



OPEN

HLA alleles and haplotypes in Sudanese population and their relationship with Mediterraneans

Fabio Suarez-Trujillo^{1,2,4}, Sayda El-Safi^{3,4}, Ignacio Juarez^{1,2}, José Palacio-Gruber^{1,2}, Alejandro Sanchez-Orta¹, José Manuel Martin-Villa^{1,2} & Antonio Arnaiz-Villena^{1,2}✉

The contribution of migrated people from once green Sahara (about 10,000–6000 years BC) towards Mediterranean area had probably a double effect: both genetic and cultural connections have been described between Western Europe and North Africa. Sudanese populations from different ethnicities have been studied for HLA-A, -B, -DRB1 and -DQB1 antigens by a standard microlymphotoxicity method. Results found show that Nubians are genetically related with African Sub-Saharan populations and distant from other Sudanese tribes, who are closer to Mediterranean populations than to Sub-Saharan ones. This is concordant with other authors and meta-analysis data. Our present work is, to our knowledge, the first and only one HLA research that studies Sudanese people according to different Sudan ethnic groups: samples were collected before Sudan partition between North and South. A prehistoric genetic and peoples exchange between Africa and the Mediterranean basin may be observed and is supported with the results obtained in this Sudanese HLA study. However, demic diffusion model of agriculture and other anthropological traits from Middle East to West Europe/Maghreb do not exist: a more detailed Sahel and North African countries ancient and recent admixture studies are also being carried out which may clearer explain pastoralists/agriculture innovations origins in Eurafrikan Mediterranean and Atlantic façade.

Sudan means “land of black people” in Arabic language. It became independent from United Kingdom in 1956. It has been involved in several civil wars mainly with North (Muslim) part of Sudan against South (Christian). Two different countries arose in 2011: Sudan at North, mostly of Muslim religion and capital at Khartoum, and South Sudan, mostly of Christian and ancestral autochthonous religions and capital at Juba^{1,2}. Samples studied in this work are from 13 different ethnicities of Sudan and were collected from Khartoum before Sudan split that was and still is being a big immigration city; it is placed at the convergence of Blue and White Nile River. Main local tribes or ethnic groups which participated in this study are defined by languages and depicted in Fig. 1 (for more details of Sudan ethnic groups see³ and www.joshuaproject.net). Also, South Sudan was somewhat neglected and stayed backwards between the Anglo (UK)-Egyptian mandate during 1899 and 1956. UK and Egypt centred benefits in North part of Sudan, particularly in Khartoum area [1,2]. Ethnicities admixture makes it difficult to assess individual genetic ancestry, although it was established mainly by language and place of origin. The term “Nubian” was used at the time of this study sample collection for African people with some Ethiopian phenotype characters close to Ethiopians or not, but coming from Nuba Mountains, which have blurred limits with South Sudan (Fig. 1).

Sahara desertification became more intense by 10–6000 years ago. Saharan people underwent then a strong migratory pressure and started to emigrate. In the context of a drying Sahara the desert settled populations there began to disperse in all directions, including North Africa and the Mediterranean basin. This gave rise to a mixture of people and genes and influenced the Ancient Classic Mediterranean cultures creating a cultural circle that includes Mediterranean populations and Guanches or Canary Islands First Inhabitants^{4,5}. The findings of the same types of Lineal Megalithic or more recent scripts and megalithic structures in the Sahara and throughout the Mediterranean Basin and the Eurafrikan Atlantic coast, including Canary Islands, support a common Ancient Megalithic culture that probably was born in the Humid Sahara period before 6000 BC⁵.

In the present study we aim: (1) to compare a mixed sample of different Sudanese ethnicities with other surrounding populations; (2) to clarify their origin according to HLA genetic background and to ascertain

¹Department of Immunology, School of Medicine, University Complutense, Pabellón 5, Planta 4. Avda. Complutense S/N, 28040 Madrid, Spain. ²Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain. ³Ibn Sina Specialized Hospital, Khartoum, Sudan. ⁴These authors contributed equally: Fabio Suarez-Trujillo and Sayda El-Safi. ✉email: arnaizantonio@gmail.com



Figure 1. Sudan and South Sudan map (split in 2011). Sudanese tribes included in this study are grey-scale remarked in the figure. Rufaa and Kawahla tribes (not shown in figure) also participated in this study and are widespread all across Central, North and South parts of Sudan. Number of individuals of each ethnicity included in present work were: Magyarab ($n=8$), Mahas ($n=7$), Danagla ($n=7$), Saigiya ($n=9$), Bedaria ($n=8$), Kenana ($n=10$), Gaahmg ($n=6$), Shukria ($n=8$), Gummuiya ($n=7$), Gaaliin ($n=8$), Rufaa ($n=11$), and Kawahla ($n=9$). Nubians ($n=101$) HLA genetic data were taken from 12th Histocompatibility International Workshop. For more information see³ and www.joshuaproject.net. This figure has been made by our own work based on a map (<https://www.rawpixel.com/image/8905866/sudan>) and performed with Adobe Illustrator 2020 software (24.3 version; <https://www.adobe.com/es/products/illustrator.html>).

their relationships with Mediterranean and Sub-Saharan populations. HLA genes are the transplantation genes in humans and, for strictly biomedical purposes, they are useful compared to other genes for studying human populations and ethnic groups origins because of the HLA very high polymorphism and correlation of HLA allele frequencies follows a geographical gradient⁴; (3) to ascertain genetic differences found between Sudanese Nubians from Nuba Mountains and Sudanese from other Sudan ethnicities (Fig. 1).

Material and methods

Populations samples

Ninety-eight healthy unrelated individuals from all Sudan territory have been included in this study belonging to different ethnicities (depicted in Fig. 1): Magyarab ($n=8$), Mahas ($n=7$), Danagla ($n=7$), Saigiya ($n=9$), Bedaria ($n=8$), Kenana ($n=10$), Gaahmg ($n=6$), Shukria ($n=8$), Gummuiya ($n=7$), Gaaliin ($n=8$), Rufaa ($n=11$), and Kawahla ($n=9$). Nubians ($n=101$) HLA genetic data were taken from 12th Histocompatibility International Workshop⁶. Blood samples were collected at Ibn Sinna Hospital, in Khartoum State, and used for HLA typing and

phylogenetic calculations. Other world populations were considered for genetic comparisons and are detailed in Table 1. The study was approved by Complutense University Ethical Committee (EC-0117.AK) and all subjects included in the study were adult unrelated blood donors who signed an informed consent to participate in it.

HLA typing, DNA sequencing and statistics

Generic HLA class I (A and B) and class II (DRB1 and DQB1) typing was done by two-step microlymphotoxicity technique on T or B lymphocytes¹⁹ by using the 10th and 11th International Histocompatibility Workshop and local reagents. This technique was used because it was necessary to study all Sudan available ethnicities and HLA low resolution was also available and it is sufficient to accurately distinguish or relate populations^{4,11}.

Two types of analyses were done to compare HLA frequencies of Sudan with those of Mediterranean and other populations: (1) with low-resolution DRB1 and DQB1 data; and (2) with low-resolution DRB1 data. This latter analysis was performed in order to compare Nubians from Sudan and other Sudanese ethnicities in a proper way by using the same level of HLA allelism distinction, who have been typed/transformed for HLA-DRB1 generic but lack HLA-DQB1 typing.

Statistical analysis was performed with Arlequin v3.0 software²⁰. In summary, this program calculates the linkage disequilibrium between two alleles at two different loci. Their level of significance (p) for 2×2 comparisons and also their relative linkage disequilibrium were calculated as previously described²¹. Computer program calculated the most frequent HLA extended haplotypes that were further analysed as follows; the most frequent complete extended haplotypes were tentatively deduced from: (1) the 4 HLA loci haplotype frequencies were obtained by software^{21,22}; (2) the previously described haplotypes in other populations were also taken into account²¹; and (3) haplotypes if they appeared in two or more individuals and the alternative haplotype was well defined were also studied^{21,22}. In order to compare phenotype and haplotype HLA frequencies with other populations, the reference tables of the 11th and 12th International HLA Workshop were used^{6,12} (Table 1). Phylogenetic trees (dendrograms) were constructed with HLA-DRB1 frequencies by using the Neighbour-Joining (NJ) method²³ and genetic distances between populations (DA)²⁴, by using DISPAN software which contains the programs GNKDST and TREEVIEW^{25,26}. Also, correspondence analysis in three dimensions and

Population	N	Reference
Sudanese	98	Present study
Lebanese—KZ ¹	93	6
Moroccans	98	7
Berbers (Souss)	98	8
Moroccan Jews	94	9
Portuguese	236	10
Spaniards	176	11
Spanish Basques	80	11
French	179	12
Algerians	102	13
Sardinians	91	12
Italians	284	12
Ashkenazi Jews	80	14
Non-Ashkenazi Jews	80	14
Cretans	135	15
Macedonians	172	16
Attika Greeks	85	6
Cypriot Greeks	101	6
Aegean Greeks	98	6
Romanians	99	6
Armenians	97	6
Iranians	100	17
Turks	89	18
Egyptians (Siwa)	101	6
Nubians	101	6
Oromo	83	6
Amhara	98	6
Fulani	38	6
Rimaibe	39	6
Mossi	42	6

Table 1. Populations used for the present work. *N* Number of individuals analysed for each population; ¹KZ Kafar Zubian, a Shia Muslim village.

its bidimensional representation was carried out using Vista v5.05 computer program²⁷. Correspondence analysis consists of a geometric technique that may be used for displaying a global view of the relationships among populations according to HLA (or other) haplotypic (or allelic) frequencies. This methodology is based on the genetic distances (DA) variance among populations (similar to the classical principal components methodology) and also a statistical visualization of these differences.

Ethical approval

This study was conducted according to the Declaration of Helsinki by the Ethics and has been approved by Complutense University Ethical Committee (EC-0117.AK).

Consent to participate

All subjects included in the study were adult unrelated blood donors who signed an informed consent to participate in it.

Results

Characteristic Sudanese HLA allele frequencies

The expected and observed genotype frequency values for HLA-A, -B, -DRB1 and -DQB1 loci do not significantly differ, and the population sample is in Hardy–Weinberg equilibrium. Table 2 shows the HLA allele frequencies found in our Sudanese analysed sample.

Figure 2 depicts a low-resolution class II (DRB1) neighbour-joining (NJ) tree constructed from genetic distances between populations (DA, Table 3). Its topology shows how compared populations cluster in two main branches: in general, western (both North African and Europeans) and other Mediterraneans group together respectively and tend to converge in the same node; in the other branch Greeks and Sub-Saharanans tend to cluster together with Nubians, Fulani, Rimaibe and Mossi (as described before in^{5,16,28} and with other genetic markers^{29,30}). Our Sudanese studied group is placed especially close to most western Mediterraneans: North African (Algerians and Moroccans) and Iberians (Table 3). Genetic distance values (Table 3) give quantitative genetic distances and relatedness between populations and show that Turks are the closest to Sudanese population, followed by Egyptian-Bedouins and Iranians. Also, Mediterranean populations like Algerians, Moroccans, Spaniards and Italians are genetically close to Sudanese population studied in present work. Finally, Greeks and Sub-Saharanans cluster together and behave as outgroups. In the same way, the correspondence analysis performed show again how Sudanese are placed within the eastern Mediterranean group, which is also related to the western one (Fig. 3). On the other hand, Nubians cluster together with Greeks and other black Sub-Saharan populations included in this work. Genetic distances between Nubians and our Sudanese mixed sample are similar to those among other Sub-Saharanans (Oromo and Amhara) and Sudanese (Tables 3 and 4). These results in Greeks have been previously confirmed by different independent research groups and genes (see “Discussion” section).

To complete the genetic relationship study of our studied Sudanese from different ethnicities with other reported populations, a second set of analyses were carried out by using low-resolution HLA-DRB1 data. The constructed NJ tree (Fig. 2) and calculated genetic distances (Tables 3 and 4) show a very similar relatedness of Sudanese population in both types of analyses. Interestingly, Nubians, have a genetic distance value with respect to our Sudanese mixed sample as far as Greeks and Sub-Saharanans (Table 3). In fact, when Correspondence analysis with generic HLA-DRB1 data is performed, Nubians are closer to sub-Saharanans than any other compared population including our Sudanese mixed-ethnicities studied group (Fig. 3 and Table 4). Table 4 also shows that Nubians are more related to Sub-Saharanans and also to Greeks, as it has been demonstrated by us and other researchers^{4,5,16}.

HLA-A, -B, -DR and -DQ linkage disequilibria

The present study reports not only HLA class I (A and B) and class II (DRB1 and DQB1) allele frequencies but also two-loci and extended haplotypes for the first time in a mixed sample of different Sudanese ethnicities, allowing their comparison with previously reported results in other populations.

The most frequent HLA extended four-loci haplotypes (HLA-A-B-DRB1-DQB1) (Table 5) are marked by the high frequency of some of the two-loci haplotypes described above. A*30 and A*31 alleles (low resolution) are particularly frequent in Mediterraneans and also in our Sudanese sample (Table 2); the same occurs with B*49 allele. The most frequent HLA extended haplotype found in our Sudanese sample is A*02-B*51-DRB1*13-DQB1*06 (4.59%) and is identified as a Mediterranean/European haplotype; it is also present in Spaniards (1.0%)¹². Also, this B-DRB1-DQB1 combination has an Iberian-Mediterranean background because is present with similar frequency in Portuguese, Italians, French, and African Americans from the US¹². The second most frequent haplotype found in present work (A*02-B*49-DRB1*03-DQB1*02, 2.56%) has been found to be specific of the Sudanese sample studied, although its partial B-DRB1-DQB1 haplotype can be found in Italians¹². Other Sudanese-specific HLA extended haplotypes with frequencies above 1% have been found in our Sudanese sample: A*02-B*51-DRB1*08-DQB1*03 (1.54%), A*03-B*52-DRB1*11-DQB1*03 (1.54%), and A*01-B*51-DRB1*13-DQB1*03 (1.02%). All other most frequent haplotypes are also shared with Mediterraneans (Table 5).

Discussion

Sudanese HLA Mediterranean background

Our present study shows that Sudanese population is related to Mediterraneans like Turks (who are genetically Mediterraneans except because of an Asian “elite” imposed language¹⁸), Egyptians, Iranians, Algerians, Italians and Iberians. This is concordant with other North African, Middle East and Iberian HLA compilation study³¹ by using another methodology (high-resolution HLA typing). On the other hand, Greeks were found initially close

Allele	%
HLA-A*	
A*02	32.65
A*30	12.24
A*31	10.2
A*03	9.18
A*68	7.15
A*24	6.63
A*01	5.63
A*25	4.08
A*23	3.57
A*32	3.57
A*11	2.04
A*33	1.53
A*29	0.51
A*34	0.51
A*69	0.51
HLA-B*	
B*51	21.43
B*15	8.67
B*49	8.67
B*52	8.17
B*41	6.12
B*57	6.12
B*14	5.63
B*13	3.57
B*38	3.57
B*45	3.57
B*07	3.57
B*44	3.06
B*18	2.04
B*35	2.04
B*50	2.04
B*08	2.04
B*37	1.53
B*42	1.53
B*53	1.53
B*58	1.53
B*55	1.02
B*67	1.02
B*27	0.51
B*39	0.51
B*73	0.51
HLA-DRB1*	
DRB1*13	25.51
DRB1*11	13.27
DRB1*15	13.27
DRB1*03	12.76
DRB1*07	8.67
DRB1*08	8.67
DRB1*04	6.63
DRB1*16	4.08
DRB1*14	3.06
DRB1*10	2.04
DRB1*01	1.53
DRB1*09	0.51
HLA-DQB1*	
Continued	

Allele	%
DQB1*06	41.84
DQB1*03	27.55
DQB1*02	20.92
DQB1*05	7.65
DQB1*04	2.04

Table 2. HLA-A, -B, -DRB1 and -DQB1 allele frequencies found in Sudanese ethnicities sample studied in present work (n = 98).

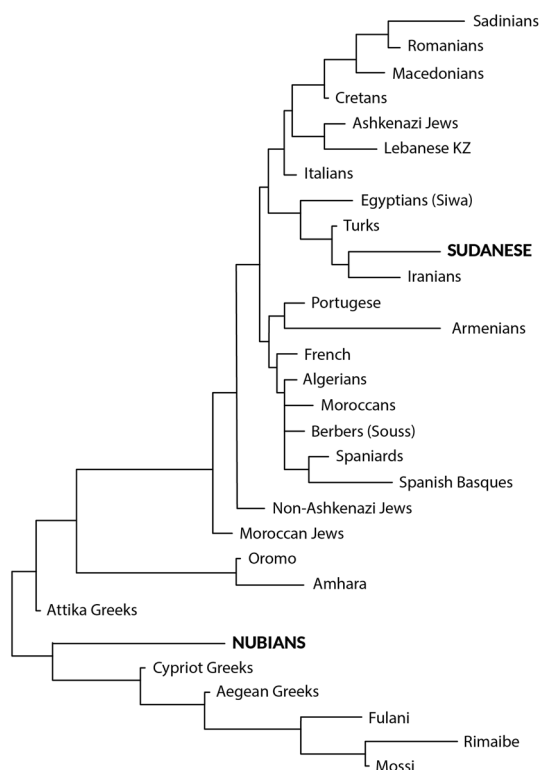


Figure 2. HLA-DRB1 dendrogram performed with Neighbour-Joining method showing genetic relatedness between Sudanese, Nubians, and other Mediterranean and Sub-Saharan populations. Genetic distances between populations were calculated with GNKDST software using low-resolution HLA-DRB1 frequencies. Bootstrap test showed values of 100 in all nodes after 1000 replicates. Note that Greek HLA relatedness with Sahel populations was detected by two independent different groups^{4,5,16,28} and it is also supported by other autosomal genetic markers for African and Greeks specifically (3120 + 1 G → A cystic fibrosis marker)^{29,30}. HLA-DRB1 allele frequencies is used because of the best discrimination between populations among HLA loci (more or less relatedness) and the fact that almost all the populations in data bases are HLA-DRB1 typed³¹.

to Sahel populations (Rimaibe, Mossi, Oromo, Amhara); it disconcerted us, and an allele-by-allele comparison was carried out between Greeks and Sub-Saharan to finally show that they both share some HLA alleles in common, confirming our phylogenetic results^{5,16}. Typing of these populations has been done by many participant laboratories in the 12th Histocompatibility International Workshop under the direction of Julia Bodmer (Oxford)^{5,6} and computer calculations were made by us. Also, another autosomal marker (3120 + 1 G → A cystic fibrosis marker) was found in Africans and also only in Greeks out of all other European populations tested^{29,30}. It was written by Aeschylus that African peoples (Danaids) emigrated northwards firstly to Egypt, where they were rejected and then fled to Peloponnesus³². This may give us a sort of historical reference for these genetic results that have been independently confirmed by other groups and other genetic markers^{5,16,28–30}.

North African migrations and contacts with the Mediterranean Basin

Sahara Desert climate has undergone sharp variations over time, ranging from wet to dry over the past hundreds of thousands of years³³. This variability is due to a 41,000-year cycle in which the axis of the Earth changes between 22° and 24.5°³³. In fact, the Columbia Shuttle detected a humid (traces of big rivers and lakes) and green

Population	DA ($\times 10^3$)
Turks	5.91
Egyptians-Siwa	5.93
Iranians	7.44
Algerians	8.42
Moroccans	8.64
Spaniards	8.75
Italians	9.52
Macedonians	10.38
Portugese	10.52
French	10.98
Berbers-Souss	11.6
Cretans	13.28
Romanians	13.30
Non-Ashkenazi Jews	13.41
Ashkenazi Jews	13.57
Moroccan Jews	14.02
Spanish Basques	15.16
Lebanese-KZ	19.71
Sadinians	19.74
Armenians	25.19
Attika Greeks	28.78
Oromo	30.93
Nubians	31.16
Amhara	36.89
Cypriot Greeks	42.71
Aegean Greeks	44.43
Fulani	48.40
Mossi	50.51
Rimaibe	54.79

Table 3. Genetic distances (DA) between studied Sudanese population and other world populations included in this work (HLA-DRB1).

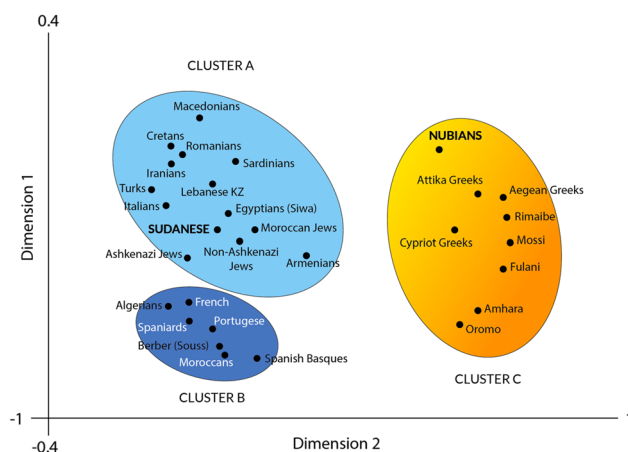


Figure 3. Correspondence analysis performed with HLA-DRB1 allele frequencies of Sudanese sample studied in present work together with those of other Mediterranean and African populations included (see Table 1). Three different clusters are obtained from the analysis: for simplicity, cluster A (light blue), cluster B (dark blue) and cluster C (yellow/orange). Mixed Sudanese population analysed in this study appears together with other Mediterraneans like Turks, Iranians, Egyptians, Italians or Jews in the eastern-Mediterraneans cluster (cluster A, light blue). However, Sub-Saharan Sahelians (Fulani, Rimaibe, Mossi) and Nubians cluster together with Greeks and other African Sub-Saharan populations (cluster C, yellow/orange), as previously obtained by two independent groups^{4,5,16,28}. Correspondence analysis supports the results previously obtained in the genetic distances analysis (DA) (Table 3) and NJ dendrogram (see Fig. 2). Italians have been chosen from North Italy⁶, a population that usually cluster with northern Balkan Peninsula or Central European groups.

Population	DA (× 10)
Greeks (Aegean)	14.84
Greeks (Cyprus)	16.13
Greeks (Attika)	16.61
Rimaibe	19.35
Mossi	23.81
Fulani	33.09
Amhara	20.01
Oromo	20.06
Moroccan Jews	20.15
Macedonians	20.44
Egyptians (Siwa)	20.52
Cretans	23.54
Italians	23.98
Non-Ashkenazi Jews	24.59
Turks	27.67
Portuguese	28.48
French	29.39
Moroccans	29.67
Algerians	30.47
Lebanese (KZ)	31.08
Sudanese	31.16
Romanians	32.52
Berbers (Souss)	33.09
Iranians	33.12
Armenians	34.27
Askenazi Jews	35.32
Sardinians	35.52
Spaniards	40.87
Spanish-Basques	46.44

Table 4. Genetic distances (DA) between studied Nubian population from 12th international histocompatibility workshop (n = 101, see⁶) and other world populations included in this work (HLA-DRB1), see Table 1.

Haplotype	%	Origin ²⁸
A*02-B*51-DRB1*13-DQB1*06	4.59	Mediterranean/European
A*02-B*49-DRB1*03-DQB1*02	2.56	Specific
A*02-B*51-DRB1*08-DQB1*03	1.54	Specific
A*03-B*15-DRB1*13-DQB1*06	1.54	European
A*03-B*52-DRB1*11-DQB1*03	1.54	Specific
A*02-B*50-DRB1*07-DQB1*02	1.02	Mediterranean/South Asian
A*02-B*51-DRB1*13-DQB1*03	1.02	Mediterranean/European
A*02-B*15-DRB1*13-DQB1*05	1.02	European
A*01-B*57-DRB1*15-DQB1*06	1.02	Mediterranean/South Asian
A*01-B*51-DRB1*13-DQB1*03	1.02	Specific

Table 5. Most frequent HLA-A-B-DRB1-DQB1 extended haplotypes in our Sudanese population (n = 98). Significant values are in bold.

Sahara Desert under the dunes³⁴. Currently it is in a dry period, but the Sahara is expected to turn green again in about 15,000 years. Due to these desertification-greening cycles, it is possible that very different human populations inhabited in the last humid stage. The Tassili N'Ajjer National Park and Ahaggar Mountains in Algeria show a record of rock art that proves human settlement in the Sahara³³. However, when the desertification began around 10,000 BC, the populations that inhabited the Sahara emigrated to other more habitable northern and other areas.

It is well established that North Africans and southern Europeans are genetically related, and this may be due to a long lasting circum-Mediterranean cultural and genetic flow particularly during the last glacial peak^{4,35–39}, being a secluded population between European North Ices and Desert. On the basis of our present day genetic and linguistic studies, we have postulated that many people coming from what is nowadays the Sahara Desert started to move towards East, West, North, and also South, being an important part of the primitive people stock of Sumerians, Egyptians, Guanches (Canary Islands), Iberians, Etruscans, Minoans, Anatolians (currently named Turks on only linguistic bases because they show a clear Mediterranean HLA profile)⁴, Kurds, and other islanders or northern Mediterraneans^{4,5,18,36–39}. In the present work, HLA genetic background of Sudanese people from different ethnicities is studied, and results obtained in part confirm these North African-Mediterranean peoples (genes) exchange.

Finally, there are common traits in Mediterranean, Iberian and North African cultures Circle, which may be conjointly named the Saharo-Canarian Circle⁴⁰. Traces of incised lineal writing have recently been found in human sites in the central Sahara (Ti-m Missaou, Algeria) in what would be the first discovery of “megalithic lineal writing” in continental Africa³⁵. This finding supports the theory that incised lineal writing was originated in fertile Sahara populations, which moved to the North when desertification began, carrying this type of lineal writing to Canary Islands (Guanche culture), the Mediterranean Basin and Europe. It is probable that these populations of the green Sahara moved towards the Mediterranean and the Canary Islands, originating common cultural and genetic traits among all the descendant populations (Guanches, Egyptians, Minoans, Etruscans, Iberians and Basques among others^{38,39}). The present study and other ones^{4,5,30,31,34}, support this theory of movements from central Africa to other regions; the appearance of Mediterranean alleles and semi-haplotypes (like HLA-DRB1*03:05, *03:07, *04:11, *04:20, *11:10 or *13:10; results not shown) in different African populations endorse bi-directional population movements from the Sahara to the Mediterranean and South Europe in pre-historic and more recent times. In some cases, Lineal Megalithic Scripts and later classical letters (or syllables) are found together; it is the case of naviform scripts that are found mixed with Iberian-Tartessian semi-syllabary signs in the same context both in Iberian Peninsula and in Canary Islands. An evolutionary pathway of writing may be observed from primitive Lineal Megalithic Scripts to other more recent and complex alphabets or syllabaries like those found in Sitovo⁴¹, Gradeshnitsa⁴² or Iberian scripts widespread in the Iberian Peninsula, Canary Islands and continental North Africa⁴³. Also, other common archaeological elements have been found in Europe and Africa (Canary Islands), including Western Sahara territory^{44,45}. The findings pointed out in this study in relation to Sudanese HLA profile support this emigrational theory from the Sahara (and Sub-Saharan areas) to the Mediterranean Basin.

However, other authors with different techniques and genetic markers (SNPs) put forward that admixture with Arabs in recent times has occurred in Sudan and that this area has been subjected to significant population movements in ancient African /Eurasian history⁴⁶. Also, East Sahelians admixture with North Africans is confirmed with other studies⁴⁷, and Fulani nomads, Tuaregs and other Arab speaking Sahelians show admixture with North Africans like most of Sahelian groups do⁴⁸.

Data availability

All genotypic data included in present work is held in Allele Frequencies Net Database (<http://www.allelefrequencies.net>) with population number 3765 and population name “Sudan Khartoum HLA”.

Received: 19 December 2022; Accepted: 6 August 2023

Published online: 27 September 2023

References

1. The New Humanitarian. *Sudan: The Nuba Mountains: Straddling the North-South Divide* (2009). <https://www.refworld.org/docid/4b013b77c.html>. Accessed 15 Feb 2022.
2. Pflanz, M. *Sudan Referendum: What is Being Voted on and What Will Happen? The Telegraph* (2011). <https://www.telegraph.co.uk/news/worldnews/africaandindianocean/sudan/8246615/Sudan-referendum-whats-being-voted-on-and-what-will-happen.html>. Accessed 16 Feb 2022.
3. Gonem, A. & Samuels, M. (eds) *The Encyclopedia of the Peoples of the World* (The Jerusalem Publishing House, 1996).
4. Arnaiz-Villena, A., Gomez-Casado, E. & Martinez-Laso, J. Population genetic relationships between Mediterranean populations determined by HLA allele distribution and a historic perspective. *Tissue Antigens* **60**(2), 111–121 (2002).
5. Arnaiz-Villena, A., Juarez, I., Palacio-Grüber, J., Lopez-Nares, A. & Suarez-Trujillo, F. The Northern Migrations from a drying Sahara (6,000 years BP): Cultural and genetic influence in Greeks, Iberians and other Mediterraneans. *Int. J. Mod. Anthropol.* **2**(15), 484–507 (2021).
6. Clayton, J. & Lonjou, C. Allele and haplotype frequencies for HLA loci in various ethnic groups. In *Genetic Diversity of HLA: Functional and Medical Implications* (ed. Charron, D.) 665–820 (EDK, 1997).
7. Gomez-Casado, E. *et al.* HLA genes in Arabic-speaking Moroccans: Close relatedness to Berbers and Iberians. *Tissue Antigens* **55**, 239–249 (2000).
8. Izaabel, H. *et al.* HLA class II DNA polymorphism in a Moroccan population from the Souss, Agadir area. *Tissue Antigens* **51**, 106–110 (1998).
9. Roitberg-Tambur, A. *et al.* Comparative analysis of HLA polymorphism at the serologic and molecular level in Moroccan and Ashkenazi Jews. *Tissue Antigens* **46**, 104–110 (1995).
10. Arnaiz-Villena, A. *et al.* Relatedness among Basques, Portuguese, Spaniards, and Algerians studied by HLA allelic frequencies and haplotypes. *Immunogenetics* **47**, 37–43 (1997).
11. Martinez-Laso, J. *et al.* The contribution of the HLA-A, -B, -C and -DR, -DQ DNA typing to the study of the origins of Spaniards and Basques. *Tissue Antigens* **45**, 237–245 (1995).
12. Imanishi, T., Akaza, T., Kimura, A., Tokunaga, K. & Gojobori, T. Allele and haplotype frequencies for HLA and complement loci in various ethnic groups. In *HLA* (eds Tsuji, K. *et al.*) 1065–1220 (Oxford University Press, 1992).
13. Arnaiz-Villena, A. *et al.* HLA allele and haplotype frequencies in Algerians: Relatedness to Spaniards and Basques. *Hum. Immunol.* **43**, 259–268 (1995).

14. Martinez-Laso, J. *et al.* HLA DR and DQ polymorphism in Ashkenazi and non-Ashkenazi Jews: Comparison with other Mediterraneans. *Tissue Antigens* **47**, 63–71 (1996).
15. Arnaiz-Villena, A. *et al.* The origin of Cretan populations as determined by characterization of HLA alleles. *Tissue Antigens* **53**, 213–226 (1999).
16. Arnaiz-Villena, A. *et al.* HLA genes in Macedonians and the sub-Saharan origin of the Greeks. *Tissue Antigens* **57**, 118–127 (2001).
17. Mehra, N. K. *et al.* Genetic diversity of HLA in the populations of India, Sri Lanka and Iran. In *Genetic Diversity of HLA: Functional and Medical Implications* (ed. Charron, D.) 314–320 (EDK, 1997).
18. Arnaiz-Villena, A. *et al.* HLA alleles and haplotypes in the Turkish population: Relatedness to Kurds, Armenians and other Mediterraneans. *Tissue Antigens* **57**, 308–317 (2001).
19. Danilovs, J. A., Ayoub, G. & Terasaki, P. I. B lymphocyte isolation by thrombin-nylon wool. In *Histocompatibility Testing* (ed. Terasaki, P. I.) 287–288 (UCLA Tissue Typing Laboratory, 1980).
20. Excoffier, L., Laval, G. & Schneider, S. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol. Bioinform.* **1**, 47–50 (2007).
21. Imanishi, T., Akaza, T., Kimura, A., Tokunaga, K. & Gojobori, T. Estimation of allele and haplotype frequencies for HLA and complement loci. In *HLA* (eds Tsuji, K. *et al.*) 76–79 (Oxford University Press, 1992).
22. Imanishi, T., Akaza, T., Kimura, A., Tokunaga, K. & Gojobori, T. Genetic relationships among various human populations indicated by MHC polymorphisms. In *HLA* (eds Tsuji, K. *et al.*) 627–632 (Oxford University Press, 1992).
23. Saitou, N. & Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).
24. Nei, M. Genetic distances between populations. *Am. Nat.* **106**, 283 (1972).
25. Nei, M. Analysis of gene diversity in subdivided populations. *Proc. Nat. Acad. Sci. USA* **70**, 3321–3323 (1973).
26. Nei, M., Tajima, F. & Tatenno, Y. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *J. Mol. Evol.* **19**, 153–170 (1983).
27. Young, F. W. & Bann, C. M. A visual statistics system. In *Statistical Computing Environments for Social Research* (eds Stine, R. A. & Fox, J.) 207–236 (Sage Publications, 1996).
28. Hajjej, A. *et al.* HLA genes in Southern Tunisians (Ghannouch area) and their relationship with other Mediterraneans. *Eur. J. Med. Genet.* **49**, 43–56 (2006).
29. Dörk, T. *et al.* Evidence for a common ethnic origin of cystic fibrosis mutation 3120+1 GA in diverse populations. *Am. J. Hum. Genet.* **63**, 656–662 (1998).
30. Padoa, C., Goldman, A., Jenkins, T. & Ramsay, M. Cystic fibrosis carrier frequencies in populations of African origin. *J. Med. Genet.* **36**, 41–44 (1999).
31. Hajjej, A., Almawi, W. Y., Arnaiz-Villena, A., Hattab, L. & Hmida, S. The genetic heterogeneity of Arab populations as inferred from HLA genes. *PLoS ONE* **13**(3), e0192269. <https://doi.org/10.1371/journal.pone.0192269> (2018).
32. Aeschylus. *The supplicants* (ed. Perea-Morales, B.) (Editorial Gredos, 2010).
33. Ehret, C. *The civilizations of Africa: A history to 1800* (University Press of Virginia, 2002).
34. McCauley, J. F. *et al.* Subsurface valleys and geology of the Eastern Sahara revealed by shuttle radar. *Science* **218**(4576), 1004–1020 (1982).
35. Arnaiz-Villena, A., Ruiz-del-Valle, V., Lopez-Nares, A. & Suarez-Trujillo, F. Iberian inscriptions in Sahara Desert rocks (Ti-m Missaou, Ahaggar Mts. area, Algeria): first evidence of incise Iberian rock scripts in continental North Africa. *Int. J. Mod. Anthrop.* **2**(15), 440–467 (2021).
36. Arnaiz-Villena, A., Martinez-Laso, J. & Alonso-Garcia, J. Iberia population genetics, anthropology and linguistics. *Hum. Biol.* **71**, 725–743 (1999).
37. Arnaiz-Villena, A. & Alonso-Garcia, J. The Usko-Mediterranean languages. In *Prehistoric Iberia* (ed. Arnaiz-Villena, A.) 205–246 (Kluwer Academic/Plenum Publishers, 2000).
38. Arnaiz-Villena, A. Prehistoric Iberia: Genetics, anthropology and linguistics. In *The Usko-Mediterranean languages* (ed. Arnaiz-Villena, A.) (Kluwer Academic/Plenum Publishers, 2000).
39. Arnaiz-Villena, A., Martinez-Laso, J. & Alonso-Garcia, J. The correlation between languages and genes: The Usko-Mediterranean peoples. *Hum. Immunol.* **62**, 1051–1061 (2001).
40. Arnaiz-Villena, A. *et al.* The Saharo-Canarian circle: The forgotten Prehistory of Euro African Atlantic façade and its lack of eastern demic diffusion evidences. *Int. J. Mod. Anthrop.* **2**(16), 586–600 (2021).
41. Morfova, Z. R. *Nadpispi Sitovo, Plovdivsk aokolia* (Godishnik Nanarodnia Arheologiches Kimuzei, 1950).
42. Fisher, S. R. *History of Writing* (Reaktion Books, 2003).
43. Arnaiz-Villena, A. *et al.* The Iberian-Tartessian semi-syllabary: Possible evolution from lineal megalithic/paleolithic scripts and the mother goddess religion. *Int. J. Mod. Anthrop.* **2**(17), 820–841 (2022).
44. Clarke, J. & Brooks, N. *The Archaeology of Western Sahara* (Oxford Books, 2018).
45. Arnaiz-Villena, A. *et al.* The Ibero-Guanche (Latin) rock inscriptions found at Mt. Tenezara volcano (Lanzarote, Canary Islands, Spain): A Saharan hypothesis for Mediterranean/Atlantic Prehistory. *Int. J. Mod. Anthrop.* **2**(13), 140–162 (2020).
46. Hollfelder, N. *et al.* Northeast African genomic variation shaped by the continuity of indigenous groups and Eurasian migrations. *PLoS Genet.* **13**(8), e1006976 (2017).
47. Triska, P. *et al.* Extensive admixture and selective pressure across the Sahel Belt. *Genome Biol. Evol.* **7**(12), 3484–3495 (2015).
48. Fortes-Lima, C. *et al.* Demographic and selection histories of populations across the Sahel/Savannah Belt. *Mol. Biol. Evol.* **39**(10), 209 (2022).

Acknowledgements

We thank all Sudanese people that acceded to participate in this work and Carlos Suarez-Sanchez for his collaboration in manuscript correction.

Author contributions

A.A.-V. performed research planning and writing article. References were revised by I.J., A.S.-O. and J.P.-G. S.E.-S. collected the samples and made part of typing procedures. F.S.-T. performed laboratory procedures, figures design and writing of the paper. All authors read and approved the final manuscript.

Funding

This work was supported by grants from the Spanish Ministry of Science and Universities (PI18/00721 and PI18/00626) and European FEDERS funds.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.A.-V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023