Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry analysis of 36 blood group alleles among 396 Thai samples reveals region-specific variants

Philaiphon Jongruamklang,[1] Christoph Gassner,[2] Stefan Meyer,[2] Aksarakorn Kummasook,[3] Marion Darlison,[1] Chayanun Boonlum,[4] Surin Chanta,[5] Beat M. Frey,[2] Martin L. Olsson,[1,6] and Jill R. Storry [1,6*]*

BACKGROUND: Blood group phenotype variation has been attributed to potential resistance to pathogen invasion. Variation was mapped in blood donors from Lampang (northern region) and Saraburi (central region), Thailand, where malaria is endemic. The previously unknown blood group allele profiles were characterized and the data were correlated with phenotypes. The high incidence of the Vel-negative phenotype previously reported in Thais was investigated.

STUDY DESIGN AND METHODS: DNA from 396 blood donors was analyzed by matrix-assisted laser desorption/ionization–time-of-flight mass spectrometry. Outliers were investigated by serology and DNA sequencing. Allele discrimination assays for *SMIM1* rs1175550A/G and *ACKR1* rs118062001C/T were performed and correlated with antigen expression.

RESULTS: All samples were phenotyped for Rh, MNS, and K. Genotyping/phenotyping for RhD, K, and S/s showed 100% concordance. Investigation of three *RHCE* outliers revealed an e-variant antigen encoded by *RHCE*02.22*. Screening for rs147357308 (*RHCE* c.667T) revealed a frequency of 3.3%. MN typing discrepancies in 41 samples revealed glycophorin variants, of which 40 of 41 were due to Mi^a. Nine samples (2.3%) were heterozygous for *FY*01W.01* (c.265C > T), and six samples (1.5%) were heterozygous for *JK*02N.01*. All samples were wildtype *SMIM1* homozygotes with 97% homozygosity for rs1175550A.

CONCLUSIONS: Matrix-assisted laser desorption/ionization–time-of-flight mass spectrometry is an efficient method for rapid routine genotyping and investigation of outliers identified novel variation among our samples. The expected high prevalence of the Mi(a+) phenotype was observed from both regions. Of potential clinical relevance in a region where transfusion-dependent thalassemia is common, we identified two *RHCE*02* alleles known to encode an e-variant antigen.

Blood group antigen polymorphism shows great variation in different world populations. The reason for this is not completely understood; however, it has been attributed to both

ABBREVIATIONS: MALDI-TOF MS = matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; PCR-ASP = polymerase chain reaction with allele-specific primers; SNP(s) = single nucleotide polymorphism(s).

From [1]Hematology and Transfusion Medicine, Department of Laboratory Medicine, Lund University, Lund, Sweden; [2]Molecular Diagnostics & Research (MOC), Blood Transfusion Service Zürich, Zürich-Schlieren, Switzerland; [3]Department of Medical Technology, School of Allied Health Sciences, University of Phayao, Phayao, Thailand; [4]Transfusion Medicine, Department of Medical Technology and Clinical Laboratory, Saraburi Hospital, Saraburi, Thailand; [5]Transfusion Medicine, Department of Medical Technology and Clinical Laboratory, Lampang Hospital, Lampang, Thailand; and [6]Clinical Immunology and Transfusion Medicine, Laboratory Medicine, Office of Medical Services, Lund, Sweden.

Address reprint requests to: Jill R. Storry, PhD, Clinical Immunology and Transfusion Medicine, Akutgatan 8, SE-22185 Lund, Sweden; e-mail: jill.storry@med.lu.se.

This work was supported by governmental grants for clinical research to university healthcare in Lund, Sweden (ALF-SKANE-446521 to MLO); the Knut and Alice Wallenberg Foundation (2014.0312 to MLO), the Swedish Research Council (2014-71X-14251 to MLO) and the Crafoord Foundation (20151060 to JRS). PJ is supported by an education grant from the University of Phayao, Thailand.

*Both these authors contributed equally.

Received for publication December 27, 2017; revision received February 24, 2018; and accepted February 26, 2018.

doi:10.1111/trf.14624

TRANSFUSION 2018;58;1752–1762

migration and potential resistance to different pathogens. Many blood group antigens have been shown to be specific receptors for various pathogens.[1,2] Specifically, malaria has been described as one of the major driving forces of human erythroid variation.[3,4] Impairments to structural proteins such as SLC4A1 (band 3; Southeast Asian ovalocytosis), and glycophorin C (elliptocytosis in the Gerbich and Yus phenotypes), globin gene defects (thalassemias, sickle cell disease), are all attributed to attempts to evade or minimize infection by *Plasmodium* spp., as is the absence of ACKR1 in the Fy(a–b–) phenotype and the generation of hybrid proteins in the MNS system.[3,4] Malaria is endemic across the equatorial belt and encompasses a large swath of Southeast Asia. In Thailand, malaria is due primarily to infection by *Plasmodium falciparum* (44%) and *Plasmodium vivax* (47%), although other *Plasmodium* spp. are accountable for the remaining 9%, including *Plasmodium knowlesi*, recently described as a human pathogen.[5] *P. falciparum* is known to use a number of different red blood cell (RBC) ligands for attachment and invasion, whereas *P. vivax* uses ACKR1 exclusively.[6,7] For these reasons, we wanted to map variation in blood group distribution in two groups of blood donors from different regions of Thailand where malaria is endemic, one northern (Lampang) and one central (Saraburi).

Phenotypic studies of Thai blood donors have been reported previously. As early as 1967, Chandanayingyong and colleagues[8] studied the blood groups of 456 blood donors and were among the first to describe the high prevalence of the Mi(a+) phenotype. Intriguingly, they observed an unexpectedly high prevalence of the Vel-negative phenotype. A retrospective review of the serologic blood grouping records of more than a million blood donations from the Thai National Blood Center was reported in 2002 by Fongsarun and colleagues.[9] From their results, the allele frequencies were calculated for the common blood group antigens. Again, a high prevalence of Mi^a was observed (9.1%) but the results were otherwise unremarkable. The Mi(a+) phenotype has been studied more closely both by serologic and molecular techniques. Kaset and colleagues[10] used the molecular screening technique first described by Palacajornsuk and coworkers.[11] to characterize 107 Mi(a+) blood donor samples and identified five different MNS hybrid genes among these donors.

Blood group genotyping as a method for predicting the blood group phenotype is now a well-accepted practice, especially in larger blood centers where the aim is to increase the number of well-characterized blood units and enable better inventory matching and management for patients with antibodies and patients at risk of alloimmunization.[12,13]

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) has been developed and used as a high-throughput blood group genotyping platform since 2013[14] and has been proven to be a reliable and accurate predictor of blood group phenotype, although mainly in white populations in Europe.

We chose the MALDI-TOF MS platform to characterize the blood group allele profile of 396 Thai blood donors by, and correlated the results with RhD, RhCE, MNS, K/k phenotype, as well as resolved apparent discrepancies. We targeted known and potential malaria receptors, for example, ACKR1, MNS, Vel, and CD147, and looked for further polymorphism among our cohort.

MATERIALS AND METHODS

Samples

Acid-citrate-dextrose–anticoagulated blood samples were obtained following informed consent and approval from 398 unrelated blood donors at two blood banks in Thailand: 199 samples were collected at Lampang Hospital located in the north (of which two samples hemolyzed during transit; valid 197), and 199 samples were collected at Saraburi Hospital located in central Thailand. Aliquots of all samples were frozen (1) in a RNA-stabilizing reagent (TRIzol LS, ThermoFisher Scientific) for subsequent RNA analysis; (2) as buffy coats for downstream genomic analysis; (c) RBCs in glycerol (1:1; S.A.L.F. S.p.A. Laboratorio Farmacologico); and (4) plasma.

DNA preparation

Genomic DNA was extracted from 396 whole blood samples using a silica-membrane–based nucleic acid purification system (QIAsymphony DSP DNA Mini Kit, Qiagen, Inc.) on a modular instrument (QIAsymphony SP, Qiagen, Inc.), as recommended by the manufacturer. Quantity and quality of DNA samples were measured by spectrophotometry (Nanodrop Technologies, Inc.).

MALDI-TOF MS

Genomic DNA was analyzed by a MALDI-TOF MS platform (MassARRAY, Agena Bioscience GmbH) that targets 36 blood group–related single nucleotide polymorphisms (SNPs) in 15 blood group systems, as described previously.[14,15]

Briefly, the platform includes two in-house multiplex reactions, allowing the parallel genotyping of 23 and 13 SNPs, respectively, in two single tubes per DNA sample. Test reagents were purchased as a kit (Complete iPLEX Pro Genotyping Reagent Set, Agena Bioscience GmbH). Multiplex reaction primers were purchased from Metabion as described previously.[14] The SNPs targeted by the platform are specific for the discrimination of (1) the wildtype versus c.261delG variation of *ABO;* (2) the erythrocyte antigens RhC/c, RhC^w, RhE/e, K/k, $Kp^{a/b}$, $Js^{a/b}$, $Jk^{a/b}$, $Fy^{a/b}$, M/N, S/s, and $Do^{a/b}$; (3) the null phenotypes K_0 (*KEL*02N.01*, *KEL*02N.06*), Jk(a–b–) (*JK*02N.01, JK*02N.02, JK*01N.06*),

Fy(a–b–) (*FY*02N.01*), Fy(+^{w}) (*FY* c.265C/T), U– (*GYPB*01N*). *RHD* is detected by SNPs at c.455, c.787 and within the 3′ noncoding region, at Position 1362 relative to the coding sequence of *RHD*, located 108 bp after the stop codon.

The RhD-negative phenotype is defined by *RHD*08N*, *RHD*01N.06*, and *RHD*03N.01*. The assay also investigates SNPs defining (4) the low- and high-prevalence antigen pairs Lu$^{a/b}$, Di$^{a/b}$, Yt$^{a/b}$, Co$^{a/b}$, Sc1/2, LW$^{a/b}$, In$^{a/b}$, and the *SMIM1* c.64_80del defining the Vel-negative phenotype; (5) the platelet antigens HPA-1$^{a/b}$ and HPA-5$^{a/b}$; and (6) sex-addressing X-chromosomal *GYG2* and its paralog on the Y-chromosome, respectively.[15] Genotyping of all DNA samples was accomplished in two typing runs within 2 working days.

Serology

ABO and RhD typing was performed at the originating blood center. RBCs from all individuals were phenotyped for C, c, E, e, and K using the Rh/K phenotyping gel cards (BioRad, Labex AB) on an analyzer (DiaMed Gelstation, BioRad). Extended typing for e antigen was performed with a panel of monoclonal anti-e (Merck Millipore).

Phenotyping for M/N, S/s, Fy^{a}/Fy^{b}, Jk^{a}/Jk^{b}, and Di^{a} antigens was performed manually with commercially available Conformité Européenne–labeled reagents according to the manufacturers' instructions. All samples were tested for M, N, and S antigens. Samples that typed S+ were tested for s antigen. Serologic testing for Fy^{a}/Fy^{b}, Jk^{a}/Jk^{b}, and Di^{a} antigens was performed to resolve the discrepancies and unexpected results from the MALDI-TOF MS testing.

Mi^{a} testing was performed with a monoclonal antibody (clone GAMA210, Immucor, Inc.). Initially, testing was performed on samples in which there was discrepancy between genotyping and phenotyping. Phenotyping was subsequently extended to confirm the results of the *GYP.Mur* polymerase chain reaction (PCR; below).

Flow cytometry

Vel antigen expression was analyzed on RBCs from 223 blood donor samples. Briefly, RBCs were incubated with human monoclonal anti-Vel, diluted 1:10 (SpG213Dc, EFS, Rennes) followed by phycoerythrin-conjugated F(ab′) 2 fragment Goat Anti-Human IgG (Jackson ImmunoResearch, #109-116–098) as described previously.[16] Cells were analyzed on a flow cytometer (FACS Canto, Becton Dickinson) using computer software (FACS Diva, version 8.0.1, Becton Dickinson). Values were taken as median fluorescence intensity and all samples were normalized against a well-characterized sample with high Vel antigen expression. In addition, a subset of the samples (n = 140) were tested with monoclonal anti-CD147 (Basigin; MCA #28822, SEROTEC) by the protocol described above to identify variation in protein expression. An allophycocyanin-conjugated goat anti-mouse IgG (Jackson ImmunoResearch, #115-136-146) was used as the secondary antibody.

ABO genotyping

Resolution of three apparent ABO phenotype/genotype discrepancies was performed by routine in-house PCR with allele-specific primers (PCR-ASP) as described previously.[17,18]

DNA sequencing

Sanger sequencing was also used to resolve discrepancies, either in-house (MNS[19]) using reagents (ABI BigDye version 3.1, ThermoFisher Scientific) on an analyzer (ABI 3500 Dx Genetic Analyser, ThermoFisher Scientific), or following exon-specific amplification (*RHCE* Exons 1 to 10[20]) by GATC Biotech. All sequences were analyzed against a reference sequence (NM_002099.5 and NM_002100.5 for *GYPA* and *GYPB*, respectively; NG_009208 for *RHCE*) using a program for sequencing assembly, and mutation detection (CodonCode Aligner, CodonCode Co.).

Genotyping for *GYP.Mur*

PCR for *GYP.Mur* was performed by using hybrid-specific primers as described previously.[21] Briefly, the PCR assay is designed to detect the junction of Exon 3 and Intron 3 of the Hil+ glycophorin hybrids *GYP.Mur*, *GYP.Hil*, *GYP.Bun*, and *GYP.HF*, and while it does not discriminate between the different hybrids, this can be achieved by sequencing of the PCR product.

Allele discrimination genotyping

Allele discrimination genotyping using TaqMan SNP genotyping assays was performed on two SNPs not incorporated into the MALDI-TOF platform. The first, *SMIM1* rs1175550A/G (catalog # 4351379, ThermoFisher Scientific) is known to be involved in the regulation of the Vel blood group antigen expression at the RBC surface.[22] The second, *ACKR1* rs118062001 (catalog # 4351379, ThermoFisher Scientific) was identified by interrogation of the 1000 Genomes Database by the Erythrogene search engine as being more prevalent in Southeast Asia.[23] The polymorphism, c.199C > T, encodes Leu67Phe in the Duffy glycoprotein.

Allele-specific real-time PCR

Screening for *RHCE* rs147357308 (c.667T) was performed in real time using the allele-specific forward primer $^{5′}$GGATGTTCTGGCCAAGTT$^{3′}$ and an *RHCE*-specific reverse primer $^{5′}$GTGACCACCCAGCACTCTT$^{3′}$ with a master mix (PowerUp SYBR Green Master Mix, ThermoFisher Scientific) according to the manufacturers' instructions.

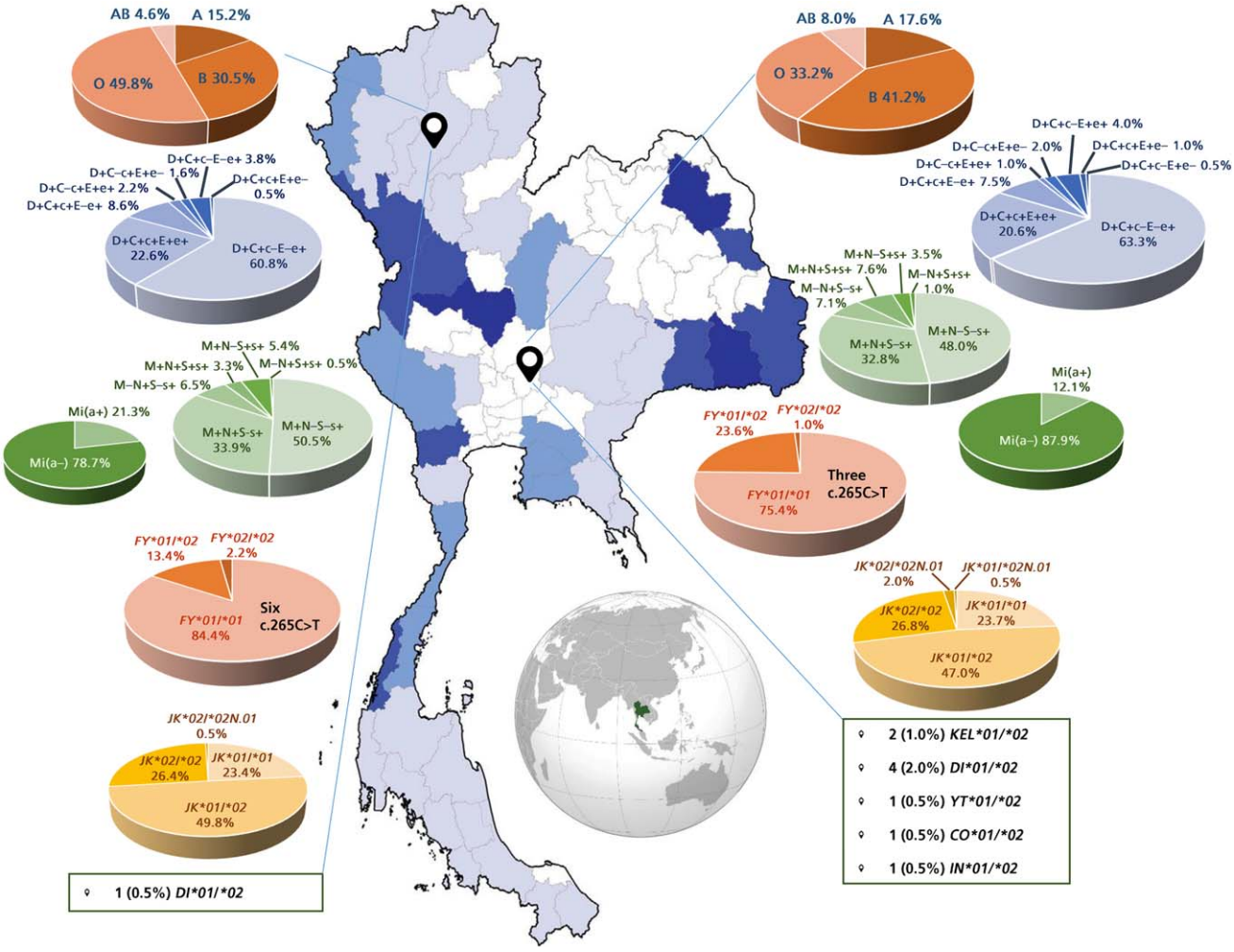

Fig. 1. The map of Thailand showing blood group distribution between two blood centers, Lampang and Saraburi. Map of Thailand modified from Corbel et al.[37]

Analysis was performed with PCR instrumentation and software (QuantStudio 3, ThermoFisher Scientific).

Cloning of *GYPB*

Messenger RNA was isolated from samples stored in Trizol LS (ThermoFisher Scientific) and converted to complementary DNA using a reverse transcription kit (High Capacity RNA-to-cDNA kit, Applied Biosystems, ThermoFisher Scientific) according to the manufacturers' instructions. The cDNA was amplified with *GYPB*-specific primers as described previously[24] and cloned into an expression vector (pEF1α-IRES-ZsGreen1 Vector, Clontech Laboratories, Inc.). Chemically competent cells (One Shot TOP10 Competent Cells, ThermoFisher Scientific) were transformed and cultured under kanamycin selection. Individual clones were picked and either directly PCR-amplified and sequenced with *GYPB*-specific primers, or cultured and purified (QIAprep Spin Miniprep Kit, Qiagen, Inc.).

RESULTS

MALDI-TOF MS analysis of 15 different blood group loci were performed on 396 DNA samples. SNPs were genotyped following the standard Agena MassARRAY iPLEX Pro genotyping protocol. Valid results were calculated using thresholds of different parameters by the software provided (TYPER 4.0, Agena Bioscience GmbH) such as height of the peak, SNP to noise ratio, and area under the peak. While all samples were analyzed, only results of the valid tests for each genotype are reported (Fig. 1, Tables 1 and S1 [available as supporting information in the online version of this paper]). The results were correlated with serologic phenotyping for ABO, RhD, RhCE, K/k, M/N, and S/s in all samples (Tables 2 and S2 [available as supporting information in the online version of this paper]). Genotyping/phenotyping for K and S/s showed 100% concordance. Only two K+k+ samples were identified.

TABLE 1. Distribution of blood group alleles by region as determined by MALDI-TOF MS and the predicted phenotype

Blood group system (n tested)	MALDI-TOF MS genotype	Predicted phenotype	Frequency*		
			Lampang	Saraburi	Total
Kell	*KEL*01/01*	K+k–	0 (0)	0 (0)	0 (0)
(n = 385)	*KEL*01/02*	K+k+	0 (0)	2 (1.0)	2 (0.5)
	*KEL*02/02*	K–k+	186 (100)	197 (99.0)	383 (99.5)
Duffy	*FY*01/01*	Fy(a+b–)	157 (84.4)	150 (75.4)	307 (79.7)
(n = 385)	*FY*01/02*	Fy(a+b+)	25 (13.4)	47 (23.6)	72 (18.7)
	*FY*02/02*	Fy(a–b+)	4 (2.2)	2 (1.0)	6 (1.6)
Kidd	*JK*01/01*	Jk(a+b–)	46 (23.4)	47 (23.7)	93 (23.5)
(n = 395)	*JK*01/02*	Jk(a+b+)	98 (49.8)	93 (47.0)	191 (48.4)
	*JK*01/02N.01*	Jk(a+b−)	0 (0)	1 (0.5)	1 (0.3)
	*JK*02/02*	Jk(a–b+)	52 (26.4)	53 (26.8)	105 (26.4)
	*JK*02/02N.01*	Jk(a–b+)	1 (0.5)	4 (2.0)	5 (1.3)
Diego	*DI*01/01*	Di(a+b–)	0 (0)	0 (0)	0 (0)
(n = 395)	*DI*01/02*	Di(a+b+)	1 (0.5)	4 (2.0)	5 (1.3)
	*DI*02/02*	Di(a–b+)	196 (99.5)	194 (98.0)	390 (98.7)
Yt	*YT*01/01*	Yt(a+b–)	197(100)	197 (99.5)	394 (99.8)
(n = 395)	*YT*01/02*	Yt(a+b+)	0 (0)	1 (0.5)	1 (0.3)
	*YT*02/02*	Yt(a–b+)	0 (0)	0 (0)	0 (0)
Dombrock	*DO*01/01*	Do(a+b–)	2 (1.0)	2 (1.0)	4 (1.0)
(n = 395)	*DO*01/02*	Do(a+b+)	31 (15.7)	49 (24.8)	80 (20.3)
	*DO*02/02*	Do(a–b+)	164 (83.3)	147 (74.2)	311 (78.7)
Colton	*CO*01/01*	Co(a+b–)	197 (100)	197 (99.5)	394 (99.8)
(n = 395)	*CO*01/02*	Co(a+b+)	0 (0)	1 (0.5)	1 (0.3)
	*CO*02/02*	Co(a–b+)	0 (0)	0 (0)	0 (0)
Indian	*IN*01/01*	In(a+b–)	0 (0)	0 (0)	0 (0)
(n = 395)	*IN*01/02*	In(a+b+)	0 (0)	1 (0.5)	1 (0.3)
	*IN*02/02*	In(a–b+)	197 (100)	197 (99.5)	394 (99.8)

* Data reported as number (%).

TABLE 2. Distribution of blood group antigens by region: results from manual phenotyping/genotyping

Blood group system (N)	Serological testing	Predicted phenotype	Frequency*		
			Lampang	Saraburi	Total
ABO	A		30 (15.2)	35 (17.6)	65 (16.4)
(n = 396)	B		60 (30.5)	82 (41.2)	142 (35.9)
	O		98 (49.8)	66 (33.2)	164 (41.4)
	AB		9 (4.6)	16 (8.0)	25 (6.3)
MNS	M+N–S–s+		63 (33.9)	80 (40.2)	143 (37.1)
(n = 385)	M+N+S–s+		92 (49.5)	82 (41.2)	174 (45.2)
	M–N+S–s+		12 (6.5)	14 (7.0)	26 (6.8)
	M+N+S+s+		8 (4.3)	15 (7.5)	23 (6.0)
	M+N–S+s+		10 (5.4)	6 (3.0)	16 (4.2)
	M–N+S+s+		1 (0.5)	2 (1.0)	3 (0.8)
Mia		Mi(a+)	42 (21.3)	24 (12.1)	66 (16.7)
(n = 396)		Mi(a–)	155 (78.7)	175 (87.9)	330 (83.3)
Rh	D+C+c–E–e+	R_1R_1	113 (60.8)	126 (63.3)	239 (62.1)
(n = 385)	D+C+c+E+e+	R_1R_2	42 (22.6)	41 (20.6)	83 (21.6)
	D+C+c+E–e+	R_1r	16 (8.6)	15 (7.5)	31 (8.1)
	D+C–c+E+e+	R_2r	4 (2.2)	2 (1.0)	6 (1.6)
	D+C–c+E+e–	R_2R_2	3 (1.6)	4 (2.0)	7 (1.8)
	D+C+c–E+e+	R_1R_z	7 (3.8)	8 (4.0)	15 (3.9)
	D+C+c+E+e–	R_2R_z	1 (0.5)	2 (1.0)	3 (0.8)
	D+C+c–E+e–	R_zR_z	0 (0)	1 (0.5)	1 (0.3)

* Data reported as number (%).

Outliers were investigated further as described below. Unexpected results obtained by MALDI-TOF MS for other alleles were also investigated serologically where reagents were available.

ABO

The MALDI-TOF MS interrogates only the *ABO* c.261delG mutation that results in an inactive ABO glycosyltransferase, and is the most common cause of the group O

TABLE 3. Extended serological analysis of three samples that gave discrepant results between genotyping and routine phenotyping

	Phenotype on Bio-Rad Rh/K cards				Phenotype predicted by MALDI-TOF MS		Source and clone of anti-e tested						
							DiaClon	Bio-Rad	Millipore (monoclonal supernate)				
Lab ID	C	c	E	e	RhCc	RhEe	MS16/MS21/MS63	MS16/MS21/MS63	MS16	MS21	MS62	MS63	MS69
Control	4+	0	4+	3+s	N/A	N/A	2+	4+	4+	4+	2+s	2+	2+w
16-050	4+	0	2+	2+	CC	EE	0	4+	4+	1+w	0	0	1+
16-220	4+	0	4+	2+	CC	EE	0	4+	4+	0	0	0	0
16-276	4+	4+	4+	2+	Cc	EE	0	4+	4+	1+	0	0	2+

N/A = not applicable.

phenotype. Therefore, *ABO*O.01* alleles can be distinguished from all other *ABO* alleles. Genotyping/phenotyping showed 99.2% concordance.

We identified three group O outliers that were heterozygous for the *ABO* c.261delG mutation.

These samples were later investigated for any mutations by routine in-house PCR-ASP and PCR with restriction fragment length polymorphism. In one outlier a typographical error was identified. The corrected phenotype was B and the result from PCR-ASP was *ABO*B.01/O.01*. The remaining two samples genotyped as *ABO*O.01/O.02*. The *ABO*O.02* allele encodes the full-length transcript but the enzyme is inactive due to an amino acid substitution in the catalytic site, p.Gly268Arg. No distinction was made between *ABO*O.01.01* and *ABO*O.01.02*.

Interestingly, the ratio of group O to group B was opposite in the two groups: in Lampang, the ratio of O:B was 49.8% versus 30.5%, whereas in Saraburi the ratio was 33.2% versus 41.2%.

RH

Serologic results were available on all samples genotyped (Table S2). MALDI-TOF MS analyzes six SNPs related to *RHD* variants; however, all samples in our cohort were wildtype for *RHD* and genotyping/phenotyping were 100% concordant.

Four SNPs related to *RHCE* can be identified by MALDI-TOF MS. Genotyping predicted the correct RhCE phenotypes in 99.2% of all samples tested. Serologic investigation of the four outliers was performed with a panel of monoclonal anti-e. Weak positive reactivity in one sample could not be confirmed on subsequent testing. The phenotype was R_zR_z in concordance with the genotyping result. Investigation of the remaining three samples revealed a qualitatively and quantitatively different e-variant antigen (Table 3), where strong reactivity was consistently observed with clone MS16, while two other antibodies (MS62 and MS63) were nonreactive. MS21 and MS69 gave variable results, possibly indicative of a quantitative difference in antigen expression. Sequence analysis identified heterozygosity for the *RHCE*02.22* allele (*RHCe* c.667C>T) in these samples.[25] This allele had been shown previously in whites associated with a weak C and partial e antigen expression. We did not observe any weak expression in C antigen, most likely due to the C antigen encoded by the R_z haplotype in *trans*.

We wanted to investigate the prevalence of this allele in the whole cohort because we suspected that it might be masked in R_1R_1 samples. Thus, we screened for the presence of *RHCE* c.667T by an allele-specific PCR and identified a further nine samples. Sequence analysis revealed that two additional samples carried *RHCE*02.22,* but in seven samples an *RHCE*02.04* allele was identified.

TABLE 4. Incidence of Mi(a+) samples in the Thai population shown regionally in this study, and compared with other studies

Source		Number of samples	Phenotype frequency (%) Mi(a+)	Mi(a–)
Serology-based methods				
National blood center (2002)[9]		20,569	9.1	90.9
Chandanayingyong and colleagues (1975)[35]		2,500	9.7	90.3
Lin-Chu and colleagues (1988)[36]			7.3	92.7
Palacajornsuk and colleagues (2007)[11]		1,041	9.03	91.97
Molecular-based methods				
Kaset and colleagues (2015)[10]		808	13.2	86.8
	North	300	23	77
	Central	508	7.5	92.5
This study		396	16.7	83.3
	Lampang (north)	197	21.3	78.7
	Saraburi (central)	199	12.1	87.9

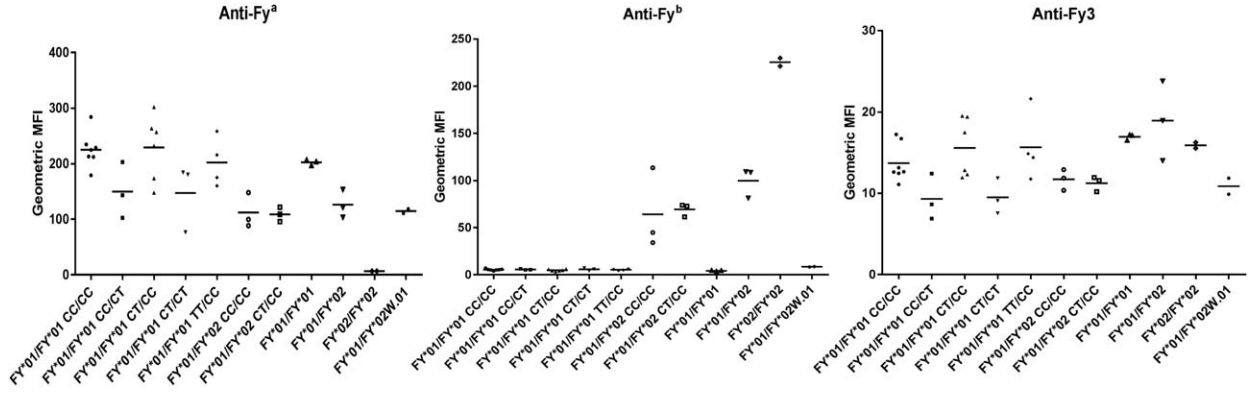

Fig. 2. Flow cytometric analysis of Fy^a, Fy^b, and Fy3 antigens was performed to investigate the effect of c.199C/T and c.265C/T polymorphism on expression. A minimum of three samples for each genotype were tested with the exception of control RBCs with the *FY*01/FY*02.01W* genotype for which we had only two examples. The genotypes are written on the X-axis and the SNP profile for c.199 and c.265 shown for each group: i.e., FY*01/FY*01 CC/CC indicates that the sample is homozygous for the Fy^a-defining SNP (c.125G); c.199CC and c.265CC. The control RBCs are shown to the right of each graph.

MNS

Comparison of the genotyping and serological results are shown in Table S3 (available as supporting information in the online version of this paper). Discrepancies in MN genotyping and the serological results were observed in 41 samples, in which the genotype was determined as *GYPA*M/M* but in which the serological phenotype was M+N+. Since glycophorin variants such as GP.Mur and GP.Bun are known to be frequent in Southeast Asia, the RBCs were tested with monoclonal anti-Mi^a. Of the 41 discrepant samples, 40 were serologically confirmed as Mi(a+). The remaining sample was investigated by DNA sequencing; however, cDNA analysis of *GYPB* revealed only wildtype transcripts, and we confirmed *MM* homozygosity by *GYPA*-specific exon 2 sequencing. The weak reaction with anti-N remains unresolved.

PCR-ASP to detect Mi^a-bearing glycophorin variants was performed in all samples. We found that 66 of 396 samples (16.7%) gave a positive result (Table 4). They were not categorized further; however, there was a difference in the prevalence of the Mi(a+) phenotype in the two regions and a higher prevalence was observed in the northern region (Lampang; 21.3%) compared to the central region (Saraburi; 12.1%), which is significantly different ($p = 0.015$).

FY

Not surprisingly, we observed a high prevalence of *FY*01* (98.4%). Nine samples (2.3%) carried the Fy^x-characteristic weakening mutation, c.265C > T, but on an *FY*01* allele, of which six were identified in blood donors from Lampang. All nine samples were homozygous for *FY*01*, thus weakened expression of Fy^a antigen was not observed by routine serologic testing. The samples were characterized further for the presence of c.298G > A, which is in linkage equilibrium with c.265T in the European *FY*02W.01* (Fy^x) allele; however,

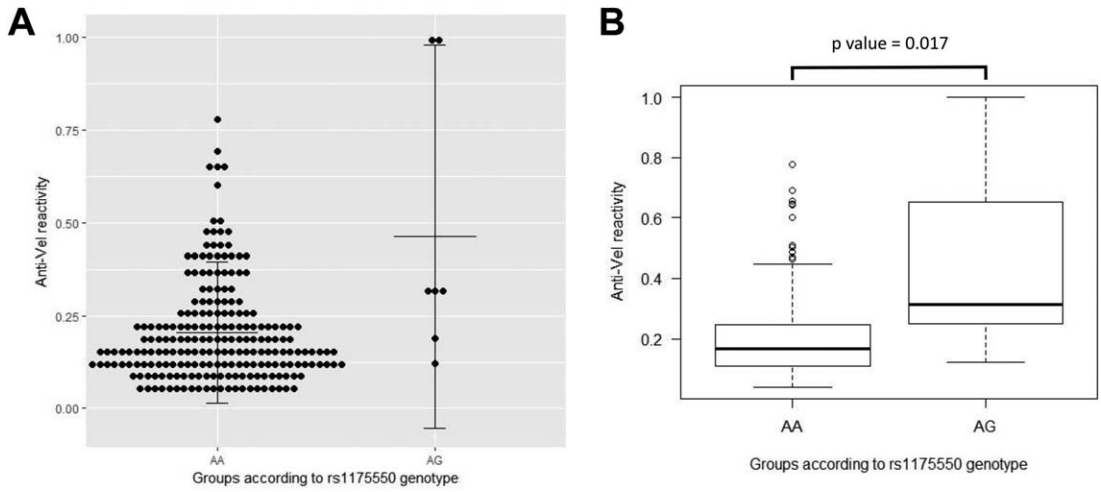

Fig. 3. Correlation between *SMIM1* genotype and Vel antigen expression. (A) Vel antigen expression on RBCs measured by flow cytometry with human monoclonal anti-Vel and PE-conjugated goat-anti-human secondary antibody. All samples were normalized against a well-characterized sample with high Vel antigen expression and plotted individually (black dots). The mean expression, and standard deviation from the mean is indicated by long and short lines, respectively. (B) All samples plotted together with the median and interquartile range. Statistical analyses were performed with Wilcoxon rank sum test with continuity correction data.

all samples were wildtype c.298G (*FY*01W.01*). Furthermore, the *FY*02* allele was less common in the northern region ($p = 0.025$).

Interrogation of the Erythrogene database[23] revealed a polymorphism in the 1000 Genomes data set that was overrepresented in Southeast Asia. This polymorphism, *ACKR1* rs118062001 (c.199C > T) encodes Leu67Phe in the Duffy glycoprotein. Using allele discrimination, we identified five c.199T homozygotes and 95 c.199C/T heterozygotes in our data set and established the prevalence of c.199T to be 13.3%. At least three samples of each genotype were tested by flow cytometry using polyclonal anti-Fy^a, -Fy^b and a monoclonal anti-Fy3 to determine if this amino acid change affected antigen expression. Samples heterozygous for c.265C/T were tested in parallel. In contrast to c.265T, the new polymorphism did not appear to affect expression of Fy antigens (Fig. 2).

JK

Six samples (1.5%) were heterozygous for the *JK*02N.01* allele, of which five genotyped as *JK*02/JK*02N.01* and one as *JK*01/JK*02N.01*. The latter typed Jk(a+b–). The null alleles were more common in the central region (5/6 samples).

Vel

MALDI-TOF MS genotyping analysis of the SMIM1 c.64_80del polymorphism underlying the Vel-negative type revealed all samples to be wildtype. Allele discrimination was performed for *SMIM1* rs1175550 to assess if SMIM1 expression may be weakened among the Thai. Weak antigen expression is associated with rs1175550A, while presence of *SMIM1* rs1175550G increases Vel at the cell surface.[22,26] To investigate Vel antigen expression, we analyzed the RBCs from 223 blood donors by flow cytometry using monoclonal anti-Vel. Expression was correlated to the rs1175550 genotype, determined by allele discrimination, where homozygosity for rs1175550A was shown in 216 of 223 samples tested (97%); the remaining seven samples genotyped as rs1175550A/G. The results are summarized in Fig. 3.

Homozygosity for rs1175550A showed lower median fluorescence intensity for Vel antigen expression than heterozygous samples ($p = 0.017$). This is consistent with data from European and African populations.[22,27]

Polymorphism in other blood group systems

Five samples genotyped *DI*01/DI*02*, and were confirmed serologically to be Di(a+b+). Of note, a single sample each of the following genotypes was detected: *CO*01/CO*02*; YT*01/YT*02*, *IN*01/IN*02*. Since basigin (CD147), the carrier of the Ok blood group, is a malaria receptor,[28] we analyzed 165 RBC samples by flow cytometry using monoclonal anti-CD147. No variation in protein expression was observed.

DISCUSSION

The application of genotyping in transfusion medicine for the prediction of blood group antigens is widespread in many blood centers, with the goal of increasing a well-characterized inventory of RBC units for transfusion. It is important therefore to understand the blood group

polymorphism distribution of the population for which blood is being provided. While our study is not novel in concept, the use of MALDI-TOF MS provided a rapid platform for genotyping our samples as a basis for further examination of variation in the Thai population. Since the MALDI-TOF MS platform is focused primarily to detect SNPs of importance in transfusion medicine, we extended our investigation to include other selected blood group–related polymorphisms with potential relevance in malaria (although even these were limited to noncarbohydrates). Thus, our data set provides new information regarding these receptors.

We have mapped the distribution of the Rh, MNS, and K/k antigens in our samples. As expected, the *RHCE*02* allele was highly prevalent. Following a discrepancy between the genotyping and phenotyping results in 3/15 R_1R_z samples, we investigated these samples further and identified a weakened e antigen that was not predicted by genotyping. Sequence analysis identified the *RHCE*02.22* allele, which had been described previously in nine samples in a study of RhCE variants, and was at that time identified in individuals of white descent.[25] Furthermore, anti-e had been stimulated in one person with the R_1R_2 phenotype carrying *RHCE*02.22*, indicating a qualitatively different e antigen. Screening of our cohort for *RHCE* rs147357308T (characteristic of this allele) revealed a further nine samples, of which two carried *RHCE*02.22* and the remaining seven samples carried another variant, *RHCE*02.04* (*RHCE*CeVA*). This finding was somewhat surprising, since *RHCE*02.04* had been described only in whites previously, and again was associated with an altered C antigen.[29] Our study shows the c.667T polymorphism to be relatively common in the Thai and amounted to 3.3% of the *RHCE*02* alleles. When investigating the prevalence of rs147257308 in tables of human variation, such as 1000 Genomes and EXAc (accessed through www.ensembl.org), the minor allele of this SNP was shown to be present only among the Chinese Dai in Xishuangbanna at approximately 1% in the East Asian superpopulation. The Thai are not represented in the East Asian superpopulation; however, this subpopulation bears the most similarity to the Thai people. The prevalence is similar to that in the African superpopulation, where c.667C > T is found on the *RHCE*01.07* alleles (approx. 1.6%) and also encodes a variant e antigen.[25]

Prediction of N antigen by MALDI-TOF MS was not reliable in this group of samples due to the presence of Mi^a. RBCs that were genotypically *GYPA*01* homozygotes but that also carried *GYP.Mur* hybrid reacted with anti-N. This is due to the hybrid that carries "N" and which cross-reacts with routine anti-N reagents. Since the MALDI-TOF MS assay did not screen for Mi^a-carrying hybrids, a specific PCR was used to screen all 396 samples.[21] In total, 66 samples (16.7%) were found to carry a hybrid allele. Interestingly, these hybrids were more common in samples from the north of Thailand, reinforcing the gradient from north to south among the different ethnic groups that had been seen also by Kaset and colleagues.[10] Furthermore, the incidence of hybrid alleles in our cohort was higher than that seen previously (Table 4), where the average has been between 9 and 10%. This may be due to a relative insensitivity of serological testing in the earlier studies, and may explain the higher than expected incidence of anti-Mi^a in transfused patients.[30]

The *FY*01* allele was highly prevalent (89.1%) in our populations. Of interest was the identification of *FY* c.265T, a polymorphism that has been shown to weaken the expression of Fy^b in Caucasian populations where it is carried on *FY*02*, most often in linkage disequilibrium with *FY* c.298A.[31] In our samples, c.298A was not present. This allele, *FY*01W.01*, was described previously in a Vietnamese individual.[32] An *FY*A* allele carrying c.265T and c.298A (*FY*01W.02*) was described in a white blood donor.[33] Consistent with its known weakening effect on Fy^b antigen, we observed a weakening in the expression of Fy^a in those samples carrying c.265T. In contrast, we did not see an effect on Fy antigen expression by *ACKR1* rs118062001 by flow cytometry. Even though there was no apparent effect on antigen expression, the Duffy protein is a known ligand for *Plasmodium vivax,* which is endemic in this region and malaria is more prevalent in the north and the northwest of Thailand. The prevalence of *FY*01* is 89.1% in our cohort and one may speculate that modification of this protein by c.265C > T or by c.199C > T may reduce susceptibility to invasion. The polymorphism c.265T on an *FY*B* background has been shown to reduce invasion by *P. vivax*[34] but the effect of c.199T and c.265T on a *FY*A* allele remains to be investigated.

Vel antigen expression on RBCs is known to vary considerably from one individual to another. This is influenced by the rs1175550 genotype, and individuals homozygous for the common rs1175550A polymorphism show weaker expression.[22] In this study, all but seven samples tested were rs1175550A homozygotes and showed weak but variable expression of Vel as determined by flow cytometry. This might explain the curious findings of Chandanayingyong and colleagues,[8] who identified four of 328 samples (1.2%) as Vel-negative by serology. Given that antigen expression is so variable, we suspect the Vel-negative samples reported might, in fact, have been weakly Vel-positive.

In summary, we have analyzed the blood group genotypes of blood donors from two different regions of Thailand and examined potential susceptibility markers for malaria, known to be endemic in this region.

ACKNOWLEDGMENTS

We thank Dr Janine Robb (Quotient Ltd.), John MacEachern, (Immucor, Inc.), Lex Visser (Merck Millipore), Dr Yannic Danger

(EFS Rennes), and Dr Makoto Uchikawa (Japanese Red Cross) for kindly providing reagent antibodies used in this study. We thank Nadine Trost for her excellent technical assistance and Dr Mattias Möller for bioinformatics support. We thank Dr Kristina Persson for helpful discussion regarding the project and the manuscript. PJ, CG, AK, MLO, and JRS conceived the project; CB and SC collected samples and performed preliminary donor testing; PJ, SM, MD, and JRS performed the experimental work; PJ, CG, SM, MLO, and JRS interpreted the data. PJ and JRS wrote the manuscript. All authors contributed to and edited the final manuscript.

CONFLICT OF INTEREST

The authors have disclosed no conflicts of interest.

REFERENCES

1. Moulds JM, Nowicki S, Moulds JJ, et al. Human blood groups: incidental receptors for viruses and bacteria. Transfusion 1996;36:362-74.
2. Cooling L. Blood groups in infection and host susceptibility. Clin Microbiol Rev 2015;28:801-70.
3. 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.
4. Weatherall DJ. Genetic variation and susceptibility to infection: the red cell and malaria. Br J Haematol 2008;141:276-86.
5. Cox-Singh J, Davis TM, Lee KS, et al. *Plasmodium knowlesi* malaria in humans is widely distributed and potentially life threatening. Clin Infect Dis 2008;46:165-71.
6. Satchwell TJ. Erythrocyte invasion receptors for *Plasmodium falciparum*: new and old. Transfus Med 2016;26:77-88.
7. Zimmerman PA, Ferreira MU, Howes RE, et al. Red blood cell polymorphism and susceptibility to *Plasmodium vivax*. Adv Parasitol 2013;81:27-76.
8. Chandanayingyong D, Sasaki TT, Greenwalt TJ. Blood groups of the Thais. Transfusion 1967;7:269-76.
9. Fongsarun J, Nuchprayoon I, Yod-in S, et al. Blood groups in Thai blood donors. Thai J Hematol Transfus Med 2002;12: 277-86.
10. Kaset C, Nathalang S, Leetrakool N, et al. Detection of MNS hybrid molecules (GP.Hut, GP.Mur, GP.Hop, GP.Bun and GP.HF) in central and northern Thai blood donors. Thai J Hematol Transfus Med 2015;25:101-6.
11. Palacajornsuk P, Nathalang O, Tantimavanich S, et al. Detection of MNS hybrid molecules in the Thai population using PCR-SSP technique. Transfus Med 2007;17:169-74.
12. Perreault J, Lavoie J, Painchaud P, et al. Set-up and routine use of a database of 10,555 genotyped blood donors to facilitate the screening of compatible blood components for alloimmunized patients. Vox Sang 2009;97:61-8.
13. Flegel WA, Gottschall JL, Denomme GA. Implementing mass-scale red cell genotyping at a blood center. Transfusion 2015;55:2610-5. quiz 09.
14. Gassner C, Meyer S, Frey BM, et al. Matrix-assisted laser desorption/ionisation, time-of-flight mass spectrometry-based blood group genotyping—the alternative approach. Transfus Med Rev 2013;27:2-9.
15. Meyer S, Trost N, Frey BM, et al. Parallel donor genotyping for 46 selected blood group and 4 human platelet antigens using high-throughput MALDI-TOF mass spectrometry. Methods Mol Biol 2015;1310:51-70.
16. Storry JR, Joud M, Christophersen MK, et al. Homozygosity for a null allele of SMIM1 defines the Vel-negative blood group phenotype. Nat Genet 2013;45:537-41.
17. Olsson ML, Chester MA. Polymorphism and recombination events at the ABO locus: a major challenge for genomic ABO blood grouping strategies. Transfus Med 2001;11:295-313.
18. Hosseini-Maaf B, Hellberg A, Chester MA, et al. An extensive PCR-ASP strategy for clinical ABO blood group genotyping that avoids potential errors caused by null, subgroup and hybrid alleles. Transfusion 2007;47: 2110-25.
19. Storry JR, Poole J, Condon J, et al. Identification of a novel hybrid glycophorin gene encoding GP.Hop. Transfusion 2000;40:560-5.
20. Hellberg A, Pesjak E, Karlsson J, et al. A novel RHCE*02 allele, containing the single-nucleotide change c.460A>G, encodes weakened expression of C and e antigens. Transfusion 2016;56:2391-2.
21. Hsu K, Lin YC, Chang YC, et al. A direct blood polymerase chain reaction approach for the determination of GP.Mur (Mi.III) and other Hil+ Miltenberger glycophorin variants. Transfusion 2013;53:962-71.
22. Christophersen MK, Joud M, Ajore R, et al. SMIM1 variants rs1175550 and rs143702418 independently modulate Vel blood group antigen expression. Sci Rep 2017;7:40451
23. Möller M, Jöud M, Storry JR, et al. Erythrogene: a database for in-depth analysis of the extensive variation in 36 blood group systems in the 1000 Genomes Project. Blood Adv 2016;1:240-9.
24. Storry JR, Reid ME, Fetics S, et al. Mutations in GYPB exon 5 drive the S-s-U+var phenotype in persons of African descent: implications for transfusion. Transfusion 2003;43:1738-47.
25. Pham BN, Peyrard T, Juszczak G, et al. Analysis of RhCE variants among 806 individuals in France: considerations for transfusion safety, with emphasis on patients with sickle cell disease. Transfusion 2011;51:1249-60.
26. Haer-Wigman L, Stegmann TC, Solati S, et al. Impact of genetic variation in the SMIM1 gene on Vel expression levels. Transfusion 2015;55:1457-66.
27. Storry JR, Peyrard T. The Vel blood group system: a review. Immunohematology 2017;33:56-9.
28. Crosnier C, Bustamante LY, Bartholdson SJ, et al. Basigin is a receptor essential for erythrocyte invasion by *Plasmodium falciparum*. Nature 2011;480:534-7.
29. Noizat-Pirenne F, Le Pennec PY, Mouro I, et al. Molecular background of D(C)(e) haplotypes within the white population. Transfusion 2002;42:627-33.

30. Kupatawintu P, Emthip M, Sungnoon D, et al. Unexpected antibodies of patients' blood samples sent for testing at NBC.TRCS. J Hematol Transfus Med 2010;20:255-62.
31. Olsson ML, Smythe JS, Hansson C, et al. The Fy(x) phenotype is associated with a missense mutation in the Fy(b) allele predicting Arg89Cys in the Duffy glycoprotein. Br J Haematol 1998;103:1184-91.
32. Arndt PA, Horn T, Keller JA, et al. First example of an FY*01 allele associated with weakened expression of Fya on red blood cells. Immunohematology 2015;31: 103-7.
33. Lopez GH, Condon JA, Wilson B, et al. A novel FY*A allele with the 265T and 298A SNPs formerly associated exclusively with the FY*B allele and weak Fy(b) antigen expression: implication for genotyping interpretative algorithms. Vox Sang 2015;108:52-7.
34. Albuquerque SR, Cavalcante Fde O, Sanguino EC, et al. FY polymorphisms and vivax malaria in inhabitants of Amazonas State, Brazil. Parasitol Res 2010;106:1049-53.
35. Chandanayingyong D, Bejrachandra S. Studies on anti-Mi-a and the MiIII complex. Vox Sang 1975;29:311-5.
36. Lin-Chu M, Broadberry RE, Chang FJ. The distribution of blood group antigens and alloantibodies among Chinese in Taiwan. Transfusion 1988;28:350-2.
37. Corbel V, Nosten F, Thanispong K, et al. Challenges and prospects for dengue and malaria control in Thailand, Southeast Asia. Trends Parasitol 2013;29:623-33.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1. Distribution of alleles encoding high-prevalence blood group antigens determined by MALDI-TOF MS.
Table S2. Distribution of MNS and Rh blood group antigens: comparison of the results of manual phenotyping versus genotyping.
Table S3. Distribution of MNS blood group antigens. Comparison of the results from manual phenotyping versus genotyping.