Extended blood group profiles for Malays, Chinese, and Indians in Peninsular Malaysia

Che Ghazali Norul Hajar1,2, Zulkafli Zefarina2,3, Nor Suhaila Md. Riffin4, Tuan Hulwani Tuan Mohammad5, Mohd Nazri Hassan2,3, Prathapan Poonachi1, Sabreena Safuan1, Gehad El Ghazali6, Geoffrey Keith Chambers7 and Hisham Atan Edinur1,2,8,9*

Abstract

Background: Blood group antigens are immunogenic polymorphic molecules presented on the surface of RBCs. This study aimed to determine extended blood group profiles (ABO, Rhesus, Kell, Kidd, Duffy, MNS, Cartwright, Dombrock, Colton, Lutheran, and Vel) in Malays, Chinese, and Indians in Peninsular Malaysia.

Results: Here, ABO Type O, DCCee, MNs, and Fy (a+b−) were the most frequent major blood group phenotypes in all three ethnic groups. Other minor blood group systems distributed differently across these ethnic groups, except for the Kell, Lutheran, Cartwright, and Vel blood group systems, where only K−k+, Lu (8+14), Yt (a+b−), and Vel (+) phenotypes were observed. Exact tests of population differentiation generally showed no significant differences between Malays included in the present study vs. other ethnically similar datasets from previous surveys. However, many significant differences were recorded in comparison between blood group datasets from ethnically unrelated populations (Malays vs. Chinese vs. Indians) especially for Rhesus, Kidd, and Duffy blood group systems. A Principal component analysis (PCA) plot showed that population groups from the Peninsular Malaysia map closely together as compared with population groups from other geographical regions.

Conclusions: Overall, our present study has successfully provided an extended blood group profiles for Malays, Chinese, and Indians in Peninsular Malaysia. These new blood group datasets can be used as guidelines for donor recruitment and as reference standards for studying diseases associated with blood group systems.

Keywords: Blood group, RBCs, Peninsular Malaysia, Malays, Chinese, Indians

Background

Blood group antigens are polymorphic and immunogenic molecules found on the surface of red blood cells (RBCs) [1]. Antibodies against blood group antigens can develop naturally following exposure to blood group-like substances from viruses or bacteria or immune antibodies due to gestation or transfusion incompatibilities. Most of the naturally occurring antibodies include anti-Cw, anti-M, and antibodies in the Li, Lewis, and P systems are not clinically significant, except for ABO antibodies. Furthermore, immune antibodies are clinically significant and produced specifically against RBC antigens such as antibodies in the K, Rh, Duffy, and Kidd systems [1–3]. From a total of 360 RBC antigens recognized by the International Society of Blood Transfusion (ISBT), 330 were grouped into 39 blood group systems [4]. However, standard transfusion practice only requires matching for ABO and Rhesus-compatible RBC units between donor and patient because transfusion of ABO and Rhesus-incompatible blood can cause haemolytic transfusion reaction, which is often fatal [3, 5]. Antigens of other blood group systems are not typed unless the units are destined for an alloimmunized patient [6]. In this respect, knowledge about the local distributions of

* Correspondence: edinur@usm.my
1 School of Health Sciences, Universiti Sains Malaysia, Health Campus, Kubang Kerian, Kelantan, Malaysia
2 Hospital Universiti Sains Malaysia, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia
3 Full list of author information is available at the end of the article
wider RBC antigens is important for antibody screening and for searching of compatible blood for these most vulnerable patients [7].

Frequencies of RBC antigens are unique, and considerable differences have been observed between unrelated ethnic groups [8, 9]. This is because, population events such as founder effects, natural selection, and gene flow may lead to changes of RBC antigen distributions in descendant populations, as compared with their related ancestral populations [10–12]. Two previous studies have reported blood group frequencies for several population groups in Peninsular Malaysia [13, 14]. These include for ABO, Rh, Kell, Kidd, and Duffy blood group frequencies in Malays, Chinese, and Indians. However, other medically relevant blood group systems (e.g. Colton and Dombrock) that are also involved in haemolytic transfusion reaction and gestation incompatibilities were not reported in either [15]. In the study reported here, we provide frequency data of 11 blood group systems screened in 222 unrelated volunteer Malay, Chinese, and Indian blood donors living in Peninsular Malaysia. The extended blood group profiles include their scores for several extra blood group systems (Cartwright, Colton, Dombrock, Lutheran, and Vel) that have not previously been reported [13, 14] can now be used as guidelines for donor recruitment and as a reference standard for studying diseases associated with blood group systems [16, 17].

**Methods**

**Study population**

Blood samples were obtained from 222 healthy individuals registered as voluntary donors at local blood transfusion units. All volunteers are unrelated and provided their written informed consent. They were then assigned as Malay (n = 97), Chinese (n = 77), or Indian (n = 48) based on their family pedigree as provided (Criterion: 3 generations with no history of intermarriage). Ethical clearance was obtained from the related Ethics Committees (NMRR-16-1399-31311 (IIR) and IRB no: IRB00010568).

**Phenotyping of ABO and RhDCE antigens**

The ABO and Rh D blood group antigens were typed using the haemaglutination-based tile method. Antisera for ABO and D phenotyping were purchased from bioCSL Pty Ltd. (Parkville, Victoria, Australia), and phenotyping was performed as per the manufacturer’s instructions. Other Rh antigens (C, E, c, e) were determined by tube technique using 5% unwashed red blood cell suspensions in isotonic saline solution with anti-C, anti-c, anti-E, and anti-e monoclonal antibodies (DIAGAST, Loos, France).

**Molecular typing of other blood group systems**

Peripheral blood sample (50 μl) was extracted for genomic DNA using Invisor® Spin Forensic Kit (STRATEC Molecular, Berlin, Germany). Extraction protocols were as per manufacturer’s recommendations and as described earlier [18]. Genomic DNA concentration and purity for downstream PCR applications were estimated with Nanodrop 2000c spectrophotometer (Thermo Scientific, MA, USA).

The DNA samples were then typed for Kell, Kidd, Duffy, Cartwright, Dombrock, Colton, Lutheran, Vel, and MNS blood group systems using polymerase chain reaction assays with sequence-specific primer (PCR-SSP) chemistries and protocols as previously reported by Rozman et al. [19], Rink et al. [20] and Heymann et al. [21], respectively. Amplified products were then fractionated by electrophoresis (on 2% agarose gels stained with ethidium bromide), and PCR product separation patterns were documented by UV photometer (Quantum ST4-1000/20M, VilberLourmat, Deutschland GmbH, Eberhardzell, Germany). In addition to a pair of primers specific to a particular blood group system, each PCR-SSP reaction mixture also contains a special pair of primers (internal positive control) targeting a 429-bp fragment of the human growth hormone (HGH) gene. Blood group type was then scored based on the size of expected PCR-SSP fragment determined using a 100 base pair DNA size standard as a reference (Bioline, London, UK).

**Statistical analysis**

Observed phenotype frequencies were obtained by direct counting. Chi-squared (χ²) analyses were used to test for Hardy–Weinberg equilibrium (HWE) using the formula χ² = Σ [(O − E)²/E], where O and E are observed and expected number of phenotypes and expected phenotypes, respectively. Significant departure from HWE is considered at a P value of <0.05. This P value was then corrected for multiple comparisons by dividing the standard significant P value (0.05) by the total number of blood group phenotypes [22]. Pairs of blood group datasets were compared using exact test with initial significance value set at P < 0.05 [23]. This P value (0.05) was then adjusted to 0.005 (0.05/10, where 10 is a total number of population datasets included for comparison) using Bonferroni correction. Blood group datasets reported in present and previous surveys [13, 24–27] were mapped on a 2-dimensional principal component plot using the algorithm implemented in Multivariate Statistical Software Package 3 (Kovach Computing Services, UK; http://www.kovcomp.com/mvs).

**Results**

Table 1 shows blood group phenotype frequencies determined in 222 Malay, Chinese, and Indian individuals. No
Table 1 Phenotype frequencies and HWE analysis of blood group systems in Malays, Chinese, and Indians

| Blood system | Phenotype | Malays \((n = 97)\) | \(\chi^2\) | \(P\) | Chinese \((n = 77)\) | \(\chi^2\) | \(P\) | Indians \((n = 48)\) | \(\chi^2\) | \(P\) |
|--------------|-----------|----------------|---------|------|----------------|---------|------|----------------|---------|------|
| **ABO**      |           |                |         |      |                |         |      |                |         |      |
| A            | 0.320     | 1.448          | 0.694   |      | 0.200          | 2.699   | 0.440 | 0.250          | 1.669   | 0.643 |
| B            | 0.250     | 0.320          | 0.190   |      |                |         |      |                |         |      |
| AB           | 0.070     | 0.090          | 0.080   |      |                |         |      |                |         |      |
| O            | 0.360     | 0.390          | 0.480   |      |                |         |      |                |         |      |
| **Rh**       |           |                |         |      |                |         |      |                |         |      |
| DCCee        | 0.031     | 1.563          | 0.955   |      | 0.030          | 7.042   | 0.317 | 0.020          | 11.158  | 0.084 |
| DCCee        | 0.639     | 0.650          | 0.540   |      |                |         |      |                |         |      |
| DCCcEE       | 0.010     | 0.000          | 0.000   |      |                |         |      |                |         |      |
| DCCcEe       | 0.180     | 0.170          | 0.150   |      |                |         |      |                |         |      |
| Dccee        | 0.010     | 0.060          | 0.020   |      |                |         |      |                |         |      |
| DCcee        | 0.000     | 0.000          | 0.000   |      |                |         |      |                |         |      |
| DccEe        | 0.040     | 0.030          | 0.060   |      |                |         |      |                |         |      |
| **Kell**     |           |                |         |      |                |         |      |                |         |      |
| K+k−         | 0.000     | na            | na      |      | 0.000          | na      | na   | 0.000          | na      | na   |
| K+k+         | 0.000     | 0.000         | 0.000   |      |                |         |      |                |         |      |
| K−k+         | 1.000     | 1.000         | 1.000   |      |                |         |      |                |         |      |
| **Kidd**     |           |                |         |      |                |         |      |                |         |      |
| Jk (a+b−)    | 0.340     | 2.893          | 0.495   |      | 0.210          | 7.042   | 0.317 | 0.300          | 4.714   | 0.194 |
| Jk (a−b+)    | 0.190     | 0.230          | 0.380   |      |                |         |      |                |         |      |
| Jk (a+b+)    | 0.440     | 0.520          | 0.310   |      |                |         |      |                |         |      |
| Jk (a−b−)    | 0.040     | 0.000          | 0.020   |      |                |         |      |                |         |      |
| **Duffy**    |           |                |         |      |                |         |      |                |         |      |
| Fy (a+b−)    | 0.750     | 0.696          | 0.706   |      | 0.900          | 8.469   | 0.014 | 0.560          | 10.705  | 0.005 |
| Fy (a−b+)    | 0.010     | 0.000          | 0.060   |      |                |         |      |                |         |      |
| Fy (a+b+)    | 0.240     | 0.090          | 0.380   |      |                |         |      |                |         |      |
| Fy (a−b−)    | 0.000     | 0.013          | 0.000   |      |                |         |      |                |         |      |
| **MNS**      |           |                |         |      |                |         |      |                |         |      |
| MNSs         | 0.130     | 0.083          | 1.000   |      | 0.132          | 0.340   | 1.000 | 0.139          | 0.750   | 0.999 |
| MNS          | 0.101     | 0.101          | 0.100   |      |                |         |      |                |         |      |
| MNs          | 0.209     | 0.221          | 0.205   |      |                |         |      |                |         |      |
| MSs          | 0.084     | 0.078          | 0.090   |      |                |         |      |                |         |      |
| MS           | 0.053     | 0.048          | 0.052   |      |                |         |      |                |         |      |
| Ms           | 0.163     | 0.167          | 0.156   |      |                |         |      |                |         |      |
| NSs          | 0.070     | 0.065          | 0.076   |      |                |         |      |                |         |      |
| NS           | 0.040     | 0.035          | 0.038   |      |                |         |      |                |         |      |
| Ns           | 0.150     | 0.154          | 0.142   |      |                |         |      |                |         |      |
| **Dombrock** |           |                |         |      |                |         |      |                |         |      |
| Do (a+b−)    | 0.120     | 1.573          | 0.455   |      | 0.090          | 0.264   | 0.876 | 0.100          | 1.496   | 0.473 |
| Do (a−b+)    | 0.410     | 0.510          | 0.560   |      |                |         |      |                |         |      |
| Do (a+b+)    | 0.470     | 0.400          | 0.340   |      |                |         |      |                |         |      |
| **Colton**   |           |                |         |      |                |         |      |                |         |      |
significant deviations from HWE were observed except for the Duffy blood group in Chinese and Indians. However, only the latter deviated significantly from HWE after Bonferroni correction. The ABO Type O, DCcEe, MNs, and Fy (a+b−) phenotypes were recorded to be the most frequent types (0.360–0.480, 0.540–0.650, 0.205–0.221, and 0.560–0.900, respectively) in all three survey groups. In contrast, the Jk (a+b+) variant of the Kidd blood group system was the most common phenotype in Malays (0.440) and Chinese (0.520), while Jk (a−b+) was the most common phenotype in groups. In contrast, the Jk (a+b+) variant of the Kidd blood group system was the most common phenotype in Malays (0.440) and Chinese (0.520), while Jk (a−b+) was the most common phenotype in groups. In contrast, the Jk (a+b+) variant of the Kidd blood group system was the most common phenotype in groups. In contrast, the Jk (a+b+) variant of the Kidd blood group system was the most common phenotype in groups.

The Malays, Chinese, and Indians (0.510 and 0.560, respectively) in all three survey groups. In contrast, the Jk (a+b+) variant of the Kidd blood group system was the most common phenotype in Malays (0.440) and Chinese (0.520), while Jk (a−b+) was the most common phenotype in groups. In contrast, the Jk (a+b+) variant of the Kidd blood group system was the most common phenotype in groups.

Table 1 Phenotype frequencies and HWE analysis of blood group systems in Malays, Chinese, and Indians (Continued)

| Blood system | Phenotype | Malays (n = 97) | Chinese (n = 77) | Indians (n = 48) |
|--------------|-----------|----------------|-----------------|-----------------|
|              |           | χ²            | P               | χ²             | P               |
| Co (a+b−)    | 0.990     | na            | na              | 1.000          | 1.000           |
| Co (a−b+)    | 0.000     | na            | na              | 0.000          | na              |
| Co (a+b+)    | 0.010     | na            | na              | 0.000          | na              |
| Lutheran     |           |               |                 |                |                 |
| Lu (8+14−)   | 1.000     | na            | na              | 1.000          | 1.000           |
| Lu (8−14+)   | 0.000     | na            | na              | 0.000          | na              |
| Cartwright   |           |               |                 |                |                 |
| Yt (a+b−)    | 1.000     | na            | na              | 1.000          | 1.000           |
| Yt (a−b+)    | 0.000     | na            | na              | 0.000          | na              |
| Vel          |           |               |                 |                |                 |
| Vel (+)      | 1.000     | na            | na              | 1.000          | 1.000           |
| Vel (−)      | 0.000     | na            | na              | 0.000          | na              |

ABO and Rh D, C, c, E, e were obtained by serological typing, while other blood groups were obtained by PCR-SSR. HWE analysis is considered significantly different if their P value is < 0.050

n sample size, na not applicable, P significance level (P value), HWE Hardy–Weinberg equilibrium

*P value for HWE analysis in Duffy blood group was then adjusted to < 0.013 using Bonferroni correction

Discussion

Peninsular Malaysia is a region occupied by a diverse range of human population groups including Malays, Chinese, Indians, and three Orang Asli groups (Semang, Senoi, and Proto-Malays). Multidisciplinary data have shown different ancestral sources for the three divisions of Orang Asli [28–30]. The Semang represent descendants of the first ‘Out of Africa’ lineage of Homo sapiens migrants [31, 32], the Senoi come from the north and share greater affinities with populations in Indo-China
| Blood Group | Present study | Previous studies |
|-------------|---------------|------------------|
|             | Malays (n=97) | Chinese (n=77) | Indians (n=48) | Banjar Malays [14] (n=30) | Jawa Malays [14] (n=30) | Maindaling [14] (n=30) | Kelantan Malays [14] (n=30) | Malays [13] (n=200) | Chinese [13] (n=274) | Indians [13] (n=120) |
| ABO         |               |                  |               |                          |                        |                        |                          |                        |                          |                          |
| A           | 0.320         | 0.200            | 0.250         | 0.230                    | 0.270                  | 0.200                  | 0.400                    | 0.310                  | 0.270                  | 0.200                  |
| B           | 0.250         | 0.320            | 0.190         | 0.470                    | 0.300                  | 0.500                  | 0.240                    | 0.280                  | 0.230                  | 0.370                  |
| AB          | 0.070         | 0.090            | 0.080         | 0.100                    | 0.130                  | 0.100                  | 0.030                    | 0.080                  | 0.110                  | 0.070                  |
| O           | 0.360         | 0.390            | 0.480         | 0.200                    | 0.300                  | 0.200                  | 0.330                    | 0.350                  | 0.380                  | 0.370                  |
| Rhesus      |               |                  |               |                          |                        |                        |                          |                        |                          |                          |
| DCCEe       | 0.031         | 0.030            | 0.020         | 0.030                    | 0.000                  | 0.030                  | 0.000                    | 0.035                  | 0.020                  | 0.008                  |
| DCCoE       | 0.639         | 0.650            | 0.540         | 0.810                    | 0.840                  | 0.430                  | 0.700                    | 0.615                  | 0.536                  | 0.500                  |
| DCcEE       | 0.010         | 0.000            | 0.000         | 0.000                    | 0.000                  | 0.000                  | 0.000                    | 0.010                  | 0.000                  | 0.000                  |
| DCcEe       | 0.180         | 0.170            | 0.150         | 0.130                    | 0.130                  | 0.330                  | 0.170                    | 0.150                  | 0.248                  | 0.125                  |
| DccEE       | 0.010         | 0.000            | 0.020         | 0.000                    | 0.000                  | 0.030                  | 0.000                    | 0.010                  | 0.090                  | 0.008                  |
| DCcEe       | 0.090         | 0.060            | 0.210         | 0.030                    | 0.030                  | 0.170                  | 0.100                    | 0.150                  | 0.066                  | 0.230                  |
| DccEe       | 0.040         | 0.030            | 0.060         | 0.000                    | 0.000                  | 0.000                  | 0.030                    | 0.025                  | 0.030                  | 0.050                  |
| dcccEe      | 0.000         | 0.000            | 0.000         | 0.000                    | 0.000                  | 0.000                  | 0.000                    | 0.005                  | 0.007                  | 0.075                  |
| Kell        |               |                  |               |                          |                        |                        |                          |                        |                          |                          |
| K+k−        | 0.000         | 0.000            | 0.000         | 0.000                    | 0.000                  | 0.000                  | 0.000                    | 0.010                  | 0.010                  | 0.010                  |
| K+k+        | 0.000         | 0.000            | 0.000         | 0.000                    | 0.000                  | 0.000                  | 0.000                    | 0.010                  | 0.000                  | 0.010                  |
| K−k+        | 1.000         | 1.000            | 1.000         | 1.000                    | 1.000                  | 1.000                  | 1.000                    | 0.990                  | 0.990                  | 0.980                  |
| Kidd        |               |                  |               |                          |                        |                        |                          |                        |                          |                          |
| Jk(a+b−)    | 0.340         | 0.210            | 0.300         | 0.300                    | 0.170                  | 0.330                  | 0.200                    | 0.200                  | 0.360                  | 0.240                  |
| Jk(a−b+)    | 0.190         | 0.230            | 0.380         | 0.370                    | 0.200                  | 0.230                  | 0.300                    | 0.180                  | 0.250                  | 0.200                  |
| Jk(a+b+)    | 0.440         | 0.520            | 0.310         | 0.470                    | 0.470                  | 0.570                  | 0.500                    | 0.430                  | 0.510                  | 0.430                  |
| Jk(a−b−)    | 0.040         | 0.026            | 0.020         | 0.000                    | 0.000                  | 0.000                  | 0.000                    | 0.040                  | 0.000                  | 0.020                  |
| Duffy       |               |                  |               |                          |                        |                        |                          |                        |                          |                          |
| Fy(a+b−)    | 0.750         | 0.900            | 0.560         | 0.770                    | 0.570                  | 0.930                  | 0.900                    | 0.740                  | 0.850                  | 0.410                  |
| Fy(a−b+)    | 0.010         | 0.013            | 0.060         | 0.000                    | 0.070                  | 0.000                  | 0.000                    | 0.030                  | 0.010                  | 0.140                  |
| Fy(a+b+)    | 0.240         | 0.090            | 0.380         | 0.230                    | 0.370                  | 0.070                  | 0.100                    | 0.230                  | 0.130                  | 0.450                  |
| Fy(a−b−)    | 0.000         | 0.000            | 0.000         | 0.000                    | 0.000                  | 0.000                  | 0.000                    | 0.010                  | 0.000                  | 0.000                  |
| MNS         |               |                  |               |                          |                        |                        |                          |                        |                          |                          |
| MNSS        | 0.130         | 0.132            | 0.139         | na                       | na                     | na                     | na                       | 0.099                  | 0.084                  | 0.132                  |
| MNS         | 0.101         | 0.101            | 0.100         | na                       | na                     | na                     | na                       | 0.076                  | 0.073                  | 0.082                  |
Table 2 Blood group frequencies for Malays, Chinese, and Indians and other population groups in Peninsular Malaysia (Continued)

| Blood Group | Present study | Previous studies |
|-------------|---------------|------------------|
|             | Malays (n=97) | Chinese (n=77)   |
|             | Indians (n=48) |                  |
| MNs         | 0.209         | 0.221            |
|             | 0.020         | 0.010            |
|             | 0.016         | 0.010            |
|             | 0.010         | 0.010            |
| MSs         | 0.038         | 0.040            |
|             | 0.012         | 0.012            |
|             | 0.012         | 0.012            |
|             | 0.012         | 0.012            |
| MS          | 0.090         | 0.090            |
|             | 0.010         | 0.010            |
|             | 0.010         | 0.010            |
|             | 0.010         | 0.010            |
| Ms          | 0.150         | 0.150            |
|             | 0.130         | 0.130            |
|             | 0.130         | 0.130            |
|             | 0.130         | 0.130            |
| NSs         | 0.070         | 0.070            |
|             | 0.060         | 0.060            |
|             | 0.060         | 0.060            |
|             | 0.060         | 0.060            |
| NS          | 0.040         | 0.040            |
|             | 0.030         | 0.030            |
|             | 0.030         | 0.030            |
|             | 0.030         | 0.030            |
| Ns          | 0.150         | 0.150            |
|             | 0.130         | 0.130            |
|             | 0.130         | 0.130            |
|             | 0.130         | 0.130            |

**Footnotes:**

- *n* sample size, *na* not available

---

MNs 0.209 0.221 0.205 na na na na 0.212 0.225 0.140
MSs 0.084 0.078 0.090 na na na na 0.088 0.069 0.143
MS 0.053 0.048 0.052 na na na na 0.065 0.058 0.093
Ms 0.163 0.167 0.156 na na na na 0.201 0.210 0.151
NSs 0.070 0.065 0.076 na na na na 0.057 0.050 0.100
NS 0.040 0.035 0.038 na na na na 0.033 0.039 0.050
Ns 0.150 0.154 0.142 na na na na 0.169 0.191 0.108
whereas the Proto-Malays are linked to putative Austronesian ancestors in Taiwan [34]. In contrast, several other Austronesian Malay sub-ethnic groups including Jawa, Banjar, Bugis, Minangkabau, and Rawa migrated to Peninsular Malaysia as recently as around 300–400 years ago from nearby regions like Java, Kalimantan, Sulawesi, and Sumatera [35].

The Orang Asli and a few of the Malay sub-ethnic groups have now become significant minorities in Peninsular Malaysia [36]. The Malays themselves represent 60% of the Malaysian population and are inferred on the basis of genetic evidence to have arisen from Proto-Malays by admixture with Malay sub-ethnic groups and other ethnicities. The latter include Chinese and Indians who settled in Peninsular Malaysia during historic times via large-scale migration during the British colonial era in the 19th century [37–39].

Blood group data for the Malays generated from the present study (Table 1) showed no significant differences between the Malays and Malay sub-ethnic groups (Table 2 and online supplementary Table S1) reported by Musa et al. [13] and Abd Ghani et al. [14]. These findings support a close genetic relationship between these ancestrally related populations as inferred from other studies [40]. However, we observed significant differences between the Malays vs. both Chinese and Indians (online supplementary Table S1). This is not surprising as these populations are ancestrally unrelated. In addition, the gene pool of present day population groups in Peninsular Malaysia might have been further shaped by evolutionary processes such as natural selection and gene flow since they settled in Peninsular Malaysia. In particular, a significant deviation from HWE was found for the Duffy blood group system in Indians even after Bonferroni correction (Table 1). This observation might be associated with the relatively small number of Indian individuals included in the present survey as well as the effects of natural selection. Several Duffy blood group phenotypes are known to be associated with diseases [41, 42] and identified as receptors for malarial parasites [43]. However, differences between population groups in Peninsular Malaysia were not so obvious in the PCA analysis due to scaling issues (Fig. 1). Here, the Malays, Chinese, and Indians are plotted together and are well separated with the population groups from other geographical regions (Europe and Oceania). The Malays are also separated with their other linguistically related populations including Maori and Polynesians. Here we note that evidence from multidisciplinary studies do show a common origin for these Austronesian-speaking populations [26, 44–49]. They are all descendants of Taiwan aboriginals who migrated out into the Asia-Pacific region 5000–7000 years ago [50, 51]. It can be concluded that our new blood group data reflect geographical proximity and ancestral relationships. However, our findings should be interpreted with some caution as the PCA analysis was carried out using only data for three of their blood group systems (ABO, Rhesus, and MNS) as many of the other blood group systems were not characterized in the available reference populations.

Blood group population datasets reported here including those from the previous studies [13, 14] are of medical significance as blood group antigens are...
determinants of transfusion and gestation compatibilities. The data show the polymorphic nature of blood group systems in various population groups in Malaysia (Tables 1 and 2). These include those blood group systems that are reported for the first time here such as Dombrock and Colton. Other blood group systems in Malays, Chinese, and Indians are monomorphic, but we cannot rule out that they might actually be polymorphic with rare low-frequency alleles, and further work using larger sample sizes are needed to reveal the true extent of allelic variations in Malaysia. Nonetheless, the available blood group data showed that there are potential risks of transfusion and gestation alloimmunization in Malaysia, particularly for polymorphic blood group systems like ABO, Rhesus, Kidd, Duffy, MNS, Colton, and Dombrock. Our predictions are supported by Al-Joudi et al. [52] and Ismahanisa Ismail et al. [53] where alloantibodies against ABO, Rhesus, Kidd, and Duffy antigens are commonly observed among transfused patients in Malaysia.

Overall, our study further supports the importance of blood group data for studying ancestry and health [43]. Several issues and limitations highlighted here including the availability or otherwise of reference population datasets for comparison purposes, number of blood group loci screened in the earlier studies [13, 14] and sample size can be used as guidelines for future study of blood groups in other populations, including in Malaysia [54]. Nonetheless, some blood group datasets including those for Cartwright, Colton, Dombrock, Lutheran, and Vel in Malays, Chinese, and Indians were reported here for the first time and should provide important information for building better healthcare services (e.g. an improved pool of blood group-typed donors and for prenatal screening) in Peninsular Malaysia and can be used as reference datasets for future investigations of diseases associated with blood groups [3, 41, 53].

Conclusions
Overall, our present study has successfully provided extended blood group profiles for Malays, Chinese, and Indians in Peninsular Malaysia. Our statistical analyses have revealed marked differences between the new and previously reported blood group datasets. These observations support the picture of complex genetic structure present in Peninsular Malaysia. The new blood group datasets reported here can be used to guide for donor recruitment strategies and provide an important and expanded reference standard for studying diseases associated with blood group systems.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s43042-020-00096-y.

Additional file 1: Table S1. Exact tests population differentiation (P-value) between pairs of population groups in Peninsular Malaysia.

Additional file 2: Table S2. Blood group frequency data used for PCA data mapping.

Abbreviations
RBCs: Red blood cells; PCA: Principal component analysis; ISBT: International Society of Blood Transfusion; Rh: Rhesus; PCR-SSP: Polymerase chain reaction reaction-sequence-specific primer; HGH: Human growth hormone; HWE: Hardy–Weinberg equilibrium; n: Sample size; na: Not available; na: Not available/applicable

Acknowledgements
We acknowledge with thanks all the participants who voluntarily donated their blood and the invaluable help and assistance given by doctors and nurses during sample collection. Geoffrey K. Chambers thanks Victoria University of Wellington for Alumnus Scholar support.

Authors’ contributions
CGNH designed and performed the research, collected the data, analysed the data, and prepared, edited, and reviewed the manuscript. ZZ helped to design the study, data acquisition, and obtained ethical permit from the Ministry of Health, Malaysia, NMR, and THTM helped with study design, sampling, and data collection. PP help performed the experiment. MNH, SF, GEG, and GKC designed the research, advised on statistics, edited the manuscript, and HAE designed the research, advised on statistics, edited the manuscript, funded the research, and obtained ethical approval from Human Ethics Committee, Universiti Sains Malaysia. The authors have read and approved the final manuscript.

Funding
We are most grateful to the Universiti Sains Malaysia (USM) for research facilities and financial support (Short Term Grant: 304/PPSK/6315142 and Bridging Grant: 304/PPSK/6316132).

Availability of data and materials
All data generated or analysed during this study are included in this published article (and its supplementary information files).

Ethics approval and consent to participate
This study was conducted in accordance with relevant ethical standards, and the study protocol was approved (USAM/EB/16/013191) by the Human Research Ethics Committee of University Sains Malaysia (registered as IRB no: IRB000100568) and the Medical Research and the Ethics Committee, Ministry of Health, Malaysia (NMRR-16-1399-31311 (IRB)), and the subjects have signed and informed written consent.

Consent for publication
Not applicable

Competing interests
The authors declare that they have no competing interests.

Authors’ details
1School of Health Sciences, Universiti Sains Malaysia, Health Campus, Kubang Kerian, Kelantan, Malaysia. 2Hospital Universiti Sains Malaysia, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia. 3School of Medical Sciences, Universiti Sains Malaysia, Health Campus, Kubang Kerian, Kelantan, Malaysia. 4Hospital Sultan Haji Ahmad Shah, Temerloh, Pahang, Malaysia. 5Hospital Seberang Jaya, Perai, Pulau Pinang, Malaysia. 6Department of Immunology, Pathology and Laboratory Medicine Services, Sheikh Khalifa Medical City, Abu Dhabi, United Arab Emirates. 7School of Biological Sciences, Victoria University of Wellington, Wellington, New Zealand. 8Environmental Futures Research Institute, Griffith University, Nathan, Queensland, Australia. 9Institute of Tropical Biodiversity and Sustainable Development, Universiti Malaysia Terengganu, Kuala Nerus, Terengganu, Malaysia.
49. Bellwood P, Chambers GK, Ross M, Hung H-C (2011) Are ‘cultures’ inherited? Multidisciplinary perspectives on the origins and migrations of Austronesian-speaking peoples prior to 1000 BC. In: Roberts BW, Marc VL (eds) Investigating archaeological cultures. Springer, Cham, pp 321–354

50. Norhalifah HK, Syaza FH, Chambers G, Edinur HA (2016) The genetic history of peninsular Malaysia. Gene 586(1):129–135

51. Chambers GK, Edinur HA, Dunn PP (2016) New insights into ancestry and health of Polynesians and New Zealand Maori. New Zealand Sci Rev 73(1): 3–10

52. Al-Joudi F, Ali AB, Ramli MB, Ahmed S, Ismail M (2011) Prevalence and specificities of red cell alloantibodies among blood recipients in the Malaysian state of Kelantan. Asian J Transfus Sci 5(1):42–45

53. Ismahanisa Ismail W, Al-Hassan FM, Shatar A, Roslan M, Rosdan NH, Zainuddin N et al (2016) Prevalence of red cell alloantibodies among multi transfused dependent thalassemia patient in the Malaysian state of Penang. J Sci Res Development 3:62–66

54. Hakim HM, Lalung J, Narayanan S, Khaw NR, Chamber GK, Edinur HA (2018) A new analysis of population history in Sabah and Sarawak. Global J Bus Soc Sci Review 6(4):106–113

**Publisher’s Note**
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.