ELSEVIER

HLA Polymorphisms in Cabo Verde and Guiné-Bissau Inferred From Sequence-Based Typing

Hélder Spínola, Jácome Bruges-Armas, Derek Middleton, and António Brehm

ABSTRACT: Human leukocyte antigen (HLA)-A, -B, and -DRB1 polymorphisms were examined in the Cabo Verde and Guiné-Bissau populations. The data were obtained at high-resolution level, using sequence-based typing. The most frequent alleles in each locus was: A*020101 (16.7% in Guiné-Bissau and 13.5% in Cabo Verde), B*350101 (14.4% in Guiné-Bissau and 13.2% in Cabo Verde), DRB1*1304 (19.6% in Guiné-Bissau), and DRB1*1101 (10.1% in Cabo Verde). The predominant three loci haplotype in Guiné-Bissau was A*2301-B*1503-DRB1*1101 (4.6%) and in Cabo Verde was A*3002-B*350101-DRB1*1001 (2.8%), exclusive to northwestern islands (5.6%) and absent in Guiné-Bissau. The present study corroborates historic sources and other genetic studies that say Cabo Verde were populated not only by Africans but also by Europeans. Haplotypes and dendrogram analysis shows a Caucasian genetic influence in today's gene pool of Cabo Verdeans. Haplotypes and allele frequencies present a differential distribution between southeastern and northwestern Cabo Verde islands, which could be the result of different genetic influences, founder effect, or bottlenecks. Dendrograms and principal coordinates analysis show that Guineans are more similar to North Africans than other HLA-studied sub-Saharans, probably from ancient and recent genetic contacts with other peoples, namely East Africans. *Human Immunology 66, 1082–1092 (2005).* © American Society for Histocompatibility and Immunogenetics, 2005. Published by Elsevier Inc.

KEYWORDS: HLA polymorphism; SBT; Cabo Verde; Guiné-Bissau

ABBREVIATIONS

HLA human leukocyte antigen

SBT senquence-based typing

INTRODUCTION

One of the most polymorphic genetic systems in humans is the human leukocyte antigen (HLA), which consists of a closely linked set of genes highly important in transplantation, anthropologic, and forensic fields [1]. The characterization of the HLA profile of a population, with high throughput techniques as sequence-based typing (SBT), is a helpful mechanism to trace genetic relationships between neighboring populations. Previous studies have shown that haplotype frequencies are characteristic of particular populations and even certain alleles are exclusively found in some ethnic groups [2–4].

Some sub-Saharan African populations have been studied from the HLA point of view [5–7]. However, high-resolution HLA allele typing data are still scarce among sub-Saharans, especially West African populations. Guiné-Bissau (Figure 1) is one of these African Atlantic coast countries that were never HLA typed. The human settlement of the West Coast of Africa is the result of a continuous complex network of migrations, invasions, and admixture of peoples from different origins that began around 40,000 years before present [8]. Before Sahara desertification took place (9,000 years before present), several Neolithic cultures flourished in the area, bringing together people of sub-Saharan and North African origin [9]. Later, around 4,000 years before present, this region became the centre of migrations by different ethnic groups, namely Fula and East Africa

From the Human Genetics Laboratory, University of Madeira, Campus of Penteada, Funchal, Portugal (H.S., A.B.), Hospital Santo Espírito de Angra do Heroísmo, Terceira, Azores, Portugal (J.B.-A.), and Northern Ireland Histocompatibility and Immunogenetics Laboratory, City Hospital and University of Ulster, Belfast, N. Ireland (D.M.).

Address reprint requests to: António Brehm, University of Madeira, Campus of Penteada, 9000-399 Funchal, Portugal

Supported by European Community Program Interreg IIIB Madeira-Açores-Canárias (project MAC/2.3/A6).

Received June 15, 2005; accepted August 31, 2005.

0198-8859/05/$–see front matter
doi:10.1016/j.humimm.2005.09.001

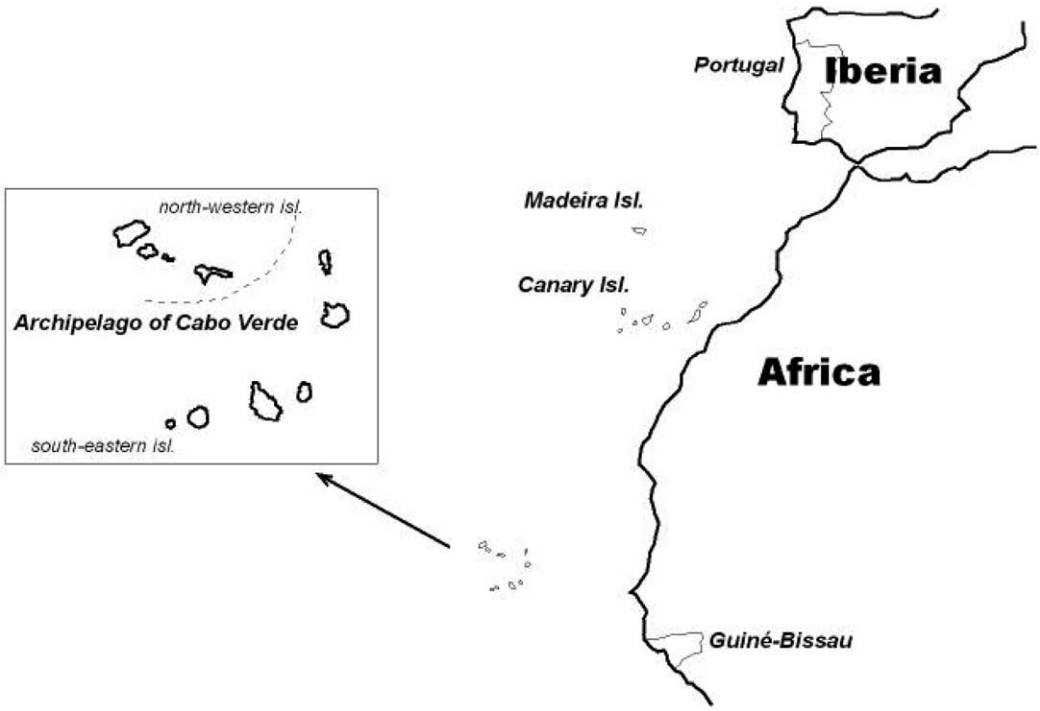

FIGURE 1 Iberian peninsula and the West African coast with Guiné-Bissau and the Atlantic archipelagos of Cabo Verde, the Canary Islands, and Madeira.

Shuwa Arabs, through the Sahel Corridor used for travel between the East and West coasts of Africa [6]. The most recent historic events, such as admixture of Berber and native populations from Muslim pressure in Northwestern Africa in the ninth century, also affected this region [10].

This region was also at the origin of most slave trade in the 15th and 16th centuries to Europe and America [11]. The Cabo Verde archipelago (Figure 1) was discovered inhabited around 1460 and settled by a few Europeans (less than 10%), whereas African slaves originating from the Guiné coast constituted most of the remaining population [12]. Northwestern Cabo Verde Islands were populated much later than southeastern islands, in the 17th century, and tradition says that settlers were probably mainly constituted by a subset of fugitive slaves from other islands [13].

TABLE 1 Ambiguous HLA alleles and their respective assignment labels under which they appear in the text

Locus	Ambiguous alleles	Reference
HLA-A	A*010101/A*0104N	A*010101
	A*020101/A*020108/A*0209/A*0243N	A*020101
	A*2301/A*2307N	A*2301
	A*240201/A*240203/A*2409N/A*2411N	A*240201
	A*680102/A*6811N	A*680102
HLA-B	B*0705/B*0706	B*0705/06
	B*180101/B*1817N	B*180101
	B*2705/B*2713	B*2705
	B*350101/B*3540N/B*3542	B*350101
	B*440201/B*4419N/B*4427	B*440201
	B*510101/B*5111N/B*5130/B*5132	B*510101
HLA-DRB1	DRB1*120101/DRB1*1206	DRB1*120101

Studies on mtDNA, STR, and Y-chromosome SNPs have been done for populations of Guiné-Bissau and Cabo Verde [13–17]. The population of Cabo Verde has been already HLA typed, but at a low-resolution level and in a small number of samples [18]. The main goal of the present work is to present a clear genetic profile of Cabo Verdeans and Guineans based on high-resolution typing of the HLA-A, -B, and -DRB1 loci and evaluate their relatedness to each other and other African and European populations. Because previous Cabo Verde HLA typing was performed in a small number of samples ($n = 64$), the present study includes for the first time Guiné-Bissau and increases substantially the number of samples typed, allowing a comparison between different groups of Cabo Verde islands (northwestern and southeastern). HLA data comparison between Cabo Verde and Guiné-Bissau may also contribute to clarify the origin of the genetic heterogeneity already found but now using

TABLE 2 HLA-A, -B, and -DRB1 allele frequencies in Guiné-Bissau and Cabo Verde populations estimated by maximum likelihood

HLA-A alleles	Frequencies				HLA-B alleles	Frequencies				HLA-DRB1 alleles	Frequencies			
	Guiné	CV (T)	CV (NW)	CV (SE)		Guiné	CV (T)	CV (NW)	CV (SE)		Guiné	CV (T)	CV-NW	CV-SE
010101	0.046	0.049	0.048	0.048	070201	0.023	0.040	0.024	0.056	0101	0	0.048	0.081	0.016
0102	0.008	0.008	0	0.016	0705/06	0.023	0.012	0.008	0.016	010201	0.023	0.052	0.016	0.089
020101	0.167	0.135	0.149	0.121	0801	0.077	0.065	0.056	0.073	0103	0	0.004	0	0.008
0202	0.062	0.016	0.008	0.024	1302	0.023	0	0	0	030101	0.046	0.081	0.056	0.105
0205	0	0.004	0	0.008	1303	0	0.004	0	0.008	030201	0	0.048	0.073	0.024
0217	0	0.004	0	0.008	1402	0.015	0.065	0.032	0.097	0303	0.008	0.008	0.008	0.008
030101	0.062	0.089	0.113	0.065	150101	0	0.044	0.048	0.040	0307	0.008	0	0	0
030103	0	0.004	0.008	0	1503	0.108	0.056	0.048	0.065	040101	0	0.02	0.024	0.016
110101	0	0.024	0.024	0.024	1510	0.038	0.004	0.008	0	0402	0	0.008	0.016	0
2301	0.156	0.113	0.081	0.148	1515	0	0.004	0.008	0	040301	0	0.008	0.016	0
240201	0.015	0.069	0.048	0.089	1516	0	0.004	0	0.008	0404	0	0.008	0	0.016
240301	0	0.008	0.008	0.008	151701	0	0.004	0	0.008	040501	0.077	0.040	0.040	0.040
2601	0.031	0.040	0.048	0.032	1518	0.008	0.008	0.008	0.008	070101	0.062	0.069	0.065	0.073
2608	0	0.004	0.008	0	180101	0.038	0.036	0.040	0.032	080101	0	0.008	0.008	0.008
2902	0.015	0.032	0.016	0.048	2702	0	0.008	0.008	0.008	080302	0	0.004	0	0.008
3001	0.038	0.044	0.056	0.032	2703	0.015	0.004	0	0.008	080401	0.008	0.008	0.008	0.008
3002	0.031	0.093	0.121	0.065	2705	0	0.020	0.024	0.016	0806	0.008	0.008	0.016	0
310102	0	0.020	0.032	0.008	350101	0.144	0.132	0.174	0.089	090102	0.077	0.060	0.032	0.089
3201	0.038	0.044	0.048	0.040	3502	0	0.004	0	0.008	1001	0.10	0.060	0.081	0.048
3301	0.046	0.020	0.032	0.008	3503	0	0.008	0.008	0.008	1101	0.108	0.101	0.106	0.089
3303	0.085	0.032	0.024	0.040	3701	0.015	0.008	0.016	0	1102	0.077	0.028	0.016	0.040
3402	0.054	0.004	0	0.008	3801	0	0.004	0	0.008	110401	0.008	0.016	0.008	0.024
6601	0.015	0.016	0.024	0.008	3910	0.008	0	0	0	1127	0	0.008	0.016	0
680101	0.023	0.024	0.040	0.008	4001	0	0.004	0.008	0	120101	0.008	0.016	0.024	0.008
680102	0	0.004	0.008	0	4002	0.008	0.016	0.024	0.008	130101	0.031	0.069	0.073	0.065
6802	0.023	0.028	0.008	0.048	4006	0	0.004	0.008	0	130201	0.077	0.088	0.089	0.081
680301	0	0.004	0.008	0	4101	0	0.008	0	0.016	130301	0	0.012	0.016	0.008
6815	0	0.004	0	0.008	4102	0.008	0.004	0.008	0	130302	0	0.004	0.008	0
6901	0	0.032	0.040	0.024	4103	0.031	0	0	0	1304	0.196	0.048	0.024	0.073
7401	0.077	0.028	0	0.056	4201	0.015	0.024	0.032	0.016	1317	0.008	0	0	0
8001	0.008	0.004	0	0.008	4202	0.008	0.004	0	0.008	140101	0.015	0.012	0.024	0
					440201	0.008	0.004	0	0.008	140701	0.008	0	0	0
					440301	0.031	0.028	0.016	0.040	1412	0.008	0	0	0
					4405	0	0.004	0	0.008	150101	0.008	0.040	0.032	0.048
					4501	0.062	0.012	0.008	0.016	150201	0	0.012	0.024	0
					4701	0	0.008	0.008	0.008	160201	0.031	0.004	0	0.008
					4901	0.031	0.032	0.024	0.040					
					5001	0.015	0.012	0.008	0.016					
					510101	0.015	0.077	0.081	0.073					
					510102	0	0.004	0.008	0					
					5108	0	0.004	0	0.008					
					520101	0	0.004	0.008	0					
					520102	0	0.004	0.008	0					
					5301	0.10	0.097	0.089	0.107					

(Continued)

TABLE 2 *(Continued)*

HLA-A alleles	Frequencies				HLA-B alleles	Frequencies				HLA-DRB1 alleles	Frequencies			
	Guiné	CV (T)	CV (NW)	CV (SE)		Guiné	CV (T)	CV (NW)	CV (SE)		Guiné	CV (T)	CV-NW	CV-SE
					5601	0.008	0	0	0					
					570101	0	0.016	0.032	0					
					5702	0.008	0	0	0					
					570301	0.008	0.028	0.032	0.024					
					5801	0.078	0.036	0.032	0.040					
					5802	0	0.032	0.056	0.008					
					7801	0.023	0	0	0					
					8201	0.008	0	0	0					

Abbreviations: CV(T) = Cabo Verde total; CV(NW) = Cabo Verde Northwestern; CV(SE) = Cabo Verde Southeastern.

markers not sex-dependent and thus qualified for analysis of admixture that is not sex-biased [13–17].

MATERIALS AND METHODS

Population Samples and HLA Typing

The present study population consisted of a total of 189 healthy unrelated males randomly distributed in Cabo Verde Northwestern islands (Santo Antão, São Vicente, and São Nicolau) ($n = 62$), Cabo Verde southeastern islands (Sal, Boavista, Maio, Santiago, Fogo, and Brava) ($n = 62$) and Guiné ($n = 65$). Blood samples were collected after informed consent, from donors whose parents and grandfathers were born and living on the same island (Cabo Verde) or belong to the same ethnic group (Guiné-Bissau). Guiné-Bissau samples were identified as belonging to seven different ethnic groups: Balanta ($n = 10$), Papel ($n = 11$), Mandinga ($n = 9$), Felupe ($n = 5$), Bijagós ($n = 10$), Fula ($n = 10$), and Mancanha ($n = 10$). Genomic DNA was isolated from whole blood containing EDTA using a phenol-chloroform procedure and frozen at −20 °C until use. All subjects were typed for HLA-A, HLA-B, and HLA-DRB1 loci. SBT of the HLA-A and -B loci was performed on exon 2 and exon 3 accordingly to Kurz *et al.* [19] and Pozzi *et al.* [20] with minor modifications. DNA fragments amplified by polymerase chain reaction were purified and sequenced using ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kits (Applied Biosystems, Foster City, CA) in an ABI Prism 310 Genetic Analyzer (Applied Biosystems), according to the manufacturer's instructions. HLA-DRB1 SBT was performed on exon 2, amplified, and sequenced with group specific primers used in sequence-specific oligonucleotide probe typing [21]. All samples were polymerase chain reaction amplified with each of the group-specific pair of primers. The positive reactions were sequenced with the amplification primers according to this description for HLA-A and -B. Sequencing was always performed in forward and reverse directions and processed using the Matchtools Allele Identification packet (Applied Biosystems) that assess the typing by an updated HLA sequence library with alleles defined up to June 2001. In the present study, we did not test for polymorphisms outside exons 2 and 3 in HLA-A and -B and outside exon 2 in HLA-DRB1. Therefore, a few groups of alleles could not be distinguished. A summary of the unresolved ambiguities and the reference used in each case in the text and for frequency analysis is given in Table 1. Likewise, some heterozygous samples could have more than one allele's combination. In these situations, for allele frequencies and haplotype calculations, we used a maximum likelihood estimation assignment, choosing the combination of the most frequent single alleles.

TABLE 3 Most common HLA-A, -B, -DRB1 three-loci haplotypes in Guiné-Bissau, Cabo Verde total, Cabo Verde Northwestern, and Cabo Verde Southeastern populations as estimated by maximum likelihood Only frequencies above 0.01, with statistical significance linkage disequilibrium ($p < 0.001$), are shown.

Haplotypes				
Guiné-Bissau	Frequencies	SD	D	D′
A*2301-B*1503-DRB1*1101	0.046	0.017	0.0442	1
A*020101-B*350101-DRB1*1102	0.031	0.017	0.0291	0.658
A*020101-B*350101-DRB1*130201	0.023	0.013	0.0211	0.477
A*020101-B*440301-DRB1*160201	0.023	0.012	0.0228	0.516
A*0202-B*350101-DRB1*1001	0.023	0.012	0.0221	0.5
A*010101-B*5801-DRB1*1304	0.015	0.012	0.0143	0.324
A*020101-B*0705/06-DRB1*1304	0.015	0.013	0.0142	0.321
A*0202-B*2703-DRB1*1304	0.015	0.007	0.0148	0.335
A*030101-B*350101-DRB1*040501	0.015	0.012	0.0143	0.324
A*2301-B*5301-DRB1*090102	0.015	0.015	0.0138	0.312
A*2301-B*5801-DRB1*1101	0.015	0.009	0.0136	0.308
A*2601-B*0801-DRB1*1304	0.015	0.015	0.0145	0.328
A*2902-B*4501-DRB1*1101	0.015	0.013	0.0149	0.337
A*3002-B*0801-DRB1*1304	0.015	0.011	0.0145	0.328
A*3301-B*5301-DRB1*090102	0.015	0.009	0.0146	0.33
A*3303-B*070201-DRB1*1102	0.015	0.007	0.0148	0.335
A*3303-B*1510-DRB1*1001	0.015	0.01	0.0147	0.333
A*3402-B*180101-DRB1*040501	0.015	0.013	0.0148	0.335
A*6601-B*4501-DRB1*1001	0.015	0.013	0.0149	0.337
A*7401-B*350101-DRB1*1001	0.015	0.019	0.0139	0.314
A*7401-B*4901-DRB1*1304	0.015	0.009	0.0145	0.328
A*7401-B*5301-DRB1*1304	0.015	0.007	0.0135	0.305
Cabo Verde total	Frequencies	SD	D	D′
A*3002-B*350101-DRB1*1001	0.028	0.011	0.0273	1
A*6901-B*150101-DRB1*130201	0.024	0.011	0.0239	0.875
A*030101-B*5802-DRB1*030101	0.023	0.011	0.0228	0.835
A*2301-B*5301-DRB1*090102	0.016	0.009	0.0153	0.56
A*020101-B*1503-DRB1*1101	0.012	0.008	0.0112	0.41
A*2301-B*1402-DRB1*1101	0.012	0.007	0.0113	0.414
A*2301-B*4901-DRB1*130201	0.012	0.007	0.0117	0.429
A*2301-B*5801-DRB1*090102	0.012	0.008	0.0118	0.432
A*2601-B*5801-DRB1*1304	0.012	0.008	0.0119	0.436
A*3002-B*0801-DRB1*040101	0.012	0.008	0.0119	0.436
A*3002-B*180101-DRB1*030101	0.012	0.006	0.0117	0.429
A*3301-B*1503-DRB1*1001	0.012	0.008	0.0119	0.436
Cabo Verde Northwestern	Frequencies	SD	D	D′
A*3002-B*350101-DRB1*1001	0.056	0.022	0.0543	1
A*030101-B*5802-DRB1*030101	0.040	0.018	0.0396	0.73
A*6901-B*150101-DRB1*130201	0.032	0.017	0.0318	0.586
A*3002-B*0801-DRB1*040101	0.024	0.014	0.0238	0.438
A*010101-B*3701-DRB1*070101	0.016	0.013	0.016	0.295
A*020101-B*2705-DRB1*0101	0.016	0.01	0.0157	0.289
A*020101-B*350101-DRB1*1101	0.016	0.014	0.0133	0.245
A*020101-B*570101-DRB1*070101	0.016	0.011	0.0157	0.289
A*110101-B*350101-DRB1*0101	0.016	0.012	0.0157	0.289
A*2301-B*1402-DRB1*0806	0.016	0.015	0.016	0.295
A*2301-B*4901-DRB1*130201	0.016	0.01	0.0158	0.291
A*2301-B*5801-DRB1*090102	0.016	0.012	0.0159	0.293
A*3001-B*510101-DRB1*1101	0.016	0.013	0.0155	0.285
A*3002-B*180101-DRB1*030201	0.016	0.014	0.0156	0.287
A*3201-B*070201-DRB1*120101	0.016	0.013	0.016	0.295
A*3301-B*1503-DRB1*1001	0.016	0.013	0.0159	0.293
A*680101-B*350101-DRB1*040501	0.016	0.012	0.0157	0.289

TABLE 3 *(Continued)*

Haplotypes

Guiné-Bissau	Frequencies	SD	D	D′
Cabo Verde Southeastern	Frequencies	SD	D	D′
A*2301-B*5301-DRB1*090102	0.032	0.017	0.0306	1
A*2301-B*1402-DRB1*1101	0.024	0.015	0.0227	0.742
A*010101-B*350101-DRB1*030101	0.016	0.012	0.0156	0.51
A*010101-B*440301-DRB1*070101	0.016	0.011	0.0159	0.52
A*0102-B*4901-DRB1*040501	0.016	0.012	0.016	0.523
A*020101-B*1503-DRB1*1101	0.016	0.012	0.0153	0.5
A*020101-B*4901-DRB1*090102	0.016	0.011	0.0156	0.51
A*020101-B*510101-DRB1*130101	0.016	0.013	0.0154	0.503
A*0202-B*350101-DRB1*1102	0.016	0.012	0.0159	0.52
A*030101-B*4101-DRB1*070101	0.016	0.012	0.0159	0.52
A*2301-B*070201-DRB1*150101	0.016	0.014	0.0156	0.51
A*240201-B*070201-DRB1*150101	0.016	0.012	0.0158	0.516
A*240201-B*2705-DRB1*040501	0.016	0.013	0.0159	0.52
A*240201-B*510101-DRB1*1102	0.016	0.012	0.0157	0.513
A*2601-B*5801-DRB1*1304	0.016	0.01	0.0159	0.52
A*2902-B*1402-DRB1*010201	0.016	0.011	0.0156	0.51
A*3002-B*180101-DRB1*030101	0.016	0.012	0.0158	0.516
A*6901-B*150101-DRB1*130201	0.016	0.012	0.0159	0.52
A*7401-B*5301-DRB1*1304	0.016	0.014	0.0156	0.51

Abbreviations: SD = standard deviation; D = linkage disequilibrium; D′ = relative LD.

Data Analysis

Basic genetic parameters (allele and haplotype frequencies, gene diversity, and Hardy-Weinberg equilibrium) at the three loci were estimated with Arlequin v2.000 [22]. In the present study, haplotypes were inferred using a maximum likelihood approach. Linkage disequilibrium (D) and relative LD (D′) between alleles at different *loci* and their level of significance (P) was calculated according to Weir [23]. Ewens-Watterson's, Slatkin's, and Chakraborty's selective neutrality statistical tests were applied to examine the presence of selective forces influencing allelic diversity at each locus.

The Cabo Verde and Guiné-Bissau data were compared with several populations that are available at the same typing resolution (www.allelefrequencies.net) [6, 7, 24–27]. To perform a principal coordinates analysis, data with high resolution were normalized to a four-digit typing. To compare our samples with previously published data from other populations, dendrograms were constructed based on allele frequencies of the same high-resolution typing but collapsed into two-digit, low-resolution typing. An analysis of molecular variance was performed with these populations based on Euclidean distances between all pairs of haplotypes [28]. The total genetic variation between Cabo Verde northwestern and Cabo Verde southeastern was estimated and the correspondent FST value was used to evaluate if there was significant difference between them. Variance components were tested for significance by nonparametric randomization tests using 10,000 permutations under the null hypothesis of no population structure. The population genetic software Arlequin v2.000 was employed in all these analyses.

Comparative analysis of the Cabo Verde and Guiné-Bissau data set with other populations available in the literature was achieved using the software included in the PHYLIP v.3.6 package (Department of Genome Sciences, University of Washington, Seattle) [29]. First, SEQBOOT was used to perform a bootstrap analysis from gene frequency data. The program generates multiple data sets resampled from the original data. Distance matrices from each replicate data set were generated using GENDIST and used as input to NEIGHBOR to produce neighbor-joining trees. A single consensus bootstrapped tree was obtained with CONSENSUS. The topology was visualized with TreeView [30]. Dendrograms were based on HLA-A, -B, and -DRB1 or on HLA-A and -B to include some relevant populations without HLA-DRB1 typing.

Principal coordinates analysis using HLA-A and HLA-B allele frequencies was carried out on the Multi-Variate Statistical Package MVSP3 for Windows (Kovach Computing Services, Anglesey, Wales; http://www.kovcomp.com/mvsp).

RESULTS

Table 2 shows the allele frequencies of HLA-A, HLA-B, and HLA-DRB1 loci in Cabo Verde and Guiné-Bissau

populations. A total of 31 HLA-A, 45 HLA-B, and 32 HLA-DRB1 alleles were found in Cabo Verde and 20 HLA-A, 31 HLA-B, and 23 HLA-DRB1 alleles were found in Guiné-Bissau. The populations were in Hardy-Weinberg equilibrium at each locus, except HLA-DRB1 in Cabo Verde. The heterozygosity was high for all examined loci (Cabo Verde HLA-A: 0.94; HLA-B: 0.97; HLA-DRB1: 0.95; Guiné-Bissau HLA-A: 0.86; HLA-B: 0.94; HLA-DRB1: 0.89). Ewens-Watterson, Slatkin's, and Chakraborty's tests of selective neutrality yielded significant results at HLA-A and HLA-DRB1 in Cabo Verde and HLA-DRB1 in Guiné-Bissau, suggesting that selection was acting.

An analysis of molecular variance between the two subpopulations of Cabo Verde (northwestern and southeastern) shows that only 0.13% of the total genetic variation found can be attributed to differences among them; the remaining is due to within subpopulation differences. The exact test of population differentiation performed by Arlequin, which tests nonrandom distribution of allele frequencies in population samples under the hypothesis of panmixia, shows significant results ($p = 0.007$) between Cabo Verde northwestern and Cabo Verde southeastern groups of islands.

Guiné-Bissau Allele Frequencies

The most frequent HLA-A alleles found in Guiné-Bissau was A*020101 (16.7%) and A*2301 (15.6%). Two alleles follow with frequencies ranging between 7% and 9% (A*3303 and A*7401). HLA-A*3303 shows a higher frequency than other sub-Saharans, except Mali (9.4%) [24]. HLA-A*7401 is a typical African allele that present similar frequencies in other sub-Saharan populations [24]. Alleles A*6801 and A*6802 present the same frequency (2.3% each) in opposition to other sub-Saharan populations that have a higher prevalence of the second [24].

HLA-B locus presents the allele B*350101 (14.4%) as the most frequent in Guiné-Bissau, a prevalence higher than in other sub-Saharans (www.allelefrequencies.net). The next most frequent alleles found in Guiné were B*1503 (10.8%) and B*5301 (10%). HLA-B*1503 is a typical sub-Saharan allele and shows one of the highest frequencies found in the sub-Sahara. HLA-B*5301 present similar frequencies in other sub-Saharans [24]. HLA-B*4103 present a relatively high frequency (3.1%) because it is a very rare allele and it was absent in sub-Saharan populations previously studied [24].

Allele DRB1*1304 (19.6%) was the most frequent at HLA-DRB1 locus, but is surprisingly absent in the Cameroon and residual in Morocco (0.5%) [7, 24]. The next most frequent alleles were DRB1*1101 (10.8%) and DRB1*1001 (10%), with frequencies much higher than in Cameroon (4.8% and 1.1%, respectively) [7].

Cabo Verde Allele Frequencies

The most frequent HLA-A alleles found in the Cabo Verde were A*020101 (13.5%) and A*2301 (11.3%), typical frequencies in sub-Saharan populations [24]. Three alleles follow with frequencies ranging between 7% and 9% (A*240201, A*030101, and A*3002). HLA-A*240201 (6.9%) shows a higher frequency than other sub-Saharans and similar to Europeans [24]. HLA-A*030101 (8.9%) presents an intermediate frequency between sub-Saharan populations (3–8%) and Europeans (8–20%) [24]. HLA-A*3002 (9.3%) has a prevalence in the extended range found in the sub-Sahara (3–23%). Alleles A*6801 and A*6802 present the same frequency (2.8% each) in opposition to other sub-Saharans that have a higher prevalence of A*6802 [24].

Allele B*350101 (13.2%) is the most frequent at HLA-B locus in Cabo Verde, a prevalence higher than other sub-Saharans except Mali, which is similar [24]. The next most frequent alleles found in Cabo Verde were B*5301 (9.7%) and B*510101 (7.7%), higher and similar to sub-Saharans, respectively [24].

The DRB1*1101 (10.1%), with a frequency similar to Guiné-Bissau, was the most frequent HLA-DRB1 allele in Cabo Verde, followed by DRB1*130201 (8.8%), similar to Guiné-Bissau, and DRB1*030101 (8.1%), similar to Cameroon [7].

Haplotype Frequencies

The exact test of linkage disequilibrium between the three pairs of loci was statistically significant in Guiné-Bissau and Cabo Verde. The most representative (frequency >1%) three-loci haplotypes are listed in Table 3. The complete list of the two- and three-loci haplotypes found in Guiné-Bissau and Cabo Verde is available from the authors on request.

A*2301-B*1503-DRB1*1101 (4.6%) was the most frequent three-loci haplotype in Guiné-Bissau, also present in Cabo Verde, but with a residual frequency (0.4%). Its partial two-loci haplotype A*2301-B*1503 (5.9% in Guiné-Bissau) has been found in several sub-Saharan populations [5, 24] and in Cabo Verde, being exclusive to the Southeastern islands (4%). The second most frequent three-loci haplotype in Guiné-Bissau A*020101-B*350101-DRB1*1102 (3.1%) is absent in Cabo Verde. The partial A*020101-B*350101 haplotype (5.4% in Guiné-Bissau), also absent in Cabo Verde, has been found in Malians at 2.2% [5]. Three other haplotypes (A*020101-B*350101-DRB1*130201; A*020101-B*440301-DRB1*160201; and A*0202-B*350101-DRB1*1001), were found in Guiné-Bissau each at a frequency of 2.3%, but absent in Cabo Verde.

The most frequent three loci haplotype in Cabo Verde was A*3002-B*350101-DRB1*1001 (2.8%), exclusive to

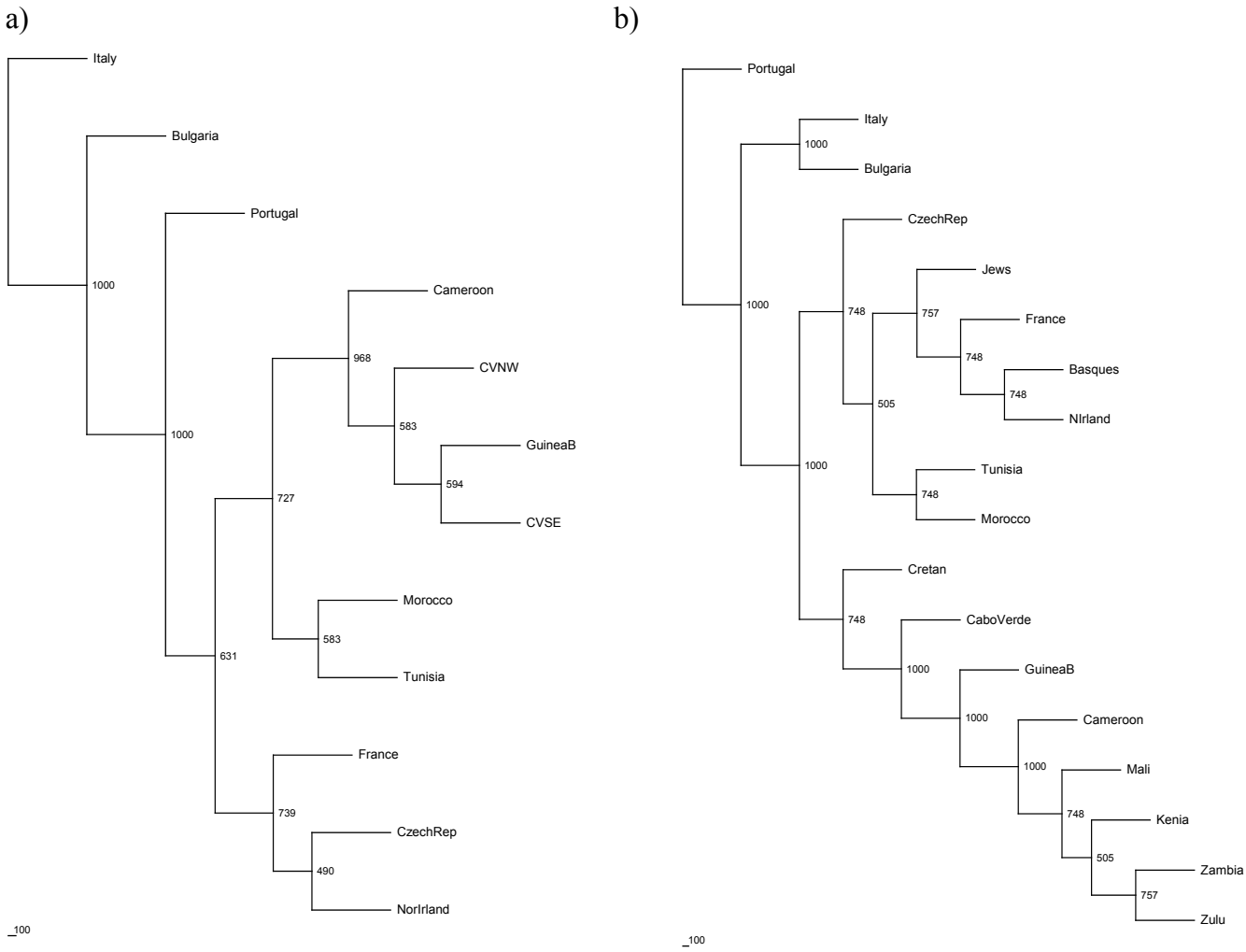

FIGURE 2 Neighbor-joining (NJ) dendrograms showing the comparative position of Cabo Verde and Guiné-Bissau populations with other populations. Standard genetic distances among populations were calculated using (A) human leukocyte antigen (HLA)-A, HLA-B, and HLA-DRB1 low-resolution allele frequencies with Cabo Verde northwestern (CVNW) and Cabo Verde southeastern (CVSE) and (B) HLA-A and HLA-B low-resolution allele frequencies. Numbers above branches are node support after the bootstrap technique implemented in PHYLIP package program BOOT. For reference of the populations used, see Data Analysis.

Northwestern islands (5.6%) and absent in Guiné-Bissau. The partial haplotype A*3002-B*350101 (7.3% in Cabo Verde Northwestern and 0.8% in Guiné-Bissau) has not been found in other sub-Saharans [5, 24]. The other two most frequent haplotypes in Cabo Verde, A*6901-B*150101-DRB1*130201 (2.4%) and A*030101-B*5802-DRB1*030101 (2.3%), were also absent in Guiné-Bissau. Partial haplotype A*6901-B*150101 was not found in other sub-Saharan and is more prevalent in northwestern (3.2%) than southeastern (1.6%) Cabo Verde islands. A*030101-B*5802 is exclusive to northwestern islands (5.6%) and despite it was not being found in Guiné-Bissau is present in Kenyans, Zulus, and Cameroons [5, 6].

From the haplotypes present in more than 1% in the Cabo Verde Islands, only two (A*2301-B*5801-DRB1*090102, A*6901-B*150101-DRB1*130201) were present on both islands. These two haplotypes are absent in Guiné-Bissau and more prevalent in northwestern than southeastern Cabo Verde islands. Only southeastern islands present haplotypes common to the most frequent found in Guiné-Bissau (A*2301-B*5301-DRB1*090102 and A*7401-B*5301-DRB1*1304).

More than nine percent (9.3%) of all three-loci haplotypes found in Cabo Verde are common to Portugal [27]. Some of these haplotypes were considered previously as European, namely A*020101-B*440301-DRB1*070101, A*2902-B*440301-DRB1*070101, and A*3002-B*180101-DRB1*030101 [26], but most of them were not yet described in other populations. Only 2.2% of Cabo Verde

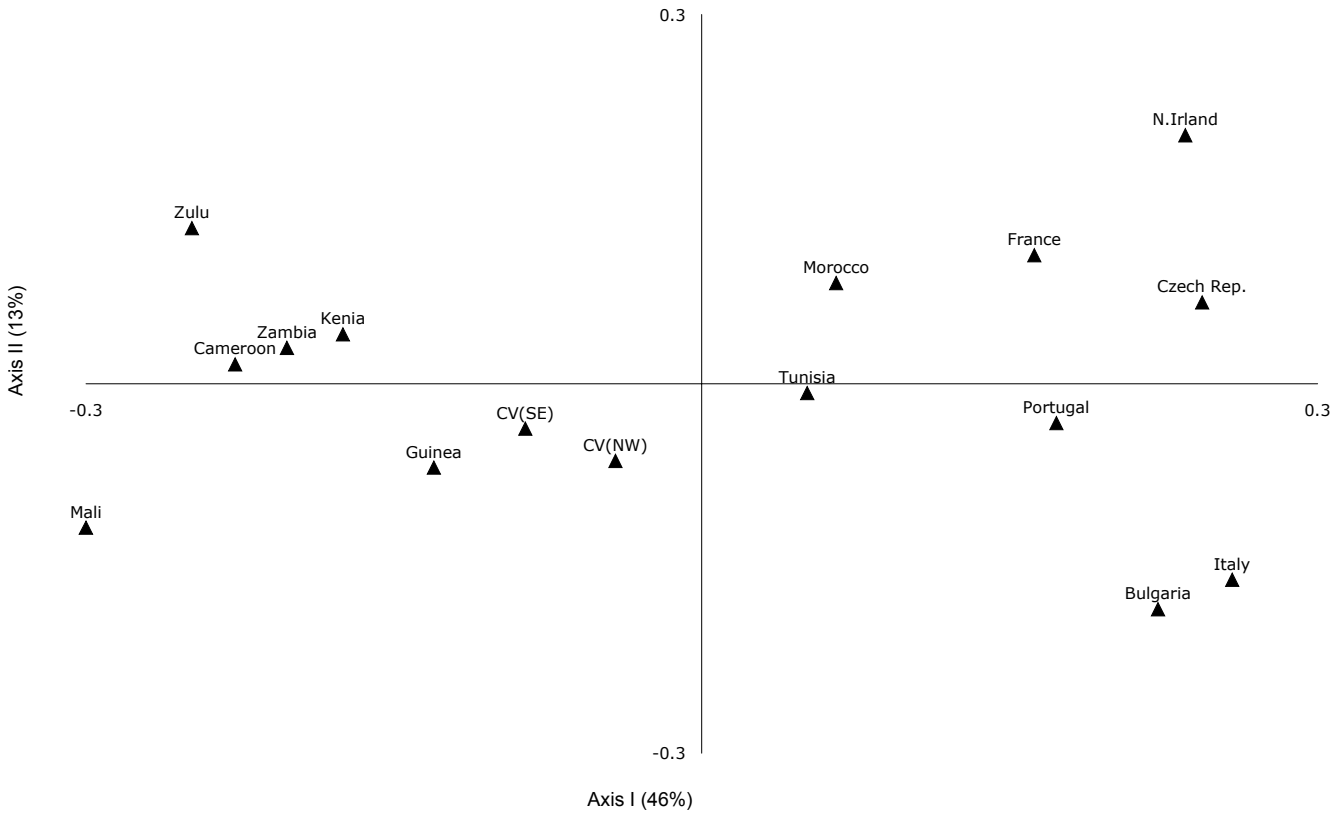

FIGURE 3 Principal coordinates analysis using the human leukocyte antigen (HLA)-A and HLA-B allele frequencies. CV(NW): Cabo Verde northwestern; CV(SE): Cabo Verde southeastern.

haplotypes are common to Guiné-Bissau, and no one Guiné-Bissau haplotype is common to Portugal (data not shown).

Phylogenetic Analyses

Phylogenetic trees constructed with HLA-A, -B, and -DRB1 or just with class I (HLA-A and -B) allele frequencies (Figure 2) show a close relationship between Guiné-Bissau and Cabo Verde, which appear clustered with other sub-Saharan populations. A principal coordinate analysis (Figure 3) is consistent with the dendrograms, but shows also that Northwestern Cabo Verde islands are not so distant from North Africans and Europeans than southeastern islands.

DISCUSSION

Guiné-Bissau was strongly influenced by different historic and prehistoric events that bring together people from different origins in Africa. In different times during several thousands of years, this area on the West Coast of Africa received people from North Africa, the Sahara, and East Africa. The coast of Guinea, known as Senegambia, was the origin of most of the slave trade in the 15th and 16th centuries in which Portugal was involved. African people were captured from the coast of Guinea and taken to Cabo Verde Islands, which served as a central outpost in the slave trade to Europe and America continents [11]. Cabo Verde archipelago, after its discovery by Portuguese in 15th century, was populated with slaves from the West Coast of Africa and some Caucasians, mostly from Iberia.

The present work describes for the first time HLA-A, -B, and -DRB1 high-resolution allele and haplotype frequencies in Guiné-Bissau and Cabo Verde populations. This high-resolution throughput shows that several HLA generic groups identified before in Cabo Verde [18] were in fact constituted by a variety of alleles (Table 2). This additional information is highly relevant for a detailed HLA structure characterization. Unfortunately, few African populations have been well characterized to HLA allele level, a situation that renders difficult the comparison of the available data sets.

Guiné-Bissau and Cabo Verde position in dendrograms (Figure 2) reflects clearly their sub-Saharan origin. However, principal coordinates analysis (Figure 3) shows that these two populations are not so distant from North Africans as other sub-Saharans. In Guiné-Bissau, this situation could result from ancient genetic contacts with

other peoples, namely North Africans before Sahara desertification and East Africa Arabs through the Sahel corridor [6]. More recent influences such as admixture with Berbers from Muslim pressure in Northwestern Africa in the ninth century [10] are possible. A previous mtDNA study in Guiné-Bissau samples showed 6% of its haplogroups characteristic of North Africa, East Africa, Arabia, the Middle East, and even Europe [17].

The observed similarity between Guiné-Bissau and Cabo Verde is consistent with the well-documented origin of the archipelago's first settlers. Despite the prevalent sub-Saharan influence in Cabo Verde, we found many more haplotypes common to Portugal (9.3%) than to Guiné-Bissau (2.2%), denoting a Caucasian contribution to the Cabo Verde settlement.

Principal coordinate analysis (Figure 3) shows that the northwestern group of islands is more similar to Caucasians than southeastern Cabo Verde islands, in agreement with previous studies based on STR and Y-chromosome typings [13, 15]. In fact, overall HLA allele frequencies are significantly different between these two groups ($p = 0.007$), and each one shows a specific haplotype profile (Table 3). Southeastern islands have also more common haplotypes to Guiné-Bissau (3%) and Portugal (10%) than do the northwestern islands (1% common to Guiné-Bissau and 7% to Portugal). This reduced number of common haplotypes to Guiné-Bissau could mean that Cabo Verde sub-Saharan settlers do not belong to a single and homogeneous group. The lower number of common haplotypes to Guiné-Bissau and Portugal in the northwestern islands is most probably the result of a founder effect, because these islands were populated later from slaves that fled from the southeastern group. Further studies on West African coast populations should clarify the origin of sub-Saharan settlers in the Cabo Verde Islands.

REFERENCES

1. Riley E, Olerup O, Arnaiz-Villena A, Benmamar D, Alvarez M, Diaz-Campos N, Varela P, Gomez-Casado E, Martinez-Lazo J. HLA polymorphism and evolution, Immunol Today 13:333, 1992.
2. Arnaiz-Villena A, Benmamar D, Alvarez M, et al. HLA allele and haplotype frequencies in Algerians. Relatedness to Spaniards and Basques. Hum Immunol 43:259, 1995.
3. Arnaiz-Villena A, Iliakis P, Gonzáles-Hevilla M, Langas J, Gomez-Casado E, Slyridaki K, Trapaza S, Silnero-Redanda C, Matsauka C, Martinez-Jasa J. The origin of Creatan populations as determined by characterization of HLA alleles. Tissue Antigens 53:213, 1999.
4. Clayton J, Lonjou C, Whittle D: Allele and haplotype frequencies for HLA loci in various ethnic groups. In Charron D (eds): Genetic Diversity of HLA. Functional and Medical Implications. Paris, EDK, 1997.
5. Cao K, Moormann AM, Lyke KE, Masaberg C, Sumba OP, Daumba OK, Kaech D, Lancaster A, Nelson M, Meyer D, Single R, Hartzman RJ, Plawe CV, Kazura S, Mann DL, Sztein MB, Thamsan G, Fernandez-Vima MA. Differentiation between African populations is evidenced by the diversity of alleles and haplotypes of HLA class I loci. Tissue Antigens 63:293, 2004.
6. Ellis JM, Mack SJ, Leke RFG, Quakyi I, Johnson AH, Hurley CK: Diversity is demonstrated in class I HLA-A and HLA-B alleles in Cameroon, Africa: description of HLA-A*03012, *2612, *3006 and HLA-B*1403, *4016, *4703. Tissue Antigens 56:291, 2000.
7. Pimtanothai N, Hurley C K, Leke R, Klitz W, Johnson AH: HLA-DR and -DQ polymorphism in Cameroon. Tissue Antigens 58:1, 2001.
8. Alimen H: Evolution du climat et des civilisations depuis 40000 ans du nord au sud du Sahara occidental. (Premiéres conceptions confrontées aux données récentes). Bull L'Assoc Franç L'Etude Quaternaire 4215, 1987.
9. Dutour O, Vernet R, Aumassip G: Le peuplement pré historique du Sahara. In Aumassip G, et al. (eds): Milieux, hommes et techniques du Sahara pré historique. Problémes actuels. Paris, L'Harmattan, 1988.
10. Moreira JM: Os Fulas da Guiné Portuguesa na panorâmica geral do mundo Fula. Boletim Cultural Guiné Portuguesa 19:417, 1964.
11. Russel-Wood AJ: The Portuguese Empire. 1415-1808. A World on the Move. Baltimore, Johns Hopkins University Press, 1998.
12. Carreira A: Migrações nas Ilhas de Cabo Verde. 2nd ed. Lisboa, Instituto Caboverdeano do Livro, 1983.
13. Gonçalves R, Rosa A, Freitas A, Fernandez A, Kirisild T, Villein R, Brehm A. Y-Chromosome lineages in Cabo Verde Islands witness the diverse geographic origin of its first male settlers. Human Genet 113:467, 2003.
14. Brehm A, Pereira L, Bandelt HJ, Prata MJ, Amorim A: Mitochondrial portrait of the Cabo Verde archipelago: the Senegambian outpost of Atlantic slave trade. Ann Hum Genet 66:49, 2002.
15. Fernandes AT, Velosa R, Jesus J, Carracedo A, Brehm A: Genetic differentiation of the Cabo Verde archipelago population analysed by STR polymorphisms. Ann Hum Genet 67:340, 2003.
16. Gonçalves R, Jesus J, Fernades AT, Brehm A: Genetic profile of a multi-ethnic population from Guiné-Bissau (west African coast) using the new PowerPlex 16 System kit. Forensic Sci Int 129:78, 2002.
17. Rosa A, Brehm A, Kivisild T, Metspalu E, Villems R: MtDNA profile of west Africa Guineans: towards a better understanding of the Senegambia region. Ann Hum Gen 68:340, 2004.
18. Spínola H, Brehm A, Williams F, Jesus J, Middleton D: Distribution of HLA alleles in Portugal and Cabo Verde.

Relationships with the slave trade route. Ann Hum Genet 66:285, 2002.

19. Kurtz B, Steiert I, Heuchert G, Müller CA: New high resolution typing strategy for HLA-A locus alleles based on dye terminator sequencing of haplotypic group-specific PCR-amplicons of exon 2 and exon 3. Tissue Antigens 53:81, 1999.
20. Pozzi S, Longo A, Ferrara GB: HLA-B locus sequence-based typing. Tissue Antigens 53:275, 1999.
21. Wiliams F, Meenagh A, Single R, McNally M, Kelly P, Nelson MP, Meyer D, Lancaster A, Thamsan G, Middleton D. High resolution HLA-DRB1 identification of a Caucasian population. Human Immunol 65:66, 2004.
22. Schneider S, Kueffer J M, Roessli D, Excoffier L. Arlequin v2.000: a software for population genetic data analysis. Geneva, Switzerland: Genetics and Biometry Laboratory, University of Geneva, 2000.
23. Weir BS: Genetic data analysis II. Sunderland, Sinauer Associates, MA, Massachusetts 1996.
24. Middleton D, Manavalan JS, Fernandez-Vina MA: Human Immunol 65:855, 2004.
25. Ayed K, Ayed-Jendoubi S, Sfar I, Labonne MP, Gebuhrer L: HLA class-I and HLA class-II phenotypic, gene and haplotypic frequencies in Tunisians by using molecular typing data. Tissue Antigens 64:520, 2004.
26. Sanchez-Velasco P, Gomez-Casado E, Martnez-Lazo J, Moscoso J, Zamora S, Lowy E, Silviera C, Cambarain A, Leyna-Cabian F, Arnaiz-Villena. HLA alleles in isolated populations from North Spain: origin of the Basques and ancient Iberians. Tissue Antigens 66:26, 2005.
27. Spínola H, Middleton D, Brehm A: HLA genes in Portugal inferred from sequence-based typing: in the crossroad between Europe and Africa. Tissue Antigens 66:26, 2005.
28. Excoffier L, Smouse PE, Quattro JM: Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479, 1992.
29. Felsenstein J: PHYLIP (Phylogeny Inference Package) version 3.6. Department of Genome Sciences, University of Washington, Seattle, 2004.
30. Page RDM: TREEVIEW: an application to display phylogenetic trees on personal computers. Comp Appl Biosci 12:357, 1996.