GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Heterogeneity and diversity of *ABO* and *Rh* blood group genes in select Saudi Arabian populations

E.S. AlSuhaibani[1], N.A. Kizilbash[2] and S. Malik[3]

[1]Department of Zoology, College of Science, King Saud University, Riyadh, Saudi Arabia
[2]Department of Biochemistry, Faculty of Medicine and Applied Medical Sciences, Northern Border University, Arar, Saudi Arabia
[3]Human Genetics Program, Department of Animal Sciences, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad, Pakistan

Corresponding authors: S. Malik
E-mail: malik@qau.edu.pk / fsd707@gmail.com

Genet. Mol. Res.14 (3): 7850-7863 (2015)
Received November 10, 2014
Accepted April 6, 2015
Published July 14, 2015
DOI http://dx.doi.org/10.4238/2015.July.14.11

ABSTRACT. In order to investigate the diversity of *ABO* and *Rh* blood group genes in the Saudi Arabian population, we assembled the phenotypic data of approximately 66,000 subjects from ten representative Saudi populations: Al-Khobar, Riyadh, Tabuk/Madina Al-Munawaara, Jeddah, Abha, South region, Sakaka, Domah, Al-Qurayat, and Sweer. The frequencies of p[A], q[B], and r[O] alleles at the *ABO* locus were observed to be 0.1688, 0.1242, and 0.7070, respectively, and the frequency of the *D* allele at the *Rh* locus was 0.7138. The heterozygosities at the *ABO* and *Rh* loci were 0.4563 and 0.4086, respectively, while the combined heterozygosity was 0.4324. Homogeneity tests revealed the population of Abha to be the most heterogeneous while that of Tabuk/Madina was found to be the least heterogeneous. Homogeneity was higher among the Northern populations while Southern populations demonstrated subdivisions and stratification. Gene diversity analyses

yielded a total heterozygosity value of 0.4449. The coefficient of gene differentiation was 0.0090. Nei's genetic distance analyses showed that there was close affinity between the populations of Al-Khobar and Riyadh. The largest differences were observed between the populations of Sakaka and Domah. Furthermore, negative correlations were found between p[A] and r[O] alleles, and between q[B] and r[O] alleles at the *ABO* locus. Clinal analyses revealed that the r[O] allele showed an increasing trend from North-East to South-West, and conversely the q[B] allele exhibited a decreasing trend at these coordinates. These analyses present interesting aspects of the blood group allele distribution across the geography of Saudi Arabia.

Key words: Saudi population; Genetic heterogeneity; Gene diversity; ABO; Rh; Blood groups

INTRODUCTION

Genetic studies regarding the populations of Saudi Arabia are rare. There are several preliminary reports available on the ABO and Rh blood group polymorphisms in Saudi subpopulations. Bashwari et al. (2001) reported blood group phenotypic data from the Eastern region of Saudi Arabia and showed that blood type "O" was the most common. Al-Himaidi and Umar (2002) generated a phenotypic record of Saudi citizens originating from various regions. The authors presented the distribution of phenotypic proportions of blood types but allelic frequencies and heterozygosities at the *ABO* and *Rh* loci were not mentioned. Sarhan et al. (2009) also observed the blood group proportions in the Southwestern Saudi population. Eweidah et al. (2011) attempted to report the distribution of blood groups in four cities of the Al-Jouf region in the North of Saudi Arabia. Overall, these studies present a disjointed picture of the distribution of blood group polymorphisms in Saudi Arabia.

Most of the previous studies employing molecular markers in Saudi populations have presented regional data and incorporated small sample sizes. An early study by Sinha et al. (1999) examined eight short tandem repeats in Saudi populations. Recent studies have employed molecular makers in order to observe the diversity in the Saudi population. Alanazi et al. (2013) explored selected polymorphisms in DNA repair genes in the population of the central region of Saudi Arabia and observed a unique pattern of alleles in that population. In another study, Mustafa et al. (2014) carried out molecular genotyping of the *Rh* locus and observed the RHD and RHCE variants in Turabah Province. For the RHD variant, 86.5% subjects were found to be RHD positive and 13.5% were RHD negative; the most common phenotype was DcE4+. Abu-Amero et al. (2013) studied four single nucleotide polymorphisms in two genes, *TLR2* and *TLR4*, and genotyped 201 unrelated Saudi individuals. The authors concluded that regional variation at these loci could have been shaped by both evolutionary pressures and bidirectional human migrations.

The most frequently studied genetic makers in the Saudi populations are the *ABO* and *Rh* blood group loci (Bashwari et al., 2001; Al-Himaidi and Umar, 2002). In order to appreciate the diversity of these markers across the geography of Saudi Arabia, we have carried out a detailed study and have investigated genetic heterogeneity and gene diversity at the *ABO* and *Rh* loci in the representative Saudi populations.

MATERIAL AND METHODS

Subjects

Blood group phenotypic data of >66,000 subjects from ten representative Saudi populations were assembled from publically available reports (Table 1) (Bashwari et al., 2001; Al-Himaidi and Umar, 2002; Sarhan et al., 2009; Eweidah et al., 2011). The Eastern province was represented by Al-Khobar; the middle province was represented by Riyadh, the North-Western region by Tabuk/Madina Al-Munawaara, the Western region by Jeddah, the South-West by Abha, and the Northern region was represented by Sakaka, Domah, Al-Qurayat, and Sweer cities (Figure 1). Only records with sample sizes ≥100 were considered. There were few populations for which multiple data sets were available in the published literature. Only the most recent records or data with the largest sample size were retained (Bashwari et al., 2001). Tribe-specific data were not included in this study. Blood group data for cohorts of diseased/morbid subjects were not included in the analyses. Data from adjoining populations, such as Bahrain, United Arab Emirates, and Yemen, were also not included in this study.

Table 1. ABO and Rh blood group proportions in ten Saudi Arabian populations.

Population	Sample (N)	Phenotype (%)						Reference
		A	B	AB	O	Rh+	Rh-	
Al-Khobar	57,396	26.43	18.44	4.07	51.07	92.10	7.90	Bashwari et al. (2001)
Riyadh	3324	25.99	20.01	4.00	50.00	92.00	8.00	Bashwari et al. (2001)
Jeddah	3924	27.01	16.00	3.49	53.49	92.00	8.00	Albaz (2001)
South region	291	33.68	11.34	2.75	52.23	89.69	10.31	Al-Himaidi and Umar (2002)
Abha	944	33.37	6.04	3.81	56.78	92.80	7.20	Sarhan et al. (2009)
Tabuk/Madina Al-Munawaara	166	30.12	12.05	4.82	53.01	92.17	7.83	Ozsoylu and Alhejaili (1987)
Sakaka	100	23.00	24.00	5.00	48.00	86.00	14.00	Eweidah et al. (2011)
Domah	100	29.00	30.00	6.00	35.00	94.00	6.00	Eweidah et al. (2011)
Al-Qurayat	100	29.00	22.00	9.00	40.00	92.00	8.00	Eweidah et al. (2011)
Sweer	100	29.00	26.00	9.00	36.00	93.00	7.00	Eweidah et al. (2011)
Total*	66,445	26.70	19.09	4.21	50.00	91.81	8.19	

*Proportions in total Saudi population were based on data from 25 sub-populations.

Descriptive analyses of the phenotypic data were performed and blood groups were presented as percentages (Gerstman, 2008). Allele frequencies were calculated at the *ABO* locus using the maximum likelihood method (Mather, 1964). The frequency of the *Rh*(d) allele was calculated from the square-root of the frequency of Rh(-) phenotypes. Prior to the analyses, Hardy-Weinberg equilibrium (HWE) was checked through goodness-of-fit tests (Silva, 2002; Malik and Amin-ud-Din, 2013). Heterozygosity at the individual *ABO* and *Rh* loci and the combined heterozygosities were calculated using the method of Nei (1987).

To observe the variability at the blood group systems, coefficients of variance (CoV) were calculated across the phenotypic and allelic estimates. Homogeneity was tested between the populations (Neel and Schull, 1954). The *Z*-test was employed to check the significance of the heterogeneity of blood group proportions among the studied populations (Gerstman, 2008). Homogeneity tests for allele frequencies were carried out to establish an eloquent grouping of studied populations (Neel and Schull, 1954). MS Excel (Microsoft, Redmond, WA, USA) and GraphPad Prism (ver. 5) (San Diego, CA, USA) were used for graphical presentations.

Figure 1. Map of the Arabian Peninsula showing the populations of Saudi Arabia employed in the present study.

The degree of differentiation according to both blood group/polymorphic systems was estimated for all populations (Nei, 1987). Nei's genetic alignment distances (DA) were calculated using the allelic frequencies (Nei and Roychoudhury, 1982). The resulting matrices were displayed by UPGMA dendrogram (Sneath and Sokal, 1973; Ota, 1993).

RESULTS

Blood group proportions and allele frequencies

For the overall Saudi population, the blood type "O" was the most common for the ABO system (50.0%), followed by types "A" (26.70%), "B" (19.09%), and "AB" (4.21%) (Table 1). For the Rhesus system, the Rh+ blood group was found in 91.81% subjects while 8.19% subjects were Rh-. Among the regional samples, type "O" was highest in Abha (56.78%) and lowest in Domah (35%); type "A" was highest in the South region (33.68%) and lowest in Sakaka (23%); type "B" was highest in Domah (30%) and lowest in Abha (6.04%); and type "AB" was observed to be highest in Al-Qurayat and Sweer (9%) and lowest in the South region (2.75%) (Table 1, Figures 2 and 3). For the Rhesus system, Rh+ was found to be highest in Domah (94%) and lowest in Sakaka (86%).

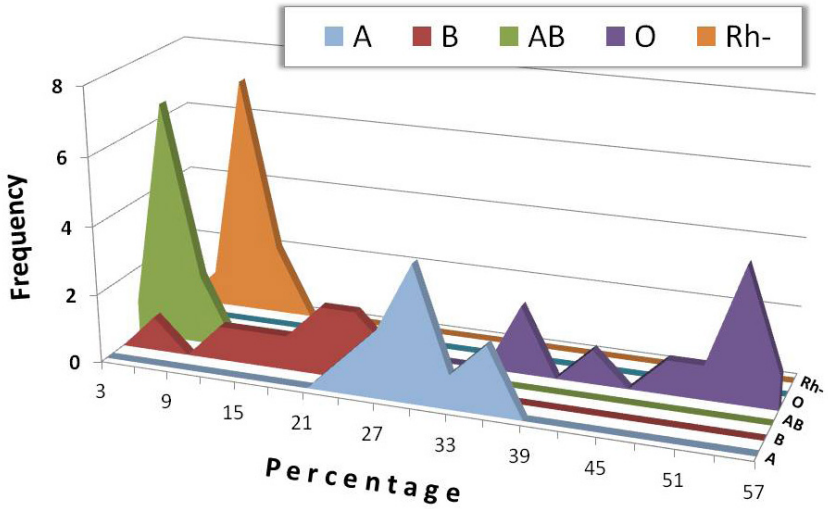

Figure 2. Three-dimensional area chart depicting the spread of blood group proportions in Saudi populations. Blood type "AB" clearly shows the least variability while types "B" and "O" depict high variability.

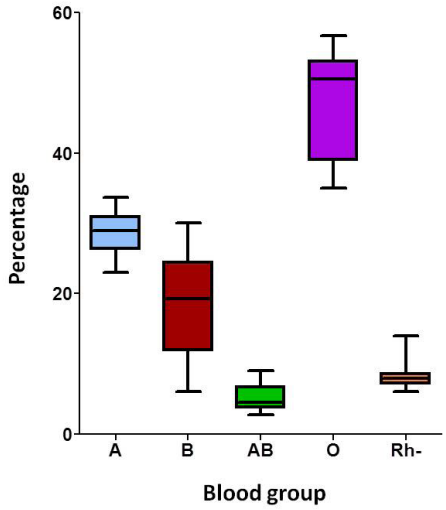

Figure 3. Box and whisker plots demonstrating the ranges of blood type proportions in Saudi populations.

Accordingly, the frequencies of p[A], q[B], and r[O] alleles at the *ABO* locus were observed to be 0.1688, 0.1242, and 0.7070, respectively (Table 2). The frequencies of the *D* and *d* alleles at the *Rh* locus were calculated to be 0.7138 and 0.2862, respectively. The distributions of allele frequencies at the *ABO* and *Rh* loci were found to be highly variable in the Saudi populations. For instance, the A[p] allele ranged from 0.1513-0.2119, B[q] between 0.0502-0.2015, and O[r] between 0.5949-0.7435 (Table 2 and Figure 4). At the *Rh* locus, the

D allele revealed highest estimates in Domah (0.7551), and lowest in Sakaka (0.6258). The samples were checked for conformity with HWE assumptions at the *ABO* locus. These analyses established that most of the samples of Saudi populations were in conformity with HWE, with the exception of Abha, which exhibited significant deviation ($\chi^2 = 18.32$; $P < 0.001$). The total sample of Saudi population was consistent with HWE (Table 2).

Table 2. Distribution of allele frequencies at the *ABO* and *Rh* loci and Hardy-Weinberg equilibrium at the *ABO* locus.

Population	*ABO* locus			HWE test statistics; P value	*Rh* locus	
	p[A]	q[B]	r[O]		Rh+ (D)	Rh- (d)
Al-Khobar	0.1662	0.1197	0.7141	1.58	0.7189	0.2811
Riyadh	0.1634	0.1283	0.7082	0.41	0.7171	0.2829
Jeddah	0.1663	0.1027	0.7309	0.08	0.7171	0.2829
South region	0.2028	0.0732	0.7240	0.06	0.6789	0.3211
Abha	0.2062	0.0502	0.7435	18.32*	0.7316	0.2684
Tabuk/Madina Al-Munawaara	0.1925	0.0878	0.7197	1.38	0.7202	0.2798
Sakaka	0.1513	0.1573	0.6914	0.02	0.6258	0.3742
Domah	0.1952	0.2015	0.6034	0.69	0.7551	0.2449
Al-Qurayat	0.2111	0.1681	0.6208	0.78	0.7172	0.2828
Sweer	0.2119	0.1932	0.5949	0.13	0.7354	0.2646
Total[#]	0.1688 ± 0.0010	0.1242 ± 0.0009	0.7070 ± 0.0012	0.07	0.7138 ± 0.0017	0.2862 ± 0.0017

*Significantly deviating from Hardy-Weinberg equilibrium (HWE) expectations. [#]Means ± SD.

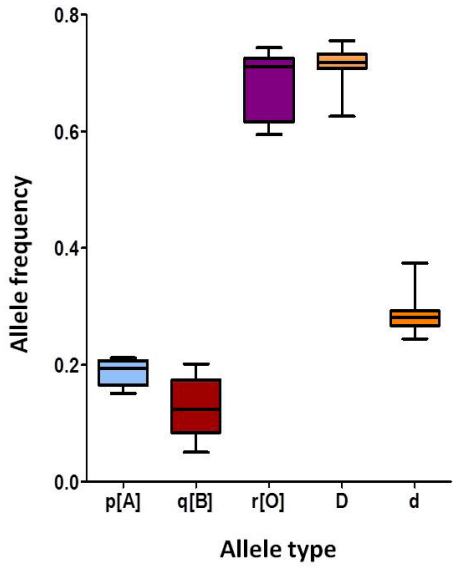

Figure 4. Box and whisker plots depicting the ranges of allele frequencies at the *ABO* and *Rh* loci in Saudi populations.

Locus heterozygosity

Heterozygosities at the individual *ABO* and *Rh* loci and combined blood group loci were established in the Saudi populations. The heterozygosities at the *ABO* and *Rh* loci were

0.4563 ± 0.0013 and 0.4086 ± 0.0010, respectively, and the combined heterozygosity was calculated to be 0.4324 ± 0.0238 (Table 3). At the *ABO* locus, heterozygosity was observed to be highest in Sweer (0.5667) and lowest in Abha (0.4024) (Figure 5). At the *Rh* locus, the highest estimate was evident in Sakaka (0.4707) and lowest in Domah (0.3717) (Table 3). The differences between populations became less remarkable when combined heterozygosities were considered. The combined heterozygosity was highest in Sweer (0.4813 ± 0.0882) and lowest in Abha (0.3979 ± 0.0047) (Figure 5).

Table 3. Heterozygosities at the studied loci in Saudi populations.

Population	Heterozygosity		
	ABO	Rh	Average
Al-Khobar	0.4481	0.4042	0.4261 ± 0.0220
Riyadh	0.4554	0.4058	0.4306 ± 0.0248
Jeddah	0.4276	0.4058	0.4168 ± 0.0109
South region	0.4301	0.4367	0.4342 ± 0.0033
Abha	0.4024	0.3929	0.3979 ± 0.0047
Tabuk/Madina Al-Munawaara	0.4386	0.4042	0.4227 ± 0.0172
Sakaka	0.4767	0.4707	0.4761 ± 0.0030
Domah	0.5600	0.3717	0.4682 ± 0.0946
Al-Qurayat	0.5445	0.4077	0.4785 ± 0.0688
Sweer	0.5667	0.3911	0.4813 ± 0.0882
Total[#]	0.4563 ± 0.0013	0.4086 ± 0.0010	0.4324 ± 0.0238

[#]Means ± SD.

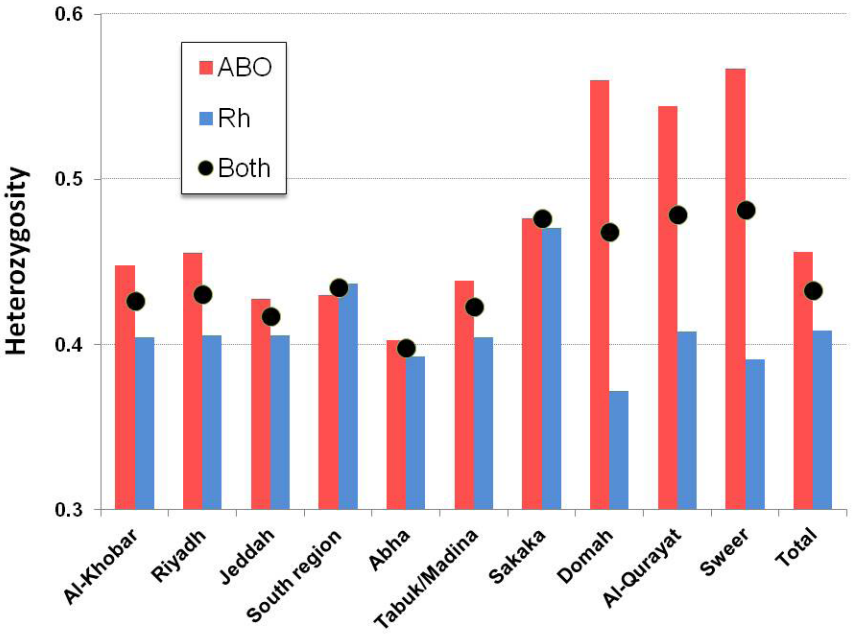

Figure 5. Comparison of individual and combined heterozygosities at the *ABO* and *Rh* loci in the studied populations of Saudi Arabia.

Variance at the phenotypic and allelic levels

In order to establish variability at the studied attributes, descriptive summaries were generated for the phenotypic proportions and allelic frequencies (Figures 2-4). The CoV was calculated to be the highest for the "AB" blood type (42.22%) followed by type "B" (39.60%), while CoVs were 16.27 and 11.44% for the "O" and "A" blood types, respectively. At the *ABO* allelic system, the q[B] allele depicted the highest CoV (39.86%), followed by p[A] and r[O] alleles (12.12 and 8.23%, respectively) (Figure 5). At the *Rh* locus, the CoV at the *d* allele was 12.4% and at *D* it was 5.02%.

Heterogeneity among samples

Homogeneity was tested among the ten Saudi populations by employing the allelic frequencies at the *ABO* locus and blood type proportions of the Rh system (Neel and Schull, 1954). Pairwise analysis for the ABO system demonstrated that the Abha population was the most heterogeneous, while Sakaka exhibited the most homogeneity. However, for the Rh system, Sakaka was the most heterogeneous sample. Z-tests were conducted through the phenotypic proportions and aggregated scores were used to compare the samples. The sample from Abha appeared most heterogeneous followed by Sweer and Al-Qurayat. Furthermore, Tabuk/Madina Al-Munawaara samples exhibited the least differences from the other populations. For the phenotypic systems, differences were most obvious for the blood type "B", followed by types "O", "AB", and "A".

Gene diversity analysis

To further establish gene differentiation among the Saudi populations the concept of gene diversity was explored at the *ABO* and *Rh* loci. Expected heterozygosities were estimated for the total population (H_T) and within the subpopulations (H_S). Absolute gene diversity (D_{ST}) was calculated from the expected heterozygosities (Nei, 1987).

For the sake of convenience, Saudi population samples were pooled into two groups: Southern and Northern. The Southern group comprised five populations, Al-Khobar, Riyadh, Jeddah, South region, and Abha *ABO*. The Northern region comprised Tabuk/Madina Al-Munawaara, Sakaka, Domah, Al-Qurayat, and Sweer (Table 4). Homogeneity was higher in the Northern group compared to the Southern, which appeared subdivided and stratified. This was evident from the behavior of "coefficient of inter-population gene differentiation" (G_{ST}) (Table 4). The D_{ST} at the pooled loci was almost four times higher in the Southern group compared to the Northern. However, it was observed that H_T was comparable with H_S at both *ABO* and *Rh* loci, which indicated only minor contributions of the inter-population components of genetic differentiation. The total genetic diversity (H_T) of the *ABO* locus was greater than that of the *Rh* locus in both groups, as well as in the total pool. Taken together, across all populations D_{ST} was 0.0040 and G_{ST} was 0.0090 (Table 4).

Genetic distance matrix through Nei's genetic distance (DA) measures

Nei's genetic distances (DA) were calculated for the recruited Saudi populations (Table 5). There were high affinities between samples obtained from Al-Khobar and Riyadh,

between Al-Khobar and Jeddah (DA = 0.0002), between Riyadh and Jeddah (DA = 0.0004), between Jeddah and Tabuk/Madina Al-Munawara (DA = 0.0004), and between Domah and Sweer (DA = 0.0005) (Table 5). Highest levels of heterogeneity (i.e., low affinities) were established between Sakaka and Domah (DA = 0.0214), between South region and Domah (DA = 0.0191), and between Abha and Domah (DA = 0.0177). Generally, the Domah population demonstrated greater differences with other populations.

Table 4. Gene diversity analysis for *ABO* and *Rh* loci in Saudi populations.

Population	Locus	H_T	H_S	D_{ST}	G_{ST}
Northern group	ABO	0.4339	0.4325	0.0014	0.0032
	Rh	0.4095	0.4089	0.0006	0.0015
	Pooled	0.4217	0.4207	0.0010	0.0024
Southern group	ABO	0.5195	0.5149	0.0046	0.0089
	Rh	0.4112	0.4072	0.0040	0.0097
	Pooled	0.4653	0.4611	0.0043	0.0092
All (10)	ABO	0.4794	0.4737	0.0057	0.0118
	Rh	0.4103	0.4080	0.0023	0.0056
	Pooled	0.4449	0.4409	0.0040	0.0090

Table 5. Genetic distance matrix showing the affinities between Saudi populations.

	Al-Khobar	Riyadh	Jeddah	South region	Abha	Tabuk/Madina Al-Munawaara	Sakaka	Domah	Al-Qurayat
Riyadh	0.0000								
Jeddah	0.0002	0.0004							
South region	0.0030	0.0034	0.0022						
Abha	0.0029	0.0037	0.0019	0.0023					
Tabuk/Madina Al-Munawaara	0.0007	0.0011	0.0004	0.0016	0.0008				
Sakaka	0.0080	0.0075	0.0085	0.0069	0.0155	0.0103			
Domah	0.0096	0.0087	0.0125	0.0191	0.0177	0.0123	0.0214		
Al-Qurayat	0.0050	0.0045	0.0070	0.0097	0.0108	0.0064	0.0121	0.0021	
Sweer	0.0091	0.0083	0.0119	0.0165	0.0166	0.0112	0.0184	0.0005	0.0009

Dendrogram analyses

Based on the standard genetic distances matrix, dendrograms were constructed by employing the unweighted pair-group method using arithmetic averages (UGPMA). An outlier with equal allele frequencies at both loci was added in the analyses. These analyses further established the close affinities between the samples obtained from the populations of Al-Khobar and Riyadh, between South and Tabuk/Madina Al-Munawaara, and between Domah and Sweer (Figure 6). The sample of Jeddah joined the cluster of Al-Khobar and Riyadh, Abha joined the cluster of South and Tabuk, and Al-Qurayat joined the cluster of Domah and Sweer. Interestingly, among the northern four populations (i.e., Sakaka, Domah, Al-Qurayat, and Sweer), Sakaka emerged distinct from the others. In contrast, Sakaka joined the clusters of the middle and Southern populations.

Principal component analyses were performed to further elucidate the sample affinities (Figure 7). These analyses iterated the clustering of the Northern populations and the Eastern/Western samples. The distinct nature of Sakaka was also evident.

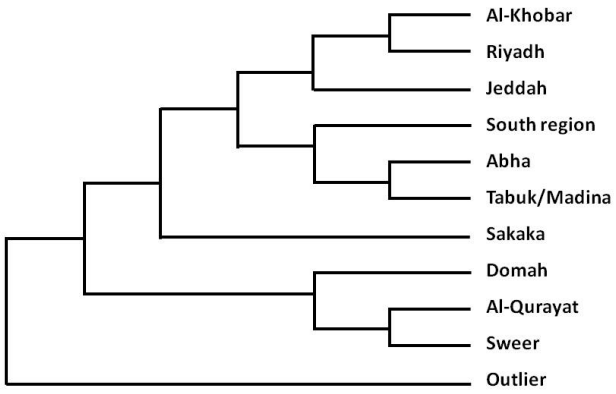

Figure 6. Dendrogram based upon DA-UPGMA showing the genetic relationships between Saudi populations. DA-UPGMA = alignment distance-unweighted pair-group method using arithmetic averages.

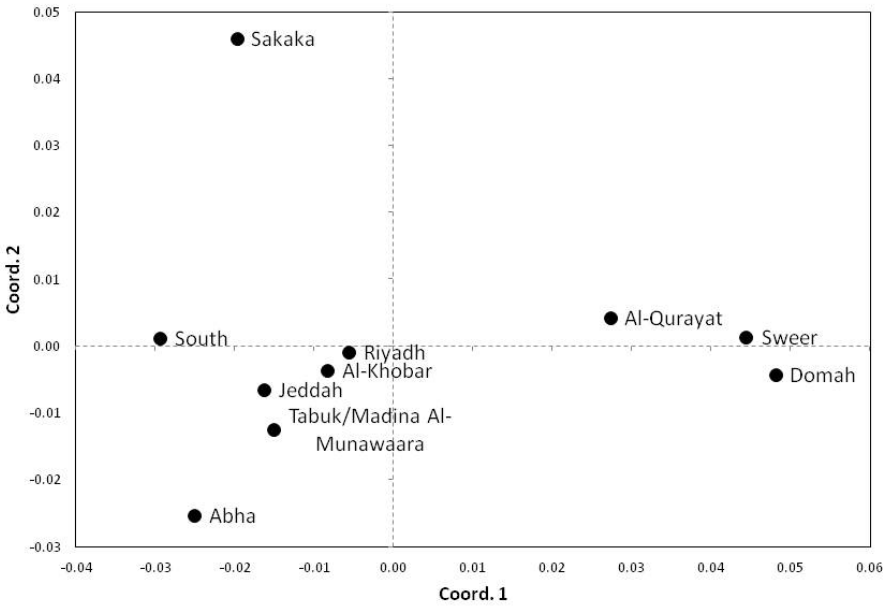

Figure 7. Scatter plot exhibiting the results of principal component analyses.

Correlation between the allelic systems

In order to observe any potential association between the alleles, Spearman's correlations were calculated. The initial hypothesis was that there was no correlation(s) between the alleles at one locus or alleles within the loci and all alleles existed independently. Alternatively, there were correlations. Interestingly, significant negative correlations were observed between the p[A] and r[O] alleles, and between the q[B] and r[O] alleles at the *ABO* locus (Table 6). The most remarkable correlation was observed between the q[B] and r[O] alleles (Pearson r^2 =

-0.9160; 95%CI = -0.9803 to -0.6765; P = 0.0002; highly significant). Furthermore, the p[A] allele also showed association with the *D* (and *d*) allele at the *Rh* locus.

Table 6. Correlation matrix between the allelic systems.

	ABO locus			*Rh* locus	
	p[A]	q[B]	r[O]	*D*	*d*
p[A]	1				
q[B]	0.02	1			
r[O]	-0.42	-0.92	1		
D	0.50	0.13	-0.32	1	
d	-0.50	-0.13	0.32	-1	1

Clinal trends of phenotypic and allele frequencies

In order to observe any clinal trends for the studied parameters, the phenotypic and allelic frequency data were mapped on the geography of Saudi Arabia (Figure 8). There were remarkable geographic trends at blood types "B" and "O" and the q[B] and r[O] alleles. Blood group "O" and the r[O] allele showed an increasing trend from North-East towards South-West. Conversely, blood type "B" and the q[B] allele exhibited a decreasing trend from North-East towards South-West (Figure 8).

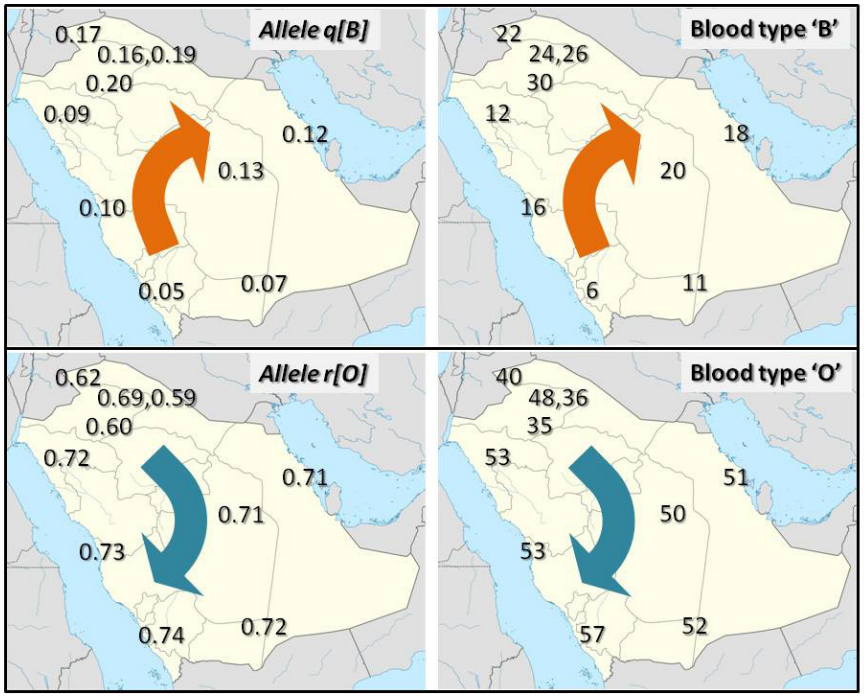

Figure 8. Clinal analyses. The upper panel shows the clinal trends for allele q[B] and blood type "B", while the lower panel depicts the trends for allele r[O] and blood type "O".

DISCUSSION

The study of distribution of blood groups is useful for blood transfusion, organ transplantation, genetic research, human evolution investigations, and forensic pathology. Blood groups show associations between certain diseases and Rh and ABO incompatibilities of newborns (Klein and Anstee, 2005). Data for *ABO* and *Rh* loci in Saudi populations are limited. This study attempted to draw a comprehensive picture of these polymorphisms in the Saudi populations by analyzing various indices of gene diversity. Of the ten studied populations, only the sample from Abha depicted deviations from HWE at the *ABO* locus. The inspection of the phenotypic proportions showed that blood type "O" was overrepresented while type "B" was strongly underrepresented in the Abha population. These variables resulted in an increase in the frequency of the r[O] allele and a decrease in the frequency of the q[B] allele (Tables 1 and 2). Deviations from HWE could result from non-random sampling and high inbreeding. It is interesting to mention that Abha had the highest loss of heterozygosity at the studied loci and a remarkable homozygosity was observed in this sample. The specific reasons for this high homozygosity remain to be discovered. It is quite likely that there is wide-spread consanguinity in this population. There are sporadic reports of high levels of consanguinity in the Saudi populations but a country-wide pattern of inbreeding coefficients has not been elucidated (el-Hazmi et al., 1995). Furthermore, compared to the Pakistani populations (which are also highly inbred), heterozygosity appeared to be very low in the Saudi populations (Malik and Amin-ud-Din, 2013; Ali and Malik, 2014; Rehman et al., 2014). It would be interesting to investigate the relationship of consanguinity and the reduction of heterozygosity in the Saudi as well as Pakistani populations (Hina and Malik, 2014; Jabeen and Malik, 2014).

The phenotypic and allelic frequencies of the populations of Al-Khobar, Riyadh, and Jeddah were observed to be in close agreement with one another. It appeared that these populations of the Eastern, middle, and Western regions of Saudi Arabia had mixed populations due to a cosmopolitan effect. They comprise the main urban assemblage of masses and therefore contain people of mixed ethnicities. On the other hand, the populations in the North and South are less admixed and more stratified. Interestingly, the stratification and sub-structuring in the Saudi samples appear to be less remarkable than the estimates available for the Pakistani populations of the Punjab Province (Shami and Rasmuson, 1994; Malik and Amin-ud-Din, 2013; Ali and Malik, 2014).

Homogeneity analyses allowed the identification of sub-clusters of Saudi populations with similar characteristics. These population sub-clusters were further iterated by Nei's measure of genetic distance and phylogenetic analyses. It is quite likely that the affinities exhibited by the UPGMA tree are real. Alternatively, these similarities might have emerged due to close geographic proximities of the studied populations, common evolutionary mechanisms, admixture, or other stochastic factors. These analyses further revealed correlations between the allelic systems q[B] and r[O]. There appeared to be a geographic trend operational at both alleles, which was not obvious at allele p[A]. It could be anticipated that both of the aforementioned alleles might have some selection pressure resulting in their characteristic distribution. It is safe to conclude that the *ABO* locus is under certain evolutionary constraints. Different factors like migrations, infectious diseases and immunological exposures, disease susceptibility, non-random mating and inbreeding, and genetic drift are factors that confer evolutionary constraints on human populations. It remains however unexplored which of these (or even other) factors have shaped the current allelic landscape of the *ABO* and *Rh* genes in the Saudi population.

The findings of this study need to be affirmed through highly polymorphic markers such as microsatellites or by high-throughput methods. In a prospective study, it would be pertinent to include a dense grid of samples from all major small and large populations, and the samples should comprise both male and female individuals. It would be very interesting to observe stratification in different Saudi populations and to estimate the loss of heterozygosity, and to explore stratification at other loci. Furthermore, it would be worthwhile to check the homozygosity existing in Saudi population at the genomic level, particularly in the context of a high inbreeding coefficient.

The present study has some limitations. For example, there was a higher representation of male subjects in the data. Due to the traditional nature of Saudi society, the women tended to stay at home, and hence, their data are under-represented. Furthermore, the blood transfusion record could be biased towards more healthy males of a younger age while the data of children and older subjects might be less represented. The positive aspect of this study is the large sample size and the assemblage of large number of representative populations from all geographic borders of Saudi Arabia.

This is the first comprehensive study documenting the distribution of ABO and Rh blood group types in ten cities of Saudi Arabia. It provided a country-wide picture of the studied polymorphic markers and an initial glimpse into the allelic diversity of the studied loci.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors would like to extend their sincere appreciation to the DeanShip of Scientific Research at King Saud University for funding this research group (#IRG14-05).

REFERENCES

Abu-Amero KK, Jaeger M, Plantinga T, Netea MG, et al. (2013). Genetic variation of *TLR2* and *TLR4* among the Saudi Arabian population: insight into the evolutionary dynamics of the Arabian Peninsula. *Genet. Test. Mol. Biomarkers* 17: 169-169.

Al-Himaidi AR and Umar U (2002). ABO blood group distribution among Saudi citizens related to their regional or original tribal location. *Kuwait J. Sci. Eng.* 29: 75-81.

Alanazi M, Pathan AA, Ajaj SA, Khan W, et al. (2013). DNA repair genes *XRCC1*, *XRCC3*, *XPD*, and *OGG1* polymorphisms among the central region population of Saudi Arabia. *Biol. Res*. 46: 161-167.

Albaz N (2001). Distribution of some blood group antigens, Western Region, Saudi Arabia [Abstract]. Update in Hematology and Transfusion Medicine Symposium. January 23-24, 2001, Jeddah.

Ali S and Malik S (2014). Phenotypic distribution, allelic diversity and degree of differentiation at ABO and Rh loci in the population of Haripur District, Khyber Pakhtunkhwa, Pakistan. *Pakistan J. Zool.* 46: 1-7.

Bashwari LA, Al-Mulhim AA, Ahmad MS and Ahmed MA (2001). Frequency of ABO blood groups in the Eastern region of Saudi Arabia. *Saudi Med. J*. 22: 1008-1012.

el-Hazmi MA, al-Swailem AR, Warsy AS, al-Swailem AM, et al. (1995). Consanguinity among the Saudi Arabian population. *J. Med. Genet.* 32: 623-626.

Eweidah MH, Rahiman S, Ali H, Mohammed A, et al. (2011). Distribution of ABO and Rhesus (RHD) blood groups in Al-Jouf province of the Saudi Arabia. *Anthropologist* 13: 99-102.

Gerstman BB (2008). Basic Biostatistics: Statistics for Public Health Practice. Jones and Bartlett Publishers, Sudbury, 380-382.

Hina S and Malik S (2014). Pattern of consanguinity and inbreeding coefficient in Sargodha district, Punjab, Pakistan. *J. Biosoc. Sci*. 9: 1-9.

Jabeen N and Malik S (2014). Consanguinity and its sociodemographic differentials in Bhimber District, Azad Jammu and Kashmir, Pakistan. *J. Health Popul. Nutr.* 32: 301-313.

Klein HG and Anstee DJ (2005). Mollison's Blood Transfusion in Clinical Medicine. 11th edn. Blackwell Scientific Publications, Oxford, 114-163.

Malik S and Amin-ud-Din M (2013). Genetic heterogeneity and gene diversity at ABO and Rh loci in the human population of southern Punjab, Pakistan. *Pakistan J. Zool.* 45: 1185-1190.

Mather WB (1964). Principles of Quantitative Genetics. Burgess Publishing Co., Minneapolis.

Mustafa MHI, Elmisbah TE, Salim AM, Ahmed MAM, et al. (2014). *RHD* and *RHCE* frequencies and gene complexes among major tribes of Turabah Province, Saudia Arabia. *Int. J. Multidiscip. Curr. Res*. 2: 573-578.

Neel JV and Schull WJ (1954). Human Heredity. University of Chicago Press, Chicago.

Nei M (1987). Molecular Evolutionary Genetics. Columbia University Press, New York.

Nei M and Roychoudhury AK (1982). Genetic relationship and evolution of human races. *Evol. Biol.* 14: 1-59.

Ozsoylu S and Alhejaily M (1987). The distribution of ABO and Rh blood groups in the Tabuk regions and Madinah Munnawwara, Saudi Arabia. *Turk. J. Pediatr*. 29: 239-241.

Ota T (1993). DISPAN: Genetic distance and phylogenetic analysis. Pennsylvania State University Park. Available at [ftp://ftp.bio.indiana.edu/molbio/ibmpc/dispan.zip].

Rehman AU, Wahab ZU, Khattak MNK and Malik S (2014). ABO and Rh (D) blood groups polymorphism in four Tehsils of Bajaur Agency (Federally Administered Tribal Areas), Pakistan. *Anthropologist* 18: 259-261.

Sarhan MA, Saleh KA and Bin-Dajem SM (2009). Distribution of ABO blood groups and Rhesus factor in Southwest Saudi Arabia. *Saudi Med. J*. 30: 116-119.

Shami SA and Rasmuson M (1994). Genetic heterogeneity and gene diversity in the population of Punjab, Pakistan, based on ABO and Rh(D) blood group frequencies. *Hum. Hered.* 44: 214-219.

Silva PJN (2002). ABO estimator. Faculdade de Ciencias da Universidade de Lisboa, Lisboa, Portugal.

Sinha S, Amjad M, Rogers C, Hamby JE, et al. (1999). Typing of eight short tandem repeat (STR) loci in a Saudi Arabian population. *Forensic Sci. Int.* 104: 143-146.

Sneath PHA and Sokal RR (1973). Numerical Taxonomy: the principles and practice of numerical classification. Freeman, San Francisco.