

## BRIEF COMMUNICATION

**HLA-A polymorphisms in four ethnic groups from Guinea-Bissau (West Africa) inferred from sequence-based typing**C. Spínola<sup>1</sup>, J. Bruges-Armas<sup>2,3</sup>, A. Brehm<sup>1</sup> & H. Spínola<sup>1</sup><sup>1</sup> Human Genetics Laboratory, University of Madeira, Funchal, Portugal<sup>2</sup> Hospital Santo Espírito de Angra do Heroísmo, SEEBMO, Azores, Portugal<sup>3</sup> Institute for Molecular and Cell Biology (IBMC), Porto, Portugal**Key words**

Guinea-Bissau; HLA-A; sequence-based typing; sub-Saharan populations

**Correspondence**

Hélder Spínola  
 Human Genetics Laboratory  
 University of Madeira  
 Campus da Penteada  
 9000-390 Funchal  
 Portugal  
 Tel: +351 29 170 5390  
 Fax: +351 29 170 5329  
 e-mail: hspinola@uma.pt

Received 10 April 2008; revised 7 July 2008,  
 1 August 2008, 20 August 2008; accepted 28  
 August 2008

doi: 10.1111/j.1399-0039.2008.01147.x

**Abstract**

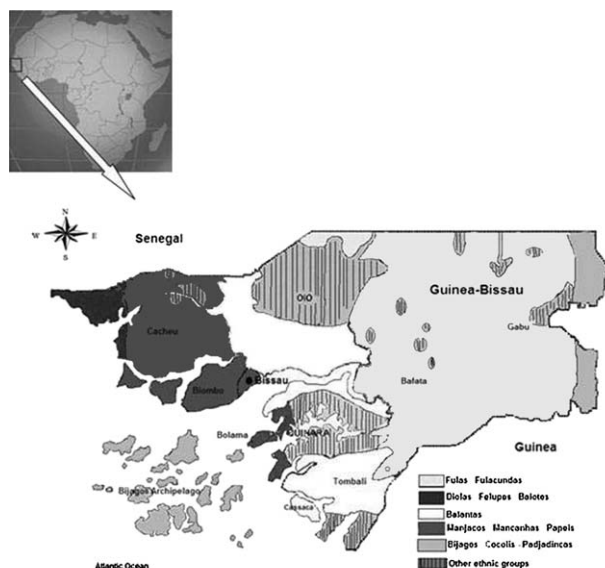
Human leukocyte antigen (HLA)-A locus polymorphisms were examined at high-resolution level, using sequence-based typing, in the four most representative Guinea-Bissau (Northwest Africa) ethnic groups: Balanta, Bijagós, Fula and Papel. Despite the Fula group having significant differences when compared with the other three ethnic groups, all four groups most likely received a genetic input from non sub-Saharan populations. The Bijagós and Papel groups showed similarities to neighboring populations from Mali and Senegal. The Balanta, despite their oral tradition of an East Africa origin, show affinities to Cameroon populations, highly influenced by Bantu migrations. These results are congruent with historical sources and other genetic studies that support the finding that the Guinea-Bissau genetic pool was influenced by several migrations from North Africa, Sahara and East Africa.

The human leukocyte antigen (HLA) system, a group of highly polymorphic loci, has been mapped to the short arm of human chromosome 6 (6p21.3), spanning over 4 Mb of DNA (1–3). Some of these HLA loci (e.g. HLA-A, HLA-B and HLA-DRB1) have been used in population studies to assess for gene flow on the basis of allele frequencies and genetic distances (4–6).

Human settlement in the northwestern coast of Africa is the result of a continuous complex network of migrations, invasions and admixture of people from different origins that began around 40,000 years before present (YBP) (7). Before the Sahara desertification took place (around 9000 YBP), several Neolithic cultures flourished in the area, bringing together people of sub-Saharan and North African origin (8). Around 4000 YBP, the delta known today as Guinea-Bissau was a region influenced by continuous migrations of different ethnic groups coming from the East Coast of Africa through the Sahel Corridor (9), which by its turn brought genetic inputs from Asia and

Middle East. Studies on Y chromosome haplotypes from Cameroon populations show evidences of genetic influence from Asians, supporting the hypothesis of an Asia to Africa migration (10). Most recently, in the 9th and 11th centuries, the region was influenced by Berber admixtures because of Muslim pressure from Northwestern Africa (11).

Guinea-Bissau is a small country (Figure 1) (12) with approximately 1.5 million inhabitants subdivided into 30 different ethnic groups in which Balanta (~500,000 people), Bijagós (~30,000 people), Fula (~300,000 people) and Papel (~140,000 people) represent the major groups (13). The origin of these ethnic groups is not clear despite research using anthropological comparisons and genetic markers. The Balanta, probably of Ethiopian or Sudanese origin (14), are the largest group, and in the first quarter of the 20th century spread over territories occupied earlier by other ethnic groups (15). Quintino refers language affinities between Balanta and Sudanese, from which they could have separated 2000 years ago with the first spread of Kushites migrations (15). There is



**Figure 1** Ethnic groups distribution in Guinea-Bissau [modified from McEvedy (9)].

also a cultural and physical resemblance with Bantu suggesting a common origin at the end of Pleistocene (16). Recent studies on Y chromosome suggest that Balanta may have experienced a particular expansion benefit through food production, which associates their ancestors with the people who implemented the farming habits in the Guinea-Bissau area (17).

The Bijagós inhabit the Archipelago of Bijagós and are referred to have cultural traditions resembling Egyptians but also related to the Senegalese Djola, a group believed to be originated from Sudan in the 15th and 16th centuries. These groups also include the Beafada, which have an oral tradition of coming from Mali (15, 16, 18). Teixeira da Mota (19) considers the Bijagós as a separated branch of Djola or at least very closely related to Papel and Nalú, the latter referred as the autochthonous people of the region.

The Papel live on the coast having Bissau Island as their main territory (13) and live from agriculture (14). A Y chromosome study suggests that Papel retains a conserved ancestral pool deriving from East Africa, from where they have supposedly migrated, but this same study clearly distinguishes them from Bijagós (17).

A mass arrival of Fula people, nomad pastoralists that spread in West Africa coming from Futa Toro (Senegal River) and Sahel regions, took place in Guinea-Bissau in the beginning of the 19th century (20). The origin of this ethnic group in unknown, but tradition relates them to Hiksos and Nubians from Near East and East Africa, respectively (18).

Recent studies on the Guinea-Bissau population with different molecular markers (HLA loci, mtDNA and Y chromosome) show a great genotypic diversity common to

other African populations, showing a sub-Saharan origin with clear influences from North and East Africa (17, 21–24). Unfortunately, the HLA-A analysis carried out previously in Guinea-Bissau (24) missed a typing analysis separated by ethnic groups.

The present study consisted of a total of 127 healthy unrelated males from different ethnic groups from Guinea-Bissau. Blood samples were collected after informed consent from donors whose parents and grandparents belong to the same ethnic group. Samples were identified as belonging to four different groups: Balanta ( $n = 48$ ), Bijagós ( $n = 23$ ), Fula ( $n = 31$ ) and Papel ( $n = 25$ ). Genomic DNA was isolated from whole blood containing ethylenediaminetetraacetic acid using a salting-out procedure according to Miller *et al.* (25), with some modifications. All subjects were high-resolution sequence based typed for HLA-A locus as previously described (26), with minor modifications, and processed using the SEQPILOT version 3.0 (JSI medical systems GmbH, Frankfurt, Germany). Heterozygous ambiguities were resolved by polymerase chain reaction-sequence specific oligonucleotide probes (PCR-SSOP) as previously described (27), except for 21 samples in which for allele frequencies calculation, we used the combination of the most frequent single alleles as performed previously [a list of these results with more than one heterozygous allele combination available from Spínola *et al.* (4, 24)]. A few groups of alleles could not be distinguished because we did not test polymorphisms outside exons 2 and 3. A summary of this unresolved ambiguities and the reference used in each case in the text for frequency analysis are given in Table 1.

An analysis of molecular variance between the different ethnic groups, performed with ARLEQUIN v3.0 (28) and based on Euclidean distances (29), shows that only 1.58% of the total genetic variation found can be attributed to differences among them (overall  $F_{st} = 0.016$ ,  $P = 0.01$ ), a value higher than that obtained between Cape Verde and Guinea-Bissau in a previous study (0.71%) (24). The exact

**Table 1** Ambiguous HLA-A alleles and their respective assignment labels under which they appear in the text

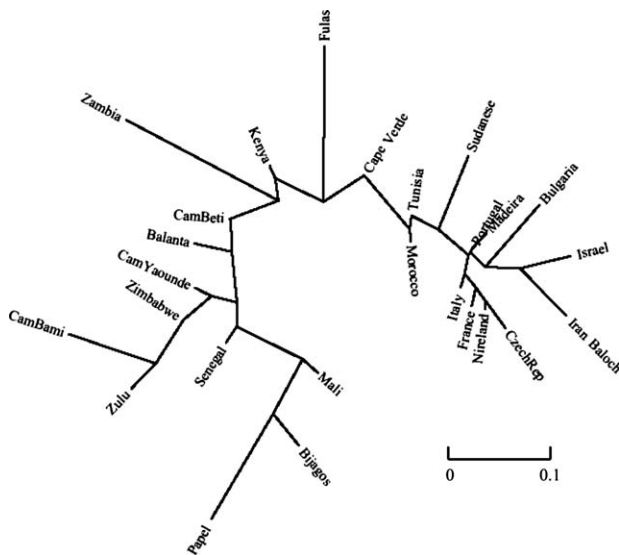
Ambiguous alleles	Reference
A*0101/A*0104N/0122N	A*0101
A*0201/A*0209/A*0243N/A*0266/A*0275/ A*0283N/A*0289/A*0297/A*9232/A*9234/A*9240	A*0201
A*0206/A*9226	A*0206
A*0301/A*0320/A*0321N/A*0326/A*0337	A*0301
A*2301/A*2307N/A*2317/A*2318	A*2301
A*2402/A*2409N/A*2411N/A*2440N/A*2476/ A*2479/A*2483N	A*2402
A*2601/A*2624/A*2626	A*2601
A*3001/A*3024	A*3001
A*3303/A*3315	A*3303
A*7401/A*7402	A*7401

test of population differentiation, performed by ARLEQUIN, shows significant results between Fula and all other ethnic groups (Bijagós,  $P = 0$ ; Balanta,  $P = 0.012$  and Papel,  $P = 0.013$ ).

The phylogenetic tree (Figure 2) constructed with HLA-A allele frequencies (Table 2) shows a close relationship between Bijagós and Papel ethnic groups, which appear clustered with other sub-Saharan populations. The Balanta and Fula cluster together with sub-Saharans, but the latter shows a clear tendency to group with Sudanese and North African populations. The principal coordinate analysis (PCO) (Figure 3) is consistent with the dendrogram, plotting Bijagós close to Papel, and Balanta and Fula not far from each other and from other sub-Saharans. Previous studies on HLA-A locus in Guinea-Bissau had shown similar results in allele frequencies, except for A\*0301, with 6.2% and 3.6%, and A\*6802, with 2.3% and 4.6%, in the previous and present study, respectively (24). In this study, we consider separately the four most representative ethnic groups and include a higher number of samples (127 vs 65). This is probably the reason why some rare alleles were

**Table 2** HLA-A allele frequencies in Balanta, Bijagós, Fula and Papel ethnic groups from Guinea-Bissau population

Frequencies	Guinea-Bissau ethnic groups			
	Balanta (n = 48)	Bijagós (n = 23)	Fula (n = 31)	Papel (n = 25)
HLA-A				
A*0101	0.021	0.022	0.081	0.08
A*0102	0.021	0	0	0.02
A*0201	0.135	0.087	0.113	0.04
A*0202	0.063	0.087	0.065	0.08
A*0205	0.01	0.022	0	0
A*0206	0	0.022	0	0
A*0220	0.01	0	0	0
A*0301	0.021	0.065	0.016	0.02
A*2301	0.166	0.196	0.211	0.2
A*2402	0	0	0.097	0
A*2431	0.01	0	0	0
A*2601	0	0	0.032	0.1
A*2902	0.042	0	0.016	0
A*3001	0.021	0.152	0.016	0.02
A*3002	0.021	0	0.016	0.02
A*3009	0	0	0.016	0
A*3201	0.052	0	0.032	0
A*3301	0.021	0.042	0.032	0.02
A*3303	0.083	0.065	0.048	0.16
A*3402	0.094	0.022	0.032	0.02
A*6601	0	0	0.016	0
A*6603	0.01	0	0	0
A*6801	0.042	0.022	0.032	0.02
A*6802	0.042	0	0.129	0.04
A*7401	0.094	0.196	0	0.16
A*8001	0.021	0	0	0

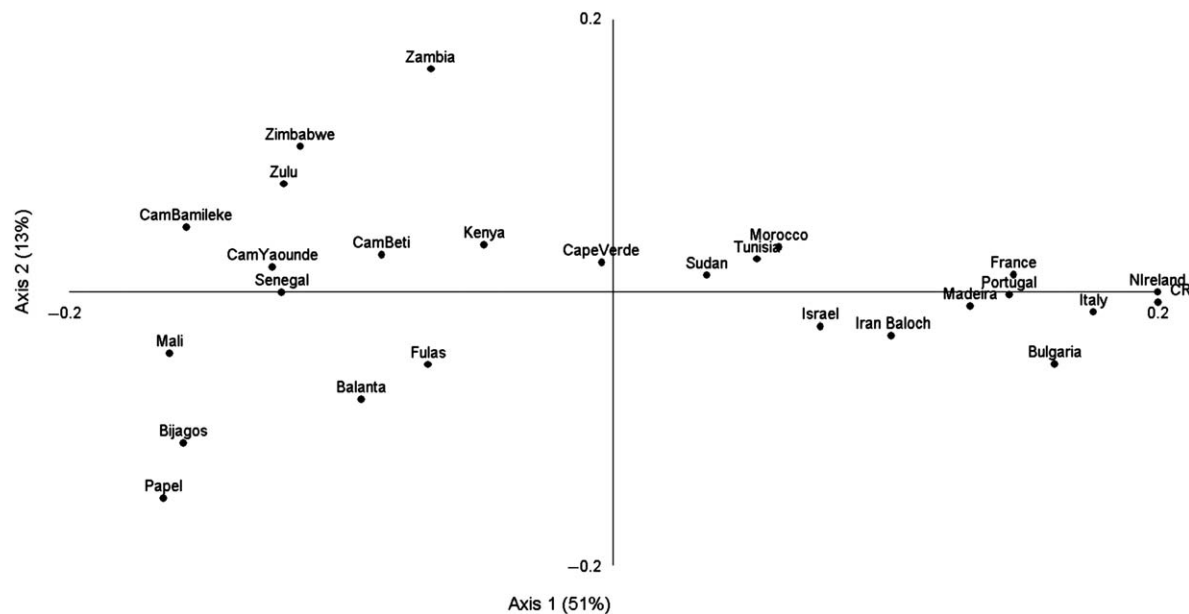


**Figure 2** Neighbor-joining (NJ) dendrogram showing the comparative position of Guinea-Bissau ethnic groups (Balanta, Bijagós, Fula and Papel) with other populations available from the literature: Kenya (5), Mali (5), Zambia (5), Cape Verde (24), Guinea-Bissau (24), Portugal (32), Madeira Island (33) and, from www.allelefrequencies.net, Italy, France, Northern Ireland (Nireland), Czech Republic (CzechRep), Bulgaria, Sudan, Cameroon Beti (CamBeti), Cameroon Yaoundé (CamYaounde), Cameroon Bamileke (CamBami), Zimbabwe, Zulu, Senegal, Tunisia, Morocco, Iran Baloch and Israel. Standard genetic distances among populations were calculated using human leukocyte antigen-A allele frequencies with the software included in the PHYLIP v.3.6 package (34). With software PHYLIP package, and using Nei genetic distances, a distance matrix was generated using GENDIST and used as input to NEIGHBOR to produce a NJ tree. The topology was visualized with DRAWTREE (34, 35).

detected that were absent in the previous work (A\*0205, A\*0206, A\*0220, A\*2431, A\*3009 and A\*6603).

While the total Guinea-Bissau whole population was not in Hardy–Weinberg equilibrium, when we considered each ethnic group separately the equilibrium is present (Table 3). This could be because of cultural and geographic isolation and of the absence of casual mating between ethnic groups, making each one a different population. In opposition to what was found in Cape Verde or Cameroon populations (24, 30), Ewens–Watterson neutrality test in all studied Guinea-Bissau ethnic groups showed observed homozygosity similar to expected (Table 3), presenting no evidence for balancing selection on HLA-A.

In Guinea-Bissau, the correlation between alleles A\*6801 and A\*6802 was not the same for all the four ethnic groups. Fula and Papel present the allele A\*6802 more frequently than A\*6801, a typical situation in sub-Saharans, but the opposite was found in Bijagós (Table 2). The absence of A\*6802 in the Bijagós ethnic group might be a size sample consequence. However, the clear prevalence of A\*6801 may be explained by the existence of genetic influences from



**Figure 3** Principal coordinate analysis (PCO) using human leukocyte antigen-A allele frequencies carried out on the MultiVariate Statistical Package MVSP3 (36). Czech Republic (CR), Northern Ireland (Nireland), Cameroon Bamileke (CamBamileke), Cameroon Yaoundé (CamYaounde), Cameroon Beti (CamBeti). For reference of the populations used, see Figure 2.

North Africa, through Mali, or as a result of a bottleneck or founder effects. The similar frequency of A\*6801 and A\*6802 in Balanta is most likely the result of genetic influences from non-sub-Saharan populations from North and East Africa.

Balanta are considered to be of Ethiopian or Sudanese origin (14), which is also supported through the presence of mtDNA markers from East Africa (22). While PCO (Figure 3) suggests that this ethnic group have similarities to Fula, another ethnic group with a possible origin of East Africa and Near East (18), Balanta shows affinities to Cameroon populations (Figure 2), highly influenced by

Bantu migrations (30). According to our data, even considering an East Africa origin for the Balanta, its genetic pool today indicates a considerable influence from other populations. Indeed, allele A\*3402 (9.4%) shows a frequency higher than that in all African and Caucasian populations studied (31), and alleles A\*6801 and A\*6802 present the same frequency (4.2% each) in opposition to other sub-Saharans that have a higher prevalence of the latter (5).

Dendrogram analyses (Figure 2) cluster Fula between sub-Saharans, and Sudanese and North Africans, a position also supported by allele A\*2402 with a frequency (9.7%) similar to Asia and Middle East populations and higher than the observed for sub-Saharans (5, 31). The significant differentiation between Fula and the other ethnic groups, besides the influence of their specific origin, could also be a consequence of a lower gene flow because of its most recent arrival in Guinea-Bissau (20). Bijagós and Papel ethnic groups show a similar behavior in phylogenetic analyses and PCO. In the phylogenetic tree, these two ethnic groups cluster together with Mali and Senegal populations in a large group with other sub-Saharans. The PCO plots the Papel and Bijagós close to Mali and not far from Senegalese. These data are consistent with the hypothesis that Bijagós are related with Senegal and Mali populations (18). The Bijagós and Papel groups present high frequencies of HLA-A\*7401 (19.6% and 16%, respectively), a typical allele from sub-Saharans found with frequencies ranging from 2% to 9% (5, 31). These high frequencies suggest the relationship between both ethnic groups showed in the dendrogram and PCO and could refer to a common ancestral population or

**Table 3** Hardy–Weinberg equilibrium (HWE) and Ewens–Watterson neutrality test<sup>a</sup> in Balanta, Bijagós, Fula and Papel ethnic groups from Guinea-Bissau population estimated with ARLEQUIN v3.0 (28)

Ethnic groups	HWE			Ewens–Watterson neutrality test				
	Obs. Het.	Exp. Het.	P-value	HWE	Fo	Fe	P-value	Selection
Balanta	0.96	0.94	0.14	Yes	0.086	0.115	0.144	No
Bijagós	0.83	0.91	0.14	Yes	0.128	0.159	0.221	No
Fulas	0.84	0.92	0.09	Yes	0.1	0.11	0.367	No
Papel	0.84	0.93	0.07	Yes	0.12	0.135	0.399	No

Obs. Het., observed heterozygosity; Exp. Het., expected heterozygosity; Fo, observed homozygosity; Fe, expected homozygosity.

<sup>a</sup> In the present study, Ewens–Watterson neutrality test was applied to examine the presence of selective forces influencing allelic diversity at this locus.

to gene flow between them. The 10% of HLA-A\*2601 found in the Papel group, an allele absent in the other three Guinea-Bissau ethnic groups and in much lower frequencies in other sub-Saharanans, can be the result of some gene flow from Middle East or Asia where this allele reach such high frequencies (31). Cruciani et al. (10) found Y chromosome haplotypes in Cameroons that are common to Asians, supporting the hypothesis of an Asia to Africa migration. In our data, the frequencies of HLA-A\*2601 (10%) in Papel and HLA-A\*2402 (9.7%) in Fula corroborate the hypothesis because these are rare alleles in sub-Saharanans but frequent in Asian and Middle East populations (31). The absence of A\*7401 in the Fula group, a typical sub-Saharan allele, also suggests a gene flow from outside sub-Saharan Africa. In conclusion, the genetic diversity found in these Guinea-Bissau ethnic groups, while small in numbers, lends support for a main sub-Saharan origin, as well as showing some specific characteristics that suggest different genetic influences along time.

### Acknowledgments

This research was supported by Portuguese Foundation for Science and Technology and European Community through the project number POCI/BIA-BCM/60440/2004. BIOFORMA and SANTOQUEIJO are sponsors of the University of Madeira Human Genetics Laboratory.

### References

1. Bodmer WF. The HLA system: structure and function. *J Clin Pathol* 1987; **40**: 948–58.
2. Klein J, Sato A. The HLA system. First of two parts. *N Engl J Med* 2000; **343**: 702–9.
3. Naik S. The human HLA system. *J Indian Rheumatol Assoc* 2003; **11**: 79–83.
4. Spínola H, Brehm A, Bettencourt B, Middleton D, Bruges-Armas J. HLA class I and II polymorphisms in Azores show different settlements in Oriental and Central Islands. *Tissue Antigens* 2005; **66**: 217–30.
5. Cao K, Moormann AM, Lyke KE et al. Differentiation between African populations is evidenced by the diversity of alleles and haplotypes of HLA class I loci. *Tissue Antigens* 2004; **63**: 293–325.
6. 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* 2002; **66**: 285–96.
7. Alimen H. Evolution du climat et des civilisations depuis 40000 ans du nord au sud du Sahara occidental. (Premières conceptions confrontées aux données récentes). *Bull L'Assoc Franç L'Etude Quaternaire* 1987; **4**: 215–27.
8. Dutour O, Vernet R, Aumassip G. Le peuplement pré historique du Sahara. In: Aumassip G, Ferhat N, Heddouche A, eds. Milieux, hommes et techniques du Sahara pré historique. Problèmes actuels. Paris: L'Harmattan, 1988, 39–52.
9. McEvedy C. *The Penguin Atlas of African History*, 2nd edn. New York: Penguin Books, 1995, 39–52.
10. Cruciani E, Santolmaza P, Shen P et al. A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *Am J Hum Genet* 2002; **70**: 1197–214.
11. Moreira JM. Os Fulas da Guiné Portuguesa na panorâmica geral do mundo Fula. *BCG (Guiné)* 1964; **19**: 417–32.
12. *United Nations, Office for the Coordination of Humanitarian Affairs*. Guinea-Bissau: Briefing Pack, 2005.
13. Nóbrega A. A luta pelo poder na Guiné-Bissau. Lisboa: Instituto Superior de Ciências Sociais e Políticas, Universidade Técnica de Lisboa, 2003.
14. Alves C. Guiné Portuguesa-Geografia. In: Editora Verbo, S.A.R.L., ed. *Enciclopédia Luso-Brasileira de Cultura*, Vol. 9. Sintra: Mem-Martins, 1976, 1348–9.
15. Quintino F. Sobrevivências da cultura etiópica no ocidente africano. *BCG (Guiné)* 1964; **19**: 5–35.
16. Quintino F. Os povos da Guiné: estrutura social. *BCG (Guiné)* 1969; **XXIV**: 861–915.
17. Rosa A, Ornelas C, Jobling MA, Brehm A, Villems R. Y-Chromosomal diversity in the population of Guinea-Bissau: multiethnic perspective. *BMC Evol Biol* 2007; **7**: 124.
18. Lopes C. Kaabunké, espaço, território e poder na Guiné-Bissau, Gambia e Caramance pré-coloniais. Lisboa: CND, 1999.
19. Teixeira da Mota A. *Guiné Portuguesa*. Lisboa: Agência Geral do Ultramar, 1954.
20. Carreira A, Quintino FR. *Antroponímia da Guiné Portuguesa*, Vol. 49. Lisboa: 52 JIU, 1964.
21. Gonçalves R, Jesus J, Fernandes 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* 2002; **129**: 78–80.
22. 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 Genet* 2004; **68**: 340–52.
23. Rosa A, Ornelas C, Brehm A, Villems R. Population data on 11 Y-chromosome STRs from Guiné-Bissau. *Forensic Sci Int* 2006; **157**: 210–7.
24. Spínola H, Bruges-Armas J, Middleton D, Brehm A. HLA polymorphisms in Cabo Verde and Guiné-Bissau inferred from sequence-based typing. *Hum Immunol* 2005; **66**: 1082–92.
25. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; **16**: 1215.
26. Kurz 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* 1999; **53**: 81–96.
27. Middleton D. PCR-SSOP class I and class II (DRB1). In Land GA, Strothman RM, Hahn AB, eds. *ASHI Laboratory Manual*, 4th edn. Lenexa: American Society for Histocompatibility and Immunogenetics, 2000, 1–23.

28. Excoffier LG, Laval G, Schneider S. Arlequin version 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform Online* 2005; **1**: 47–50.
29. 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* 1992; **131**: 479–91.
30. 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* 2000; **56**: 291–302.
31. Middleton D, Menchaca E, Rood H, Komerofsky H. New allele frequency database: <http://www.allelefrequencys.net>. *Tissue Antigens* 2003; **61**: 403–7.
32. Spínola H, Middleton D, Brehm A. HLA genes in Portugal inferred from sequence-based typing: in the crossroad between Europe and Africa. *Tissue Antigens* 2005; **66**: 26–36.
33. Spínola H, Bruges-Armas J, Mora MG, Middleton D, Brehm A. HLA genes in Madeira Island (Portugal) inferred from sequence-based typing: footprints from different origins. *Hum Immunol* 2006; **43**: 1726–8.
34. Felsenstein J. *PHYLIP (Phylogeny Inference Package) version 3.6*. Distributed by the author. Seattle: Department of Genome Sciences, University of Washington, 2004.
35. Dereeper A, Guignon V, Blanc G *et al.* Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 2008; **36**: 465–9.
36. Kovach WL. *MVSP – A Multivariate Statistical Package for Windows, version 3.6*. Pentraeth: Kovach Computing Services, 2006.