study are available within the manuscript and its Supplementary information files. Information on data access for the Leffler et al. study is available from the MalariaGEN resource at https://www.malariagen.net/resource/23. The sequence data generated for GYPB DEL1, DEL2 deletions can be found in the supplementary files.

DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ETHICAL APPROVAL

This study was performed in accordance with the Declaration of Helsinki. Ethical approval for two ongoing studies on glycophorins and malaria was granted by the Ethics Committee for Basic and Applied Sciences, College of Basic and Applied Sciences, University of Ghana (CPN: ECBAS 037/18–19) using adults, and the Noguchi Memorial Institute for Medical Research Institutional Research Board, University of Ghana (NMMIR-IRB CPN 004/11–12). Written informed consent was obtained from all the study participants or their parents/guardians in the case of the children.

FUNDING

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The study was supported by the DELTAS Africa grant (DEL-15–007: Awandare). The DELTAS Africa Initiative is an independent funding scheme of African Academy of Sciences (AAS)’s Alliance for Accelerating Excellence in Science in Africa (AESA) which is supported by the New Partnership for Africa’s Development Planning and Coordinating Agency (NEPAD Agency) with funding from Wellcome (107755/Z/15/Z: Awandare) and the UK government.

The Malariagen Genotyping centre in Oxford is supported by Wellcome (090770/Z/09/Z, 204911/Z/16/Z; Kwiatkowski) and Wellcome also provides a core award to the Wellcome Centre for Human Genetics (203141/Z/16/Z) and the Wellcome Sanger Institute (206194).

The funders were not involved in the research design, method development, data collection and analysis, manuscript preparation, publication decision or choice of journal. The opinions and interpretations expressed in this publication are those of the author(s) and not the funders.

ORCID iDs

Dominic SY Amuzu https://orcid.org/0000-0002-8920-0708
Kirk A Rockett https://orcid.org/0000-0002-6369-9299
Collins M Morang’a https://orcid.org/0000-0002-6988-150X
Christina Hubbard https://orcid.org/0000-0001-9576-9581
Anna E Jeffreys https://orcid.org/0000-0002-6454-3093
Gordon A Awandare https://orcid.org/0000-0002-8793-3641

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

REFERENCES

1. WHO. World malaria report 2018. Luxembourg: World Health Organization, 2018, p. 210
2. Cowman AF, Crabb BS. Invasion of red blood cells by malaria parasites. Cell 2006;124:755–66
3. Gaur D, Mayer DG, Miller LH. Parasite ligand–host receptor interactions during invasion of erythrocytes by plasmodium merozoites. Int J Parasitol 2004;34:1413–29
4. Tham W-H, Wilson DW, Lopaticki S, Schmidt CQ, Tetteh-Quarcoop TB, Barlow PN, Richard D, Corbin JE, Beeson JG, Cowman AF. Complement receptor 1 is the host erythrocyte receptor for Plasmodium falciparum PfRh4 invasion ligand. Proc Natl Acad Sci USA 2010;107:17327–32
5. Kwiatkowski DP. How malaria has affected the human genome and what human genetics can teach us about malaria. Am J Hum Genet 2005;77:171–92
6. Ndila CM, Uyoga S, Macharia AW, Nyutu G, Peshu N, Ojai J, Shebe M, Awoondo KO, Mturi N, Tsoka B. Human candidate gene polymorphisms and risk of severe malaria in children in Kilifi, Kenya: a case-control association study. Lancet Haematol 2018;5:e333–e45
7. Malaria Genomics Epidemiology N. Reappraisal of known malaria resistance loci in a large multicenter study. Nat Genet 2014;46:1197–204
8. Band G, Rockett KA, Spencer CC, Kwiatkowski DP. Malaria Genomic Epidemiology NA novel locus of resistance to severe malaria in a region of ancient balancing selection. Nature 2015;526:253-7

9. Leffler EM, Band G, Busby GB, Kivinen K, Le QS, Clarke GM, Bojang KA, Conway DJ, Jallow M, Sisay-Joof F. Resistance to malaria through structural variation of red blood cell invasion receptors. Science 2017;356:eaam6393

10. Malaria Genomic Epidemiology N. Insights into malaria susceptibility using genome-wide data on 17,000 individuals from Africa, Asia and Oceania. Nat Commun 2019;10:5732

11. Jaskiewicz E, Jodlowska M, Kaczmarek R, Zerka A. Erythrocyte glycoporphins as receptors for plasmodium merozoites. Parasit Vectors 2019;12:317

12. Eder AF, Spitalnik S. Blood group antigens as receptors for pathogens. Molecular biology and evolution of blood group and MHC antigens in primates. Berlin: Springer, 1997, pp. 268-304

13. Matrosovich M, Herrler G, Klenk HD. Antigen 175 mediates invasion in Plasmodium falciparum Plasmodium falciparum of the GPC receptor site for phorins as receptors for Plasmodium merozoites. Parasit Vectors 2019;12:317

14. Kathan RH, Winzler RJ, Johnson CA. Preparation of an inhibitor of viral hemagglutination from human erythrocytes. J Exp Med 1961;113:37-45

15. Burness A. Glycophorin and sialylated components as receptors for virus. Virus receptors. Berlin: Springer, 1981, pp. 63-84

16. Camus D, Hadley TJ. Plasmodium falciparum antigen that binds to host erythrocytes and merozoites. Science 1985;230:533-6

17. Orlandi PA, Klotz FW, Haynes JD. A malaria invasion receptor, the 175-kilodalton erythrocyte binding antigen of Plasmodium falciparum recognizes the terminal Neu5Ac(alpha 2-3)gal- sequences of glycophorin A. J Cell Biol 1992;116:901-9

18. Duraisingham MT, Maier AG, Triglia T, Cowman AF. Erythrocyte-binding antigen 175 mediates invasion in Plasmodium falciparum utilizing sialic acid-dependent and-independent pathways. Proc Natl Acad Sci USA 2003;100:4796-801

19. Mayer DC, Cofie J, Jiang L, Hartl DL, Tracy E, Kabat J, Mendoza LH, Miller LH. Glycophorin B is the erythrocyte receptor of Plasmodium falciparum erythrocyte-binding ligand, EBL-1. Proc Natl Acad Sci USA 2009;106:5348-52

20. Peterson DS, Wellemes TE. EBL-1, a putative erythrocyte binding protein of Plasmodium falciparum, maps within an favored linkage group in two genetic crosses. Mol Biochem Parasitol 2002;105:105-13

21. Maier AG, Duraisingham MT, Reeder JC, Patel SS, Kazura JW, Zimmerman PA, Cowman AF. Plasmodium falciparum erythrocyte invasion through glycophorin C and selection for gerbich negativity in human populations. Nat Med 2003;9:87-92

22. Thompson JK, Triglia T, Reed MB, Cowman AF. A novel ligand from Plasmodium falciparum that binds to a sialic acid-containing receptor on the surface of human erythrocytes. Mol Microbiol 2001;41:47-58

23. Ashline DJ, Duk M, Lukasiewicz J, Reinhold VN, Lisowska E, Jaskiewicz E. The structures of glycophorin C N-glycans, a putative component of the GPC receptor site for Plasmodium falciparum EBA-140 ligand. Glycobiology 2015;25:570-81

24. Rahuel C, Elouet JF, Cartron JP. Post-transcriptional regulation of the cell surface expression of glycophorins A, B, and E. J Biol Chem 1994;269:52752-8

25. Chen J-M, Ferenc C, Cooper DN. Gene conversion in human genetic disease. Genes 2010;1:550-63

26. Polin H, Danzer M, Reiter A, Brisner M, Gaszner W, Weinberger J, Gabriel C. MN typing discrepancies based on GYP-A-B-a hybrid. Vox Sang 2014;107:393-8

27. Kariuki SN, Marin-Menendez A, Intonti V, Ravenhill BJ Lin Y-C, Macharia A, Makale J, Tendwa M, Nyamu W, Kotar J. Red blood cell tension controls Plasmodium falciparum invasion and protects against severe malaria in the dantu blood group. bioRxiv 2018;475442

28. Amoah LE, Opong A, Ayantuf-Torgby R, Abankwa J, Acquah FK, Prevention of G6PD deficiency and Plasmodium falciparum parasites in asymptomatic school children living in Southern Ghana. Malar J 2016;15:388

29. Williams TN, Wambua S, Uyoga S, Macharia A, Mwacharo JK, Newton CR, Maitland K. Both heterozygous and homozygous alpha + thalassemias protect against severe and fatal Plasmodium falciparum malaria on the Coast of Kenya. Blood 2005;106:368-71

30. Willemetz A, Natal J, Thonier V, Peyrad T, Arnaud L. Gene conversion events between GYPB and GYPE abolish expression of the S and s blood group antigens. Vox Sang 2015;108:410-6

31. Gassner C, Denomme GA, Portmann C, Bensing KM, Mattle-Greminger MP, Meyer S, Trost N, Song Y-L, Engstrom C, Jungbauder C. Two prevalent ~100-kb GYPB deletions causative of the GPB-Deficient blood group MNS phenotype S–s–U–in black Africans. Transf Med Hemother 2020;47:326-36

32. Faruqi AF, Hosono S, Driscoll MD, Dean FB, Alsamadi O, Bandaru R, Kumar G, Grimwade B, Zong Q, Sun Z. High-throughput genotyping of single nucleotide polymorphisms with rolling circle amplification. BMC Genom 2001;2:4

33. Oliphant A, Barker DL, Stuelpnagel JR, Chee MS. BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. Biotechniques 2002;32:56-8

34. Pieb FB, Patil AP, Howes RE, Nyanjiri OA, Gething PW, Williams TN, Weatherall DJ, Hay SI. Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis. Nat Commun 2010;1:104

35. Pieb FB, Howes RE, Patil AP, Nyanjiri OA, Gething PW, Bhatt S, Williams TN, Weatherall DJ, Hay SI. The distribution of haemoglobin C and its prevalence in newborns in Africa. Sci Rep 2013;3:1671

36. Clarke GM, Rockett K, Kivinen K, Hubbart C, Jeffreys AE, Rowlands K, Jallow M, Conway DJ, Bojang KA, Pinder M, Usen S, Sisay-Joof F, Sirugo G, Toure O, Thera MA, Konate S, Sissoko N, Nangaly A, Poudiouougou B, Mangano VD, Bougouma EC, Sirima SB, Modiano D, Amenga-Etego LN, Ghansah A, Koram KA, Wilson MD, Enimil A, Evans J, Amodu OK, Olaninya S, Apinjoh T, Mugri R, Ndi A, Ndila CM, Uyoga S, Macharia A, Peshu N, Williams TN, Manjurano A, Sepulveda N, Clark TG, Riley E, Drakeley C, Reyburn H, Nyirongo V, Kachala D, Molyneux MD, Dunstan SJ, Phu NH, Quyen NN, Thai CY, Hien TT, Manning L, Laman M, Siba P, Karunajeewa H, Allen S, Allen A, Davis TM, Michon P, Mueller I, Molloy SF, Campino S, Kerasidou A, Cornelius VJ, Hart L, Shah SS, Band G, Spencer CC, Agbenyega T, Achidi E, Doumbo OK, Farrar J, Marsh K, Taylor T, Kwiatkowski DP, Malaria GENSC. Characterisation of the opposing effects of G6PD deficiency on cerebral malaria and severe malarial anaemia. Elife 2017;6:e15085

(Received September 20, 2020, Accepted October 5, 2020)