bjh research paper

Genetic background of the rare Yus and Gerbich blood group phenotypes: homologous regions of the *GYPC* gene contribute to deletion alleles

Elise Gourri,[1] Gregory A. Denomme,[2] Yvonne Merki,[1] Erwin A. Scharberg,[3] Cedric Vrignaud,[4,5,6] Beat M. Frey,[1] Thierry Peyrard[4,5,6] and Christoph Gassner[1]

[1]Department of Molecular Diagnostics and Research & Development, Blood Transfusion Service Zürich, Swiss Red Cross (SRC), Zürich-Schlieren, Switzerland, [2]Diagnostic Laboratories and Blood Research Institute, Blood Center of Wisconsin, Milwaukee, WI, USA, [3]Institute of Transfusion Medicine and Immunohematology, Baden-Baden, German Red Cross Blood Service Baden-Württemberg - Hessen, Baden-Baden, Germany, [4]Département Centre National de Référence pour les Groupes Sanguins, Institut National de la Transfusion Sanguine (INTS), [5]Inserm UMR_S1134, and [6]Laboratoire d'Excellence GR-Ex, Paris, France

Received 27 October 2016; accepted for publication 11 December 2016

Correspondence: Christoph Gassner, Molecular Diagnostics and Research & Development (MOC), Blood Transfusion Service, SRC, Zürich, Rutistr. 19, 8952-Schlieren, Switzerland.
E-mail: c.gassner@zhbsd.ch
EG and GD contributed equally to this work. TP and CG also share equal responsibility for the study.

Summary

The *GYPC* gene encodes the glycophorins C and D. The two moieties express 12 known antigens of the Gerbich blood group system and functionally stabilize red blood cell membranes through their intracellular interaction with protein 4.1 and p55. Three *GYPC* exon deletions are responsible for the lack of the high-frequency antigens Ge2 (Yus type, exon 2 deletion), Ge2 and Ge3 (Gerbich type, exon 3 deletion), and Ge2 to 4 (Leach type, exons 3 and 4 deletion), but lack exact molecular description. A total of 29 rare blood samples with Yus (GE:-2,3,4) and Gerbich (GE:-2,-3,4) phenotypes, including individuals of Middle-Eastern, North-African or Balkan ancestry were examined genetically. All phenotypes could be explained by 4 different Yus alleles, characterized by deletions of exon 2 and adjacent introns, and 3 different Gerbich alleles, with deletions of exon 3 and adjacent introns. A 3600 base pair *GYPC* region, encompassing exon 2 and flanking region, shares a high degree of sequence homology with a region flanking exon 3, probably representing an evolutionary duplication event. Defining the expression of Gerbich variants presently relies on rare serological reagents. Our approach substitutes the serological characterization with a precise genotype approach to identify the rare Yus and Gerbich alleles.

Keywords: blood groups, red cell antigens, immunohaematology, transfusion medicine, immunogenetics, molecular genetics.

Antigens of the Gerbich blood group system are the result of the expression of the integral membrane sialoglycoproteins glycophorin C (GPC) and glycophorin D (GPD). These glycophorins are important in the maintenance of the red blood cell (RBC) plasma membrane structure through their interaction with protein 4.1 and p55, which links GPC and GPD to spectrin and actin (Alloisio *et al*, 1993; Marfatia *et al*, 1994). Both proteins result from the expression of one single gene, *GYPC* (Le Van Kim *et al*, 1987), with GPD representing a 21 N-terminal amino acid truncated version of GPC and encoded through an alternate transcription mechanism (Tanner, 1988; El-Maliki *et al*, 1989; Le Van Kim *et al*, 1996). The *GYPC* gene is comprised of 4 exons and spans approximately 48 kilobases on the long arm of chromosome 2.

Unlike multi-cistronic blood group system loci, like *RHD/RHCE* and *GYPA/GYPB/GYPE* that are subject to a significant antigen diversity due to intergenic recombination events (Wagner & Flegel, 2014; Meyer *et al*, 2016), *GYPC* has no homologous gene and it shares little similarity with *GYPA*, *GYPB* and *GYPE*. It shows an alternate start codon resulting in the expression of GPD sialoglycoproteins (Cartron *et al*, 1990). Exons 2 and 3 of *GYPC* and their flanking regions of approximately 3600 base pairs share 95% identity (Fig 1), suggesting that the exon organization originally arose from a

  doi: 10.1111/bjh.14578

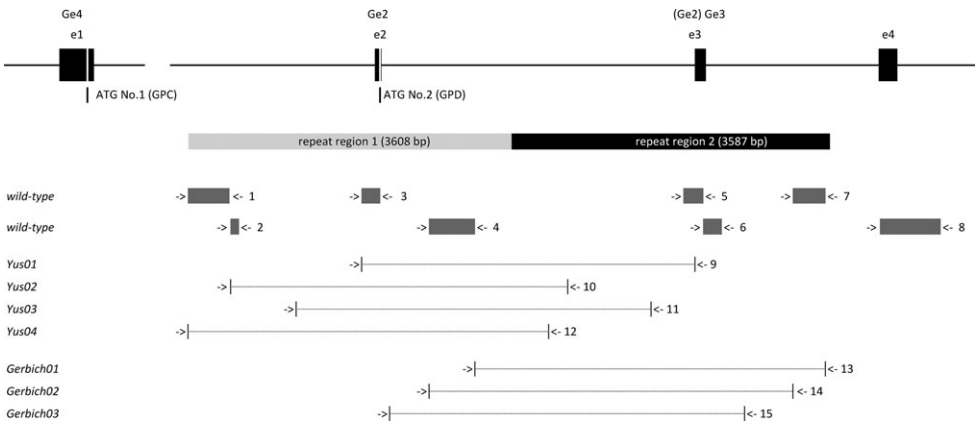


Fig 1. Schematic representation of *GYPC* and its two homologous repeat regions (light grey and black blocks). The deletions found in the *Yus* and *Gerbich* alleles are depicted as dashed lines. Diagnostic PCR-SSP products of wild-type-reactions are shown as dark grey blocks and numbered 1–8. Diagnostic gap PCRs are indicated as numbers 9–12 and 13–15, for the four *Yus* and three *Gerbich* alleles, respectively. Locations of primers for diagnostic PCRs are indicated as arrows.

duplication event. Intragenic recombination events result in nucleotide deletions encompassing one or more exons and, sometimes, their flanking regions. These regions of identity misalign during meiosis to create new alleles and consequently lead to new variant proteins (for a review, see Daniels, 2013).

The Gerbich blood group system [International Society of Blood Transfusion (ISBT) 020] is comprised of 12 antigens; 5 low-frequency and 7 high-frequency antigens (International Society of Blood Transfusion, 2008). There are at least 3 variants with affected GPC expression and with or without co-expression of GPD that result in the loss of some high-prevalence antigens (Walker & Reid, 2010). The Yus (Yussef, GE:-2,3,4) and Gerbich (GE:-2,-3,4) phenotypes are the result of a deletion of exon 2 (*GE*01.-02*) and exon 3 (*GE*01.-03*) respectively. Another exceptional form, the Leach phenotype (GE:-2,-3,-4) with a complete absence of all GPC and GPD proteins, is due to the absence of exons 3 and 4 (*GE*01N.01)* or a missense mutation in the exon 3 (*GE*01N.02)* (International Society of Blood Transfusion, 2008; Daniels, 2013). The malarial merozoite form of *Plasmodium falciparum* uses GPC and GPD as a receptor to invade human RBCs (Patel *et al*, 2001; Mayer *et al*, 2002). Therefore, variants that alter the expression of exofacial domain probably arose from the natural selective pressure of the negative health effects of malaria (Maier *et al*, 2003). Worldwide epidemiological studies have shown that variants of the Gerbich blood group system are observed in two major areas, Middle East/North Africa/Balkans, and Southeast Asia/Indonesia/Papua New Guinea. *P. falciparum* infestations are common in both of these areas (Mendis *et al*, 2001).

Antibodies to Gerbich blood group antigens are important to consider in transfusion and pregnancy. The antibodies may be associated with haemolytic transfusion reactions and haemolytic disease of the newborn. Gerbich antibodies are particularly of clinical significance in pregnancy because they are often associated with a late-onset anaemia (Arndt *et al*, 2005). The mechanism of this anaemia is similar to the one observed with antibodies to the Kell antigens (Denomme *et al*, 2006). *In vitro* studies have shown that antibodies to glycophorin C inhibit growth of early erythroid progenitor cells by inducing mitochondrial depolarization and perturbing intracellular phosphorylation pathways; this leads to actin polymerization and the induction of caspase-independent apoptosis (Micieli *et al*, 2010). This is consistent with the early expression of glycophorin C on erythroid progenitor cells during ontogeny. As a result, during pregnancy and postpartum, Gerbich antibodies not only cause classical macrophage-dependent erythroid cell destruction, but also inhibit early erythroid progenitor cell proliferation. The absence of reticulocytosis in haemolytic disease of the fetus and newborn due to Gerbich maternal antibodies is consistent with the effect on erythroid cell proliferation (Arndt *et al*, 2005).

The anti-Gerbich antibodies are potentially clinically relevant and can cause haemolytic transfusion reactions (Baughn *et al*, 2011). Alloimmunization among persons with the Yus and Gerbich phenotypes is complex. It appears that some individuals of the Gerbich phenotype make anti-Ge3 without anti-Ge2. Furthermore, the Gerbich phenotype (GE:-2,-3,4) seems to present with predominantly anti-Ge2 or anti-Ge3. However, antibodies found in the Yus phenotype (GE:-2,3,4) consistently show an anti-Ge2 specificity. The antibodies to the high-prevalence antigens are the most challenging to characterize. For Gerbich, no commercial antisera are available and antibody identification relies on rare red cells and alloantibodies shared among investigators. Sometimes these reagents are not as they appear and errors are made in the classification of new samples. Such antibody studies can benefit from samples that are well characterized at the molecular level to ensure identity between proband samples.

The nucleotide sequence homology between exons and flanking regions, together with deletion variant alleles, can make nucleotide interrogation (genotyping) for the prediction

of blood group antigens a challenge. In addition, variation in the location of the deletion, i.e. the specific breakpoint of the deletion, can place the remaining information in a different context for mRNA processing and expression. We used a collection of alloimmunized individuals of the Yus and Gerbich phenotypes to characterize their nucleotide breakpoint regions in the corresponding *GYPC* alleles. We show that both the Yus and Gerbich phenotypes result from exon deletions due to a recombination event starting in either intron 1 (Yus phenotype) or intron 2 (Gerbich phenotype). Most importantly, a small set of different deletion mutant alleles was identified for each phenotype. We propose to classify these unique breakpoint 'deletion' alleles using a set of 'diagnostic' polymerase chain reactions to categorize the rare Yus and Gerbich phenotypes. Precise categorization will help understand the evolution of *GYPC*, provide nucleotide information that may impinge on expression, and assist with the search for similar variants for antigen-matched transfusions.

Materials and methods

Samples

Twenty-eight blood samples with a Yus or Gerbich phenotype from three blood transfusion services from Switzerland, Germany and France were investigated. Alloimmunization to Gerbich (CD236) antigens, and *GYPC* gene structure based on nucleotide sequencing were studied. The blood samples were from self-identified Caucasian individuals of Middle Eastern origin including Erythrea, Northern Africa (Maghreb area) and the Balkan region of South-East Europe. Included in this report is one genomic DNA sample from the Serum Cells and Rare Fluids (SCARF) repository (Judd, 2015). Among the 28 blood samples analysed, 3 were collected in Zürich, Switzerland, 5 in Baden-Baden, Germany and 20 in Paris, France. Included in the analysis are 2 samples from Zürich who were parents (presumed carriers) of a Gerbich phenotype proband. A total of 60 Gerbich alleles were analysed; 56 from blood samples of the Yus or Gerbich phenotype, 2 of heterozygous carrier individuals and 2 from a reference Gerbich sample of the SCARF repository. The samples were not subject to informed consent authorization in accordance with the ethical committees of Switzerland. Samples from France and Germany were collected as clinical samples over time and patient/blood donor identity was protected by blinding all respective data to all other study participants.

Phenotypic and serological investigations

Samples were tested for expression of Ge2 and Ge3. Standard manual tube and gel techniques for the indirect antiglobulin test were performed using sera derived from individuals previously alloimmunized from transfusion or pregnancy. Alloimmunization to Gerbich antigens was confirmed by the Institut National de la Transfusion Sanguine (INTS) (Paris) using standard reagent red cells, previously characterized Yus and Gerbich red cells and red cells available from a cryopreserved collection.

Positional PCR, Sanger sequencing, and diagnostic PCR-SSP

Conventional and long-range high fidelity polymerase chain reactions for positional PCRs (Table SIA) were designed to evaluate genomic DNA for the presence or absence of specific sequences to locate potential *GYPC* nucleotide deletions. The final reaction volume of positional PCRs (Table SI A) and diagnostic PCRs (Table SID) was 10 μl and contained Ready PCR Buffer (Inno-train Diagnostik GmbH, Kronberg, Germany) and 0.4 units of AmpliTaq DNA polymerase (Applied Biosystems, Foster City, CA, USA). PCR products for sequencing (Table SI B) were amplified in 25-μl reactions as described above and sequenced using the primers (Table SI C) and standard Sanger sequencing procedures by Microsynth AG (Balgach, Switzerland). Cycling conditions for all positional PCRs and diagnostic PCRs were described previously (Crottet *et al*, 2014). Cycling conditions for fragments between 1500 and 2800 base pairs in size were adapted as previously described (Gassner *et al*, 2005). For fragments exceeding 4000 base pairs, the PCR Extender System from 5PRIME GmbH (Hilden, Germany) was used following the manufacturer's instructions. All in-house PCR procedures were performed on either GeneAmp PCR System 9700 or the Verity Dx automated thermocyclers (Applied Biosystems, Thermo Fisher Scientific, Life Science Group, Zug, Switzerland). PCR amplicons were visualized by agarose gel electrophoresis and documented by digital imaging.

PCR and BigDye Terminator sequencing (Sanger sequencing) were performed with the primers listed in Table SI B and C to confirm deletion breakpoints. Basic Local Alignment Search Tool (BLAST) algorithms from the National Center for Biotechnology Information (http://blast.ncbi.nlm.nih.gov/Blast.cgi) were used to identify the positions of the 2 repeat regions within *GYPC*. The reference sequence NG_007479.1 was used throughout the report for nucleotide positions. Repeat region 1 represents 3608 base pairs, from nucleotide 36979 to 40586, and repeat region 2 represents 3587 base pairs, from nucleotide 40587 to 44173. An additional set of sequence-specific priming (SSP) PCR assays was designed to classify alleles (deemed 'diagnostic PCR-SSP') on the basis of the deletions characterized (Table SI D). Primers and their concentrations are given in the supplemental information.

Results

Samples analyzed

Twenty-eight samples showed lack of expression of Ge2 or Ge3 antigens using a routine indirect antiglobulin test in the tube and gel techniques. Alloimmunization was detected in

Table I. List of the 31 samples (62 alleles) investigated.

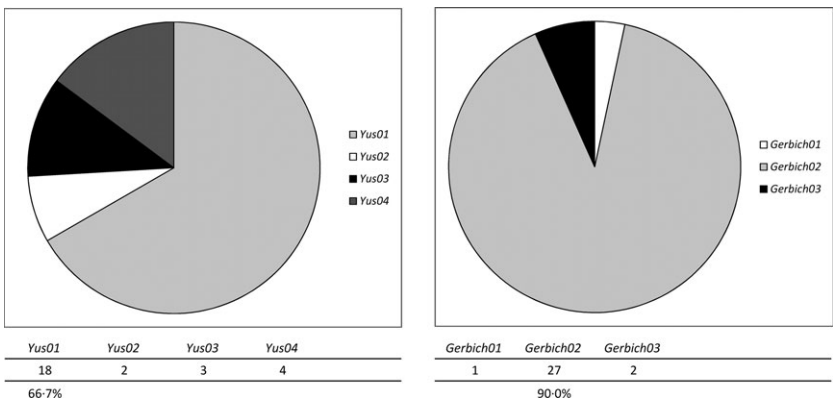
		Antibodies		Antigens		*GYPC* alleles								
Patient ID	Institute of origin	anti-Ge2	anti-Ge3	GE:2 expression	GE:3 expression	*wild-type*	*Yus01*	*Yus02*	*Yus03*	*Yus04*	*Gerbich01*	*Gerbich02*	*Gerbich03*	*GETI negative*
B-01302	Zürich	anti-Ge2	no	no	GE:3		2*							
B-01412	Zürich	no	anti-Ge3	no	no							2*		
H-00171	Zürich	anti-Ge2	no	no	no							2*		
MA-Iv (Father B-01412)	Zürich	no	no	GE:2	GE:3	1						1		
MA-Ja (Mother B-01412)	Zürich	no	no	GE:2	GE:3	1						1		
SC94-043	SCARF	n.a.	n.a.	no	no							2		
R-00448	Baden-Baden	anti-Ge2	no	n.a.	n.a.		2							
R-00449	Baden-Baden	anti-Ge2	no	n.a.	n.a.		2*							
R-00450	Baden-Baden	anti-Ge2	no	n.a.	n.a.							2*		
R-00517	Baden-Baden	anti-Ge2	no	n.a.	n.a.			1*	1*					
R-00596	Baden-Baden	anti-Ge2	no	n.a.	n.a.		1			1				
25540AYD	Paris	anti-Ge2	no	no	no							2*		
2555DEM	Paris	anti-Ge2	no	no	GE:3		2							
23266AHA	Paris	no	anti-Ge3	no	no						1*	1		
1510LAG	Paris	anti-Ge2	no	no	no							2		
31396GIR	Paris	anti-Ge2	no	no	GE:3		1							1*
25091OZD	Paris	anti-Ge2	no	no	no								2*	
612PER 1	Paris	anti-Ge2	no	no	no							2		
29154OUS	Paris	no	anti-Ge3	no	no							2*		
1097BRI	Paris	anti-Ge2	no	no	GE:3		1			1*				
1184PER 2	Paris	no	anti-Ge3	no	no							2*		
2206MEL	Paris	anti-Ge2	no	no	no							2		
5709SEG	Paris	anti-Ge2	no	no	GE:3		2							
1712SIM	Paris	anti-Ge2	no	no	GE:3		2							
29907LEG	Paris	anti-Ge2	no	no	GE:3					1*		1		
31936POY	Paris	no	anti-Ge3	no	no							2*		
23102BOU	Paris	anti-Ge2	no	no	GE:3		1					1		
3593MON	Paris	anti-Ge2	no	no	GE:3				2*					
2589PIC	Paris	anti-Ge2	no	no	GE:3		1			1				
1073MIL	Paris	†	†	no	GE:3		1	1*						
27704PIS	Paris	anti-Ge2	no	no	no							2		
						wild-type	*Yus01*	*Yus02*	*Yus03*	*Yus04*	*Gerbich01*	*Gerbich02*	*Gerbich03*	*GETI negative*
					allele count:	2	18	2	3	4	1	29	2	1

n.a.: Data not available.

Numbers marked with an asterisk (*) represent alleles that were sequenced to confirm the nucleotide breakpoints and one *GETI* allele. †anti-Ge2 or anti-Ge3 specificity not investigated.

Table II. Details of *Yus* and *Gerbich* deletion alleles and exon 2 and 3 of the *GYPC* wild-type sequence. The 5-prime and 3-prime breakpoints relative to the consensus sequence are shown along with unassignable sequences within each deletion. Positions are given relative to the nucleotide counts on NG_007479.1. Each deletion allele has its unique GenBank accession number. *Gerbich01* was published previously by Scott B & Eastal S., and deposited by them under Accession No. EF434170. We suggest allele names following the ISBT guidelines for naming of blood groups alleles.

Description	Suggested allele-name	Genbank accession number	Sample	Repeat region 1 5-prime break on NG_007479.1	Position	Not assignable to either repeat region 1, or 2, on NG_007479.1	Position	Repeat region 2 3-prime break on NG_007479.1	Deletion in bp on NG_007479.1
Yus01	*GE*01.-02.01*	LN901212	B-01302	TTGGGCCAAG GTGCTGCTAG GCATGGAGAG	39107	TCTTCCTCTC TGACCTCAGA TTCTTGTCCT CTGTTCACAG AGCCTGATCC	42767	AGGGATGTCT GGATGGCCGG ATGGCAGAAT	3609
Yus02	*GE*01.-02.02*	LN901213	R-00517	GGTTCCCAAG GTGCTGACTC CAGACCCAGA	37510	GTTCACTGGG TC	41133	CCTGGGCTGC CTCAGTGCCT TGTATTATCT	3610
Yus03	*GE*01.-02.03*	LN901214	3593MON	GACCCTGCAA AAGCTCATGT TGAAATTTGG	38369	TTCCCAAGGT GGCAGTGATG GGAGGTGGGA CCTAGTGGG	42020	AGATATTTGA GTCACAGGGG CAGATCCCTC	3611
Yus04	*GE*01.-02.04*	LN901215	29907LEG	AACTCCACCA CTAAATCCAT TCATTTATAC	37289	ATGCATGCAT TTACTCATGC ATTCAGTGAA TATTTATTGA GTACCTACTA TG	40949	TGCTAGATAC TAAACTGGAG GCTGCACTTG	3607
Gerbich01	*GE*01.-03.01*	EF434170	23266AHA	CCTGCCCCAA GCACAGTTAA TAGCCTCATT	40492	TTTCCCCCTG TATAGTGTAC ATA	44103	GAGCAACCAG CTCTGCCCTG TTTTCTGGGA	3587
Gerbich02	*GE*01.-03.02*	LN901216	H-00171	AATAAACCTG CATTCTAAGA AAAGACAGAG	40117	GCTTGCTCTG CAGCTTCAAA GTAAATGCCC AAACTCAAAA CACCCCCACC CCATGAGACT GCTGCCCCCT CACT	43777	TTCAAGGTCC TGAGTGATCT GACCCCTCCT	3585
Gerbich03	*GE*01.-03.03*	LN901217	25091OZD	GGGGAGAACT GACCTAAGGA CTTGGACAGG	39347	GGTGGCTGTG GCCATTTTTT CTCTCCCTCA GAGGTGGTTT TGAATGTGGA ATGGAGGAGT CTA	42996	GCCCTGGGTA AATATATATA TACATGCATA	3585
wild-type	*GE*01*	NG_007479.1	exon 2	n.a.	39148	AGCCTGATCC GGGGATGGCC TCTGCCTCCA CCACAATGCA TACTACCACC ATTGCAG	39205	n.a.	n.a.
wild-type	*GE*01*	NG_007479.1	exon 3	n.a.	42757	AGCCTGATCC AGGGATGTCT GGATGGCCGG ATGGCAGAAT GGAGACCTCC ACCCCCACCA TAATGGACAT TGTCGTCATT GCAG	42841	n.a.	n.a.



Yus01	Yus02	Yus03	Yus04
18	2	3	4
66·7%			

Gerbich01	Gerbich02	Gerbich03
1	27	2
	90·0%	

Fig 2. Summary of the proportion of *Yus* and *Gerbich* deletion alleles observed in this report. The most frequent alleles are given in light grey (*Yus01*, *Gerbich02*), the least frequent alleles are given in white (*Yus02*, *Gerbich01*). *Yus01* represents 66.7% of all *Yus* alleles observed in this study, *Gerbich02* represents 90.0% of all *Gerbich* alleles observed in this study.

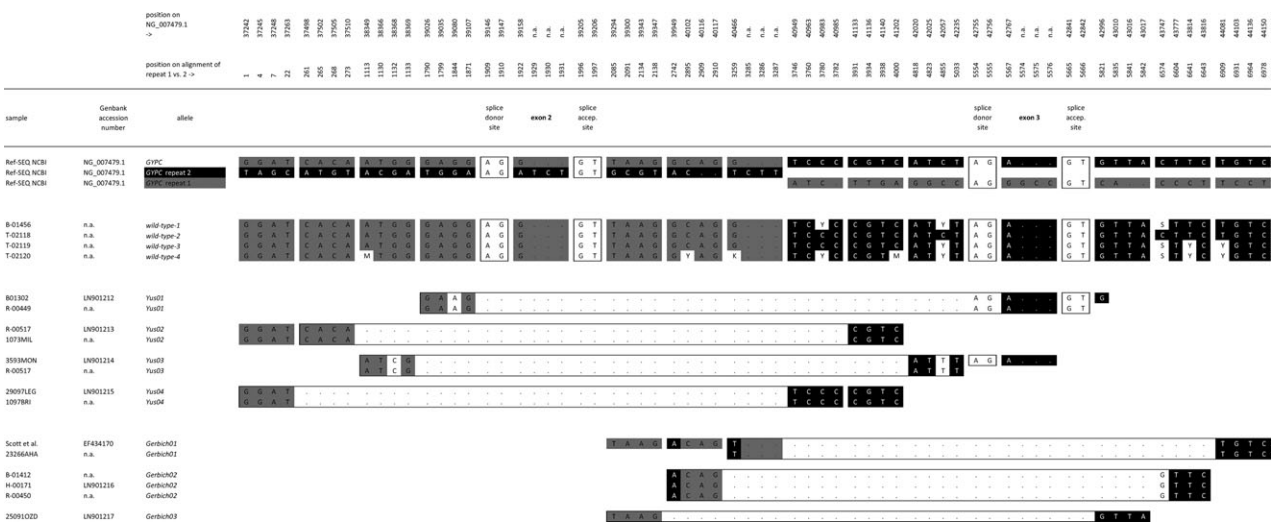


Fig 3. *Yus* and *Gerbich* deletion alleles aligned to sequences of the homologous *GYPC* repeat region 1 and 2 with the reference sequence and with 4 genomic DNA samples from individuals with wild-type GPC expression. *GYPC* reference sequence (NG_007479.1) spanning the repeat 1 and 2 regions is shown in the top row. The homologous *GYPC* repeat region 2 (black) and region 1 (grey) are aligned with their alternate homologous regions in rows 2 and 3. Exon regions are flanked by their splice sites shown in boxes. Apart from the splicing sites, only the nucleotide positions that differed between the 2 homologous regions are shown. Three gaps were introduced (indicated by dots) to maximize the alignments due to the different lengths of repeat regions 1 and 2. *GYPC* repeat regions cover 7185 nucleotides, from 37424 to 44150. DNA from 4 Gerbich variant samples was sequenced (second block) to identify potential common nucleotide deletions. International Union of Pure and Applied Chemistry (IUPAC) codes were used to indicate heterozygous positions (white cells) found in 3 of 4 sequences. The span of the 4 *Yus* and 3 *Gerbich* deletions (boxes with dots) are shown, with their flanking regions in grey (intron region 2) and black (intron region 3).

27 of 28 samples using either commercially available reagent red cells or thawed-deglycerolized Yus and Gerbich red cells from cryopreserved collections. One sample lacked information about the specificity of the antibody (anti-Ge2 or anti-Ge3). Two samples from parents of a proband (B-01412) analysed in Zürich were not alloimmunized, consistent with their presumed *Gerbich* deletion allele carrier status. Anti-Ge2 was observed in the sera of all samples with *Yus* deletion alleles (Table I). Consistent with previous serological observations (Gourri *et al*, 2015), anti-Ge2 or anti-Ge3 was demonstrated in all samples homozygous for Gerbich alleles. Anti-Ge3 was not observed among the *Yus* deletion alleles (Table I).

Molecular analysis of GYPC organization

The *Yus* and *Gerbich* deletion alleles characterized in this report are summarized in Table II. The 3 samples from Zürich identified 2 different deletion alleles, which we designated as *Yus01* and *Gerbich02*, and each were homozygous for the respective deletion. The genomic deletion of the alleles was in line with the observed Yus or Gerbich phenotypes;

Table III. Summary of the proposed blood group Gerbich alleles and names.

Phenotype description variant 1	Phenotype description variant 2	Trivial allele names	Allele names according to ISBT terminology	Suggested allele names according to ISBT terminology	Accession number	References
wild-type	GE:2,3,4	*wild-type* allele	*GE*01*	unchanged	NG_007479·1	
Yus	GE:-2,3,4	*Yus01*	*GE*01.-02*	*GE*01.-02.01*	LN901212	this study
		Yus02		*GE*01.-02.02*	LN901213	this study
		Yus03		*GE*01.-02.03*	LN901214	this study
		Yus04		*GE*01.-02.04*	LN901215	this study
Gerbich	GE:-2,-3,4	*Gerbich01*	*GE*01.-03*	*GE*01.-03.01*	EF434170	Scott & Easteal, 2008
		Gerbich02		*GE*01.-03.02*	LN901216	this study
		Gerbich03		*GE*01.-03.03*	LN901217	this study
GETI-	GE:-12	*GETI negative* allele	*GE*01.-12*	unchanged	LT605061	Poole, 2008

Yus01 presents a deletion of exon 2 (GenBank accession LN901212), and *Gerbich02* a deletion of exon 3 (LN901216). The parents of proband B-01412 were carriers of the deletion allele (*Gerbich02*) identified in their child. Three of the 5 samples from Baden-Baden were homozygous for either the *Yus01* or *Gerbich02* alleles identified in Zürich, as confirmed by positional PCR and Sanger sequencing. The remaining 2 samples were compound heterozygous for 2 new deletion alleles. One sample had 2 different Yus deletion alleles, *Yus02* (LN901213) and *Yus03* (LN901214), and the other sample had the previously observed *Yus01* deletion allele and a new *Yus04* allele (LN901215). The 20 Paris samples represented a collection of homozygous or compound heterozygous individuals carrying *GYPC* deletion alleles. Four of 20 samples were homozygous for *Yus01* ($N = 3$) or *Yus03* ($N = 1$), and 9 samples were homozygous for *Gerbich02* ($N = 8$) or a new *Gerbich03* deletion allele ($N = 1$) (LN901217). The remaining 7 samples represented compound heterozygotes for either *Yus* ($N = 3$) or *Gerbich* ($N = 1$) deletion alleles, including a previously reported *Gerbich01* deletion allele (EF434170) (Scott & Easteal, 2008), *Yus* plus *Gerbich* deletion allele ($N = 2$), or *Yus* plus a *GETI-negative* allele GE*01.-12 (LT605061) ($N = 1$) (Poole, 2008). The proportions of the different Yus and Gerbich deletion alleles are summarized in Fig 2. The SCARF sample was unambiguously homozygous for the *Gerbich02* deletion allele.

Characterization of GYPC deletions

Positional PCRs and targeted Sanger sequencing were used to characterize *GYPC* deletions (Table SI). The location of the Yus and Gerbich deletions were identified relative to reference sequence NG_007479.1. On the basis of nucleotide polymorphisms between repeat region 1 and repeat region 2, the deletion of the 4 *Yus* deletion alleles overlapped by 1841 base pairs (nucleotides 39108 to 40948, on NG_007479.1) and comprised of 40 base pairs from the 3′ region of intron 1 and 1744 base pairs from the 5′ region of intron 2 (Fig 3). The deletion alleles each contained core 12–52 base pairs that could be assigned to either side of the breakpoint, e.g. identical sequences in either intron 1 or intron 2 and had a length of 3609, 3610, 3611 and 3607 base pairs for *Yus01* to *Yus04*, respectively (Table II). A similar pattern was observed for the 3 *Gerbich* deletion alleles. They all have their deletions overlapped with 2503 base pairs (40493 to 42995, on NG_007479.1) and were comprised of 2264 base pairs from the 3′ region of intron 2 and 155 base pairs from the 5′ region of intron 3 (Fig 3). The *Gerbich* deletion alleles contained 23–74 base pairs that could be assigned to either intron 2 or intron 3 and had a length of 3587, 3585 and 3585 base pairs for *Gerbich01* to *Gerbich03*, respectively (Table II). We suggest that the ISBT Gerbich blood group system table be updated to subcategorize the *Yus* and *Gerbich* deletion alleles (Table III).

PCR-SSP for Yus and Gerbich deletion allele genotyping

A set of 15 diagnostic PCR-SSP assays were designed for the unambiguous detection of all *Yus* and *Gerbich* deletion alleles. Eight diagnostic PCR-SSP assays were designed to identify the presence of specific wild-type sequences of *GYPC* including exons 2, 3 and 4, and 7 PCR-SSP assays were designed in order to be only positive for each deletion allele (Table IV). The 15 diagnostic assays unambiguously recognized all homozygous and compound heterozygous genotypes, and made them clearly distinguishable from carriers with a wild-type allele in all cases. Representative ethidium bromide stained agarose gel electrophoresis images of PCR-SSP amplified fragments are shown in Figure S1.

Discussion

GYPC is the gene responsible for the expression of both GPC and GPD, due to an alternate start codon in exon 2. There is a region of duplication or sequence homology between exon 2 and 3 and their flanking regions (repeat regions) of approximately 3600 base pairs each. These regions (Fig 1) are responsible for the emergence of exon deletion and

Table IV. Interpretation matrices for diagnostic PCR using sequence-specific priming based assays to identify *Yus* and *Gerbich* deletion alleles.

		Wild-type reactions								*Yus* allele reactions				*Gerbich* allele reactions		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
		5′ of *Yus04*	5′ of *Yus02, Yus04*	e2	5′ of *Gerbich02, Gerbich03*	3′ of *Yus01*	e3	3′ of *Gerbich01*	e4	*Yus01*	*Yus02*	*Yus03*	*Yus04*	*Gerbich01*	*Gerbich02*	*Gerbich03*
Single alleles																
wild-type		+	+	+	+	+	+	+	+	–	–	–	–	–	–	–
Yus01		+	+	–	–	–	+	+	+	+	–	–	–	–	–	–
Yus02		+	–	–	–	+	+	+	+	–	+	–	–	–	–	–
Yus03		+	+	–	–	+	+	+	+	–	–	+	–	–	–	–
Yus04		–	–	–	–	+	+	+	+	–	–	–	+	–	–	–
Gerbich01		+	+	+	+	–	–	–	+	–	–	–	–	+	–	–
Gerbich02		+	+	+	–	–	–	+	+	–	–	–	–	–	+	–
Gerbich03		+	+	+	–	–	–	+	+	–	–	–	–	–	–	+
Genotypes																
wild-type	*wild-type*	+	+	+	+	+	+	+	+	–	–	–	–	–	–	–
wild-type	*Yus01*	+	+	+	+	+	+	+	+	+	–	–	–	–	–	–
wild-type	*Yus02*	+	+	+	+	+	+	+	+	–	+	–	–	–	–	–
wild-type	*Yus03*	+	+	+	+	+	+	+	+	–	–	+	–	–	–	–
wild-type	*Yus04*	+	+	+	+	+	+	+	+	–	–	–	+	–	–	–
wild-type	*Gerbich01*	+	+	+	+	+	+	+	+	–	–	–	–	+	–	–
wild-type	*Gerbich02*	+	+	+	+	+	+	+	+	–	–	–	–	–	+	–
wild-type	*Gerbich03*	+	+	+	+	+	+	+	+	–	–	–	–	–	–	+
Yus01	*Yus01*	+	+	–	–	–	+	+	+	+	–	–	–	–	–	–
Yus01	*Yus02*	+	+	–	–	+	+	+	+	+	+	–	–	–	–	–
Yus01	*Yus03*	+	+	–	–	+	+	+	+	+	–	+	–	–	–	–
Yus01	*Yus04*	+	+	–	–	+	+	+	+	+	–	–	+	–	–	–
Yus01	*Gerbich01*	+	+	+	+	–	+	+	+	+	–	–	–	+	–	–
Yus01	*Gerbich02*	+	+	+	–	–	+	+	+	+	–	–	–	–	+	–
Yus01	*Gerbich03*	+	+	+	–	–	+	+	+	+	–	–	–	–	–	+
Yus02	*Yus02*	+	–	–	–	+	+	+	+	–	+	–	–	–	–	–
Yus02	*Yus03*	+	+	–	–	+	+	+	+	–	+	+	–	–	–	–
Yus02	*Yus04*	+	–	–	–	+	+	+	+	–	+	–	+	–	–	–
Yus02	*Gerbich01*	+	+	+	+	+	+	+	+	–	+	–	–	+	–	–
Yus02	*Gerbich02*	+	+	+	–	+	+	+	+	–	+	–	–	–	+	–
Yus02	*Gerbich03*	+	+	+	–	+	+	+	+	–	+	–	–	–	–	+
Yus03	*Yus03*	+	+	–	–	+	+	+	+	–	–	+	–	–	–	–
Yus03	*Yus04*	+	+	–	–	+	+	+	+	–	–	+	+	–	–	–
Yus03	*Gerbich01*	+	+	+	+	+	+	+	+	–	–	+	–	+	–	–
Yus03	*Gerbich02*	+	+	+	–	+	+	+	+	–	–	+	–	–	+	–
Yus03	*Gerbich03*	+	+	+	–	+	+	+	+	–	–	+	–	–	–	+
Yus04	*Yus04*	–	–	–	–	+	+	+	+	–	–	–	+	–	–	–

Table IV. (*Continued*)

		Wild-type reactions										*Yus* allele reactions						*Gerbich* allele reactions					
		1	2		3	4				5			6	7		8	9	10	11	12	13	14		15
		5′ of *Yus04*	5′ of *Yus02*, *Yus04*		e2	5′ of *Gerbich02*, *Gerbich03*				3′ of *Yus01*			e3	3′ of *Gerbich01*		e4	*Yus01*	*Yus02*	*Yus03*	*Yus04*	*Gerbich01*	*Gerbich02*		*Gerbich03*
Yus04	*Gerbich01*	+	+		+	+				+			+	+		+	−	−	−	+	+	−		−
Yus04	*Gerbich02*	+	+		+	−				+			+	+		+	−	−	−	+	+	+		−
Yus04	*Gerbich03*	+	+		+	−				+			+	+		+	−	−	−	+	−	−		+
Gerbich01	*Gerbich01*	+	+		+	+				−			−	−		+	−	−	−	−	+	−		−
Gerbich01	*Gerbich02*	+	+		+	+				−			−	+		+	−	−	−	−	+	+		−
Gerbich01	*Gerbich03*	+	+		+	+				−			−	+		+	−	−	−	−	+	−		+
Gerbich02	*Gerbich02*	+	+		+	−				−			−	+		+	−	−	−	−	−	+		−
Gerbich02	*Gerbich03*	+	+		+	−				−			−	+		+	−	−	−	−	−	+		+
Gerbich03	*Gerbich03*	+	+		+	−				−			−	+		+	−	−	−	−	−	−		+

duplication alleles through non-homologous recombination events. The deletion alleles are found in endemic areas of *Plasmodium falciparum* including the Middle East countries and Papua New Guinea. Glycophorin C has been shown to bind the parasite EBA140 receptor (Maier *et al*, 2003). In addition, murine embryonic stem cell studies have shown that deletion of glycophorin C confers resistance to RBC malaria parasite invasion (Yiangou *et al*, 2016). Thus, the emergence of Gerbich variants is probably the result of selective pressure in endemic regions of *Plasmodium falciparum*. Subsequently, individuals expressing GYPC variants may become alloimmunized due to pregnancy or transfusion.

We sought to evaluate the *GYPC* organization of deletion alleles of the Gerbich blood group system. The Yus phenotype is characterized by a deletion of exon 2 and flanking region. Without the genetic information for the alternate start codon, the deletion results in a truncated GPC protein and no expression of GPD, and RBCs from *Yus* deletion allele lack the Ge2 antigen. Alloimmunization to the high-prevalence Ge2 antigen is possible. On the other hand, the Gerbich phenotype is characterized by a deletion of exon 3 and flanking region, responsible for the GE:-2,-3,4 phenotype. Surprisingly, the sera of these alloimmunized individuals can contain either anti-Ge2 or anti-Ge3. Apparently, a combination of both of these antibodies has not been reported. This phenomenon might be due to the immunodominance of the Ge2 antigen. For the Gerbich variant, earlier studies have shown that GPC and GPD are not present, as predicted by the deletion construct. However a variant form of the GPC protein is being expressed (Daniels, 2013). It is not clear whether GPD.Gerbich is produced since exon 2 of the *Gerbich* allele possess the alternate start codon which could initiate the transcription of the *GYPA* mRNA that results in GPD.Gerbich protein, but may not for more than a few reasons. It is conceivable that the context of the start codon in deletion alleles may be lost or other critical intron sequence information might be missing. It has also been hypothesized that the GPD.Gerbich protein produced might not be correctly transported to the erythrocyte membrane or might be unstable and quickly degraded (Colin *et al*, 1989).

Our study revealed that the *Yus* and *Gerbich* deletion alleles are defined by breakpoints that occur before or after exon 2, respectively. However, a plethora of breakpoints may be possible given the repetitive sequence flanking exons 2 and 3 is large. Currently, the Yus phenotype is defined by at least 4 deletion breakpoints. A similar observation is noted for the Gerbich phenotype, defined by 3 breakpoints after exon 2 and subject to deletions that include exon 3, and consequently lack the high-prevalence Ge3 antigen. Whether they express an altered form of GPD is presently unknown, and a subset of individuals ($N = 4$) had evidence of anti-Ge3 alloimmunization. The deletions associated with the Gerbich phenotype did not provide information to explain why individuals develop predominately anti-Ge2 or anti-Ge3.

Our set of diagnostic PCRs defined all observed *Yus* and *Gerbich* deletion alleles and provided an innovative means to subcategorize the respective GE:-2,3,4 and GE:-2,-3,4 rare RBC phenotypes. It will thereby help with transfusion decisions, e.g. allow the transfusion of GE:-2,-3,4 phenotype to GE:-2,3,4 individuals, but not vice versa. The approach can identify heterozygous (carriers) of these deletion alleles, confirm fetal genotypes and resolve compound heterozygotes. Finally, the approach will prove useful in unravelling massive parallel sequence data involving deletion alleles of the Gerbich blood group system.

With the necessity to understand the similarities and differences among the *GYPC* deletion alleles *Yus* and *Gerbich*, we performed extensive positional PCR sequence analysis to identify the breakpoints. Then, we developed a set of diagnostic PCRs that can be used to classify both Yus and Gerbich phenotypic variants on the basis of their nucleotide deletion. The Yus phenotype was defined by 4 different breakpoints and the Gerbich phenotype by 3.

Conclusions

The extensive analysis and development of diagnostic molecular tools to define the Gerbich blood group system variants that lack the high-prevalence Ge2 and Ge3 antigens provide a reliable system for categorizing these antigens. The knowledge of the precise breakpoints for the most represented *Yus* and *Gerbich* deletion alleles presents the possibility to implement a direct and reliable search for the GE:-2,3,4 and GE:-2,-3,4 rare phenotypes in large populations using high-throughput red cell genotyping platforms. There is no clear distinction in the alloimmunization pattern among the Yus and Gerbich phenotypes on the basis of their breakpoint. However, although Gerbich phenotype individuals do not express Ge2 and Ge3, they do not always produce anti-Ge3 and more frequently develop anti-Ge2. This work further supports the original notion that the alternate start site in exon 2 is leaky (Cartron *et al*, 1990), and therefore potentially subject to loss of utilization in some deletion alleles.

Acknowledgements

This work was exclusively financed by the Blood Transfusion Service Zürich, Swiss Red Cross (SRC), Zürich-Schlieren, Switzerland. The authors want to thanks Dr. Joann Moulds, Grifols Immunohematology Center, San Marcos, USA, for providing the samples from the serum, cells & rare fluids exchange program.

Author contribution

E.G. and C.G. designed the technical approach, analysed data and created tables and figures. GA.D. wrote the paper. E.G. and Y.M. conducted research and analysed results. BM.F, EA.S., C.V. and T.P. contributed sample material and data of patients. All authors reviewed the data, provided comments, and approved the final manuscript.

Conflict of interest

The authors do not disclose any conflicts of interest.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig S1. Diagnostic PCRs.

Table SI. PCR amplification and Sanger sequencing primers used in this study.

References

Alloisio, N., Dalla Venezia, N., Rana, A., Andrabi, K., Texier, P., Gilsanz, F., Cartron, J.P., Delaunay, J. & Chishti, A.H. (1993) Evidence that red blood cell protein p55 may participate in the skeleton-membrane linkage that involves protein 4.1 and glycophorin C. *Blood*, **82**, 1323–1327.

Arndt, P.A., Garratty, G., Daniels, G., Green, C.A., Wilkes, A.M., Hunt, P., Do, J., Glenn, S. & Peterson, D. (2005) Late onset neonatal anaemia due to maternal anti-Ge: possible association with destruction of eythroid progenitors. *Transfusion Medicine*, **15**, 125–132.

Baughn, M.R., Whitacre, P., Lo, G.S., Pandey, S. & Lane, T. A. (2011) A mild acute hemolytic transfusion reaction in a patient with alloanti-Ge3: a case report and review of the literature. *Transfusion*, **51**, 1966–1971.

Cartron, J., Colin, Y., Kudo, S. & Fukuda, M. (1990) Molecular genetics of human erythrocyte sialoglycoproteins A, B, C, and D. In: Blood Cell Biochemistry (ed. by Harris, J.R.), pp. 299–355. Springer Science, New York.

Colin, Y., Le Van Kim, C., Tsapis, A., Clergets, M., Aurioly, L., London, J., Galibertll, F. & Cartron, J. (1989) Human erythrocyte glycophorin C gene structure and rearrangement in genetic variants. *The Journal of Biological Chemistry*, **264**, 3773–3780.

Crottet, S.L., Henny, C., Meyer, S., Still, F., Stolz, M., Gottschalk, J., Neuenschwander, K., Taleghani, B.M., Gowland, P., Frey, B.M., Fontana, S., Hustinx, H., Niederhauser, C. & Gassner, C. (2014) Implementation of a mandatory donor RHD screening in Switzerland. *Transfusion and Apheresis Science : Official Journal of the World Apheresis Association : Official Journal of the European Society for Haemapheresis*, **50**, 169–174.

Daniels, G. (2013) Gerbich blood group system. In: Human Blood Groups (ed. by Wiley-Blackwell), pp. 410–426. Wiley-Blackwell Publishing Ltd, West Sussex, UK.

Denomme, G.A., Shahcheraghi, A., Blackall, D.P., Oza, K.K. & Garratty, G. (2006) Inhibition of erythroid progenitor cell growth by anti-Ge3. *British Journal of Haematology*, **133**, 443–444.

El-Maliki, B., Blanchard, D., Dahr, W., Beyreuther, K. & Cartron, J.-P. (1989) Structural homology between glycophorins C and D of human erythrocytes. *European Journal of Biochemistry*, **183**, 639–643.

Gassner, C., Doescher, A., Drnovsek, T.D., Rozman, P., Eicher, N.I., Legler, T.J., Lukin, S., Garritsen, H., Kleinrath, T., Egger, B., Ehling, R., Körmöczi, G.F., Kilga-Nogler, S., Schoenitzer, D. & Petershofen, E.K. (2005) Presence of RHD in serologically D-, C/E+ individuals: a European multicenter study. *Transfusion*, **45**, 527–538.

Gourri, E., Scharberg, E., Peyrard, T., Frey, B. & Gassner, C. (2015) First detailed molecular characterization of a Yus allele (GE*01.-02) and an

novel Gerbich (GE*01.-03) allele responsible for rare phenotypes in the Gerbich blood group system. *Vox Sanguinis*, **109**, 59.

International Society of Blood Transfusion (2008) Names for GE (ISBT 020) Blood Group Alleles General description. Available at: http://www.isbtweb.org/fileadmin/user_upload/files-2015/red cells/blood group allele terminology/allele tables/020 GE Alleles v2.0 119014.pdf (accessed August 23, 2016).

Judd, W.J. (2015) Serum, Cells And Rare Fluid Exchange - SCARF Exchange. Available at: http://scarfex.jove.prohosting.com/(accessed January 1, 2016).

Le Van Kim, C., Colin, Y., Blanchard, D., Dahr, W., London, J. & Cartron, J.-P. (1987) Gerbich blood group deficiency of the Ge:-1,-2,-3 and Ge:-1,-2,3 types. Immunochemical study and genomic analysis with cDNA probes. *European Journal of Biochemistry*, **165**, 571–579.

Le Van Kim, C., Piller, V., Cartron, J.P. & Colin, Y. (1996) Glycophorins C and D are generated by the use of alternative translation initiation sites. *Blood*, **88**, 2364–2365.

Maier, A.G., Duraisingh, M.T., Reeder, J.C., Patel, S.S., Kazura, J.W., Zimmerman, P. A & Cowman, A.F. (2003) Plasmodium falciparum erythrocyte invasion through glycophorin C and selection for Gerbich negativity in human populations. *Nature Medicine*, **9**, 87–92.

Marfatia, S.M., Lue, R.A. & Branton, D.C.A. (1994) In vitro binding studies suggest a membrane-associated complex between erythroid p55, protein 4.1, and glycophorin C. *Journal of Biological Chemistry*, **269**, 8631–8634.

Mayer, D.G., Mu, J.-B., Feng, X., Su, X. & Miller, L.H. (2002) Polymorphism in a Plasmodium falciparum erythrocyte-binding ligand changes its receptor specificity. *The Journal of Experimental Medicine*, **196**, 1523–1528.

Mendis, K., Sina, B., Marchesini, P. & Carter, R. (2001) The neglected burden of Plasmodium vivax malaria. *The American Journal of Tropical Medicine and Hygiene*, **64**, 97–106.

Meyer, S., Vollmert, C., Trost, N., Sigurdardottir, S., Portmann, C., Gottschalk, J., Ries, J., Markovic, A., Infanti, L., Buser, A., Amar El Dusouqui, S., Rigal, E., Castelli, D., Weingand, B., Maier, A., Mauvais, S.M., Sarraj, A., Braisch, M.C., Thierbach, J., Hustinx, H., Frey, B.M. & Gassner, C. (2016) MNSs genotyping by MALDI-TOF MS shows high concordance with serology, allows gene copy number testing and reveals new St(a) alleles. *British Journal of Haematology*, **174**, 624–636.

Micieli, J.A., Wang, D. & Denomme, G.A. (2010) Anti-glycophorin C induces mitochondrial membrane depolarization and a loss of extracellular regulated kinase 1/2 protein kinase activity that is prevented by pretreatment with cytochalasin D: implications for hemolytic disease of the fetus and newborn ca. *Transfusion*, **50**, 1761–1765.

Patel, S.S., Mehlotra, R.K., Kastens, W., Mgone, C.S., Kazura, J.W. & Zimmerman, P.A. (2001) The association of the glycophorin C exon 3 deletion with ovalocytosis and malaria susceptibility in the Wosera, Papua New Guinea. *Blood*, **98**, 3489–3491.

Poole, J. (2008) Novel mutations in GYPC giving rise to lack of Ge epitopes and anti-Ge production. *Vox Sanguinis*, **95**, 181.

Scott, B. & Easteal, S. (2008) A single-step assay for the Gerbich-negative allele of glycophorin C. *Blood Cells, Molecules & Diseases*, **41**, 1–4.

Tanner, M. (1988) Genetic variants of human red-cell membrane sialoglycoprotein beta. *Biochemistry Journal*, **250**, 407–414.

U.S. National Library of Medicine. BLAST: Basic Local Alignment Search Tool. Available at: http://blast.ncbi.nlm.nih.gov/Blast.cgi (accessed January 1, 2016).

Wagner, F.F. & Flegel, W.A. (2014) The rhesus site. *Transfus Med Hemother*, **41**, 357–363.

Walker, P. & Reid, M. (2010) The Gerbich blood group system: a review. *Immunohematology*, **26**, 60–65.

Yiangou, L., Montandon, R., Modrzynska, K., Rosen, B., Bushell, W., Hale, C., Billker, O., Rayner, J.C., Pance, A. (2016) A stem cell strategy identifies glycophorin C as a major erythrocyte receptor for the rodent malaria parasite plasmodium berghei. *PLoS ONE*, **11**, e0158238.

