

Rita Gonçalves · Alexandra Rosa · Ana Freitas
Ana Fernandes · Toomas Kivisild · Richard Villems ·
António Brehm

Y-chromosome lineages in Cabo Verde Islands witness the diverse geographic origin of its first male settlers

Received: 14 May 2003 / Accepted: 16 July 2003 / Published online: 26 August 2003

© Springer-Verlag 2003

Abstract The Y-chromosome haplogroup composition of the population of the Cabo Verde Archipelago was profiled by using 32 single-nucleotide polymorphism markers and compared with potential source populations from Iberia, west Africa, and the Middle East. According to the traditional view, the major proportion of the founding population of Cabo Verde was of west African ancestry with the addition of a minor fraction of male colonizers from Europe. Unexpectedly, more than half of the paternal lineages (53.5%) of Cabo Verdeans clustered in haplogroups I, J, K, and R1, which are characteristic of populations of Europe and the Middle East, while being absent in the probable west African source population of Guiné-Bissau. Moreover, a high frequency of J* lineages in Cabo Verdeans relates them more closely to populations of the Middle East and probably provides the first genetic evidence of the legacy of the Jews. In addition, the considerable proportion (20.5%) of E3b(xM81) lineages indicates a possible gene flow from the Middle East or northeast Africa, which, at least partly, could be ascribed to the Sephardic Jews. In contrast to the predominance of west African mitochondrial DNA haplotypes in their maternal gene pool, the major west African Y-chromosome lineage E3a was observed only at a frequency of 15.9%. Overall, these results indicate that gene flow from multiple sources and various sex-specific patterns have been important in the formation of the genomic diversity in the Cabo Verde islands.

Introduction

Mitochondrial polymorphisms have been extensively used to study the maternal composition and relationships of human populations and past migrational events. Recently, Y-chromosomal markers have acquired a special interest because they provide an independent and paternal historic view of those relationships (Cruciani et al. 2002; Hurles et al. 1999; Malaspina et al. 2001; Scozzari et al. 1999; Semino et al. 2000; Underhill et al. 2000, 2001). In Europe, Y-chromosome structure has been shown to be strongly predetermined by the geographic location of the populations, irrespective of their linguistic affiliation (Rosser et al. 2000). At a more general resolution, Y-chromosome variation can be grouped into continent-specific haplogroups or lineages (Y Chromosome Consortium 2002). The existence of a geographic correlation of specific Y-chromosome polymorphisms is helpful in uncovering past historic events and in explaining the genetic composition of extant populations. The Archipelago of Cabo Verde presents an interesting area for studying the degree of admixture of populations from genetically different backgrounds (Europe and Africa) by using Y-chromosome bi-allelic polymorphisms.

The Cabo Verde Archipelago was inhabited at the time of its discovery in the 15th century. The first settlers were mostly European males recruited from Portuguese nobles, Genovese adventurers, exiles, and convicts from the Crown. In 1497, following a royal edit, all non-converted Jews were expelled from Portugal and exiled in Madeira and Cabo Verde. The first islands to be settled were Santiago and nearby Fogo (Carreira 1983). Since few women arrived with the European settlers, men formed liaisons with slave women brought from the west African coast in the region, now known as Senegambia, thus creating a new group of individuals, the “mulattos” or “crioulos”, who would become the majority of the population (Godinho 1965). Among the first slaves brought into the Archipelago were also Guanches from the Canary Islands and north Africans. The population grew quickly, mainly because of the slave trade from the present-day Guiné-Bissau region (Russel-Wood 1998; Barry 1998).

R. Gonçalves · A. Rosa · A. Freitas · A. Fernandes · A. Brehm (✉)
Human Genetics Laboratory, Center of Macaronesian Studies,
University of Madeira, Campus of Penteadá,
9000-390 Funchal, Portugal
Tel.: +351-291-705383, Fax: +351-291-705393,
e-mail: brehm@uma.pt

A. Rosa · T. Kivisild · R. Villems
Tartu University, Estonian Biocenter, Riia 24, Tartu, Estonia

The European population in Cabo Verde has never been numerous. The southeastern islands Santiago and Brava had little more than 100 Europeans but 13,000 African slaves in 1582 (Godinho 1965, Russell-Wood 1998). Following the discovery of São Tomé and Príncipe Islands in the Gulf of Guinea, Cabo Verde started to decline as a slave-trading center, which led to a minimal input of new European settlers. The settlement process of the northwestern islands of Santo Antão and São Nicolau began much later than that in Santiago, viz., in the 17th century. Tradition says that the settlers were probably a sub-set of fugitive slaves from the southeastern islands together with a few Europeans. Several drought and famine periods substantially decreased the number of inhabitants and promoted inter-island migrations.

A strong geographic genetic differentiation within the Archipelago of Cabo Verde has previously been detected by using mitochondrial DNA (Brehm et al. 2002) and nuclear markers (Fernandes et al. 2003; Lessa and Rufié 1960; Spínola et al. 2002; A. Freitas, A. Brehm, J. Jesus, T. Kivisild, R. Villems, in preparation). The main aim of the present work has been to analyze the Y-chromosome gene pool of the present-day Cabo Verde population and to quantify the relative paternal input of European and African origin. We have used 32 Y-chromosome bi-allelic markers to characterize the Y-chromosome profile (lineages) of the Cabo Verde and Guinean populations. Given the particular interest of these islands as a melting pot of sub-Saharan slaves and European male colonizers, we have further sought inter-island differences that may reveal spatial differences in the colonization process.

Materials and methods

Population samples

The populations included in this study consisted of a total of 201 unrelated males from the Archipelago of Cabo Verde. We followed the same approach of Brehm et al. (2002) and, for historical and geographic reasons, divided the Archipelago of Cabo Verde into two groups: the leeward group (southeastern islands) comprising Brava, Fogo, Santiago, Maio, and the later settled islands of Boavista and Sal (CVS, $n=100$), and the windward group (northwestern islands) consisting of São Nicolau, São Vicente, and Santo Antão (CVN, $n=101$). Blood samples were collected from individuals who were also subjected to an interview in order to select those that could unambiguously certify that all relatives extending back for three generations were from the same island. In order to compare the origin of Y-lineages present in Cabo Verde, we also analyzed 276 males from the Republic of Guiné-Bissau (GU), on the west African coast and the putative origin of the sub-Saharan population of Cabo Verde (Fig. 1). The results were placed into the context of European, north African, and Middle Eastern (Jewish) populations taken from published sources (Hammer et al. 2000; Semino et al. 2000; Underhill et al. 2000; Bosch et al. 2001; Nebel et al. 2001).

DNA extraction and Y-chromosome typing

Genomic DNA was isolated from whole blood containing EDTA, by using the Chelex standard method (Lareu et al. 1994). The Y-chromosome single-nucleotide polymorphisms (SNPs) analyzed were amplified by using the DNA primers and methodology de-

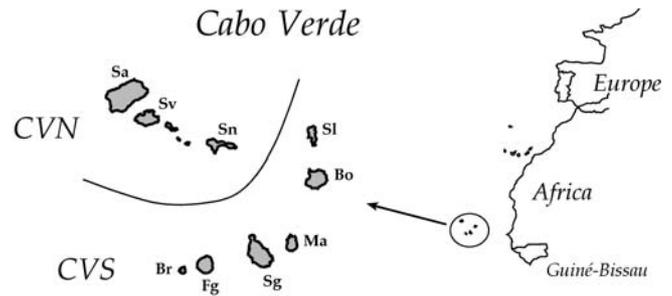


Fig. 1 Map of the Cabo Verde Archipelago and Guiné-Bissau. CVN (Cabo Verde northwestern group: Sa Santo Antão, Sv São Vicente, Sn São Nicolau) and CVS (Cabo Verde southeastern group: Sl Sal, Bo Boavista, Ma Maio, Sg Santiago, Fg Fogo, Br Brava) are the two sub-groups of islands

scribed in Underhill et al. (2000, 2001). The following SNPs were screened: M2, M9, M31, M32, M33, M35, M45, M62, M70, M75, M91, M124, M147, M168, M170, M172, M173, M175, M201, M207, P2, P15, P25, P36, SRY-1532, SRY-8299, LLY22 g, Apt, YAP (M1), 12f2 deletion, Tat, and SRY-9138. The SNPs were genotyped by restriction fragment length polymorphism analysis or direct sequencing.

Phylogeography of Y-haplotypes and data analysis

In this study, we follow the nomenclature of Y-chromosome haplogroups given by the Y Chromosome Consortium (2002). A subset of 32 markers potentially informative in classifying Y-chromosomes of African and European descent were used in genotyping the Cabo Verdean and Guinean samples.

The frequencies of Y-haplogroups for each island and the gene diversity measure for the pooled CVN and CVS groups and Guiné-Bissau were obtained by using Arlequin v2.000 (Schneider et al. 2000). These frequencies were employed in an analysis of molecular variance (AMOVA) by using Euclidean distances between all pairs of haplotypes (Excoffier et al. 1992). The total genetic variation between the populations was partitioned into hierarchical levels of grouping, and variance components were tested for significance by nonparametric randomization tests with 10,000 permutations under the null hypothesis of no population structure. Principal component analysis (PCA) of Y-haplogroup frequencies from Cabo Verde and Guinea and published data from European and African populations was performed by using the MVSP v.3.12 statistical package, and the position of each population was plotted in two dimensions. The relative contribution of sub-Saharans to the present-day populations of Cabo Verde taking as “parental” the populations of Portugal and Guinea was evaluated by employing two conventional frequency-based estimators m_R and m_C (Roberts and Hiorns 1965; Chakraborty et al. 1992; Long 1991) as implemented in ADMIX v1.0 (Bertorelle and Excoffier 1998). Both estimators are presumed to give the proportion of each parental population to a “hybrid” population.

Results and discussion

Y-chromosome SNPs

The markers used in this study identified 16 haplogroups among the populations surveyed (Fig. 2). Haplogroups A and B are the basal clades in the Y-human tree and occur only in sub-Saharan Africans at modest frequencies (Y Chromosome Consortium 2002; Underhill et al. 2000,

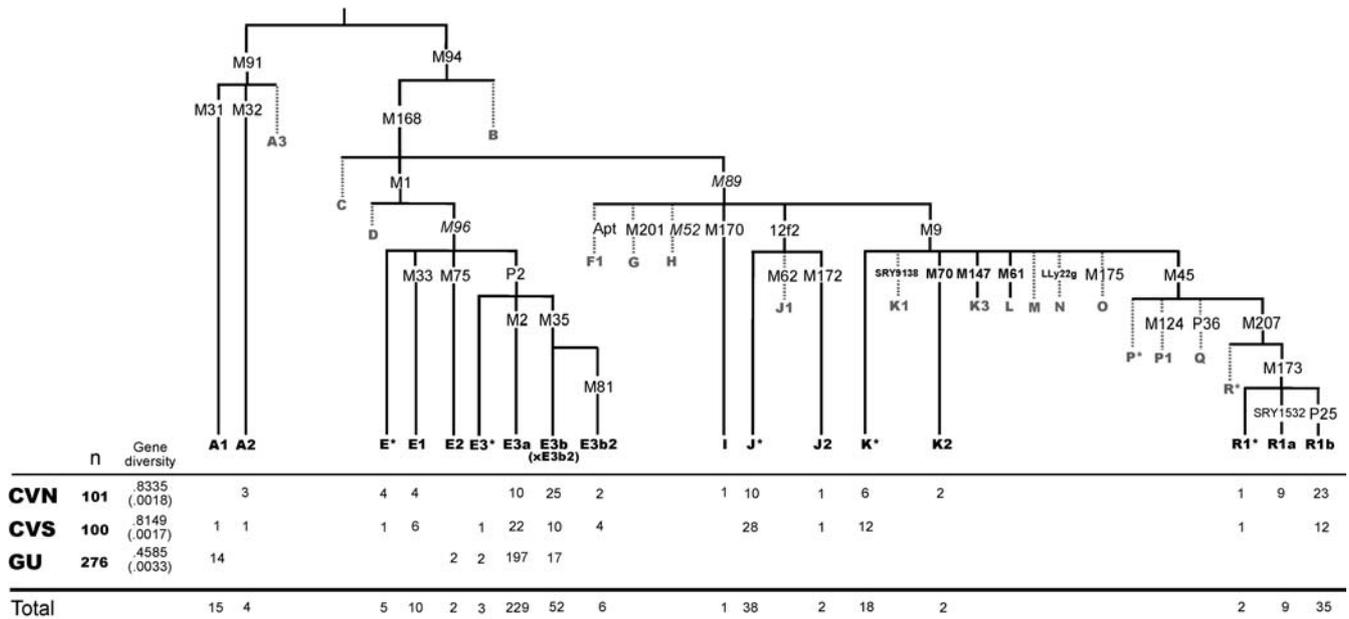


Fig. 2 Phylogenetic tree of Y-chromosome haplotypes found in this study. Mutations examined here are in *bold* and *black*. Mutations that were skipped are in *italics*. Haplogroup terminology is according to the Y Chromosome Consortium (2002). *Gray branches, gray haplogroup designations* Clades not found in our samples

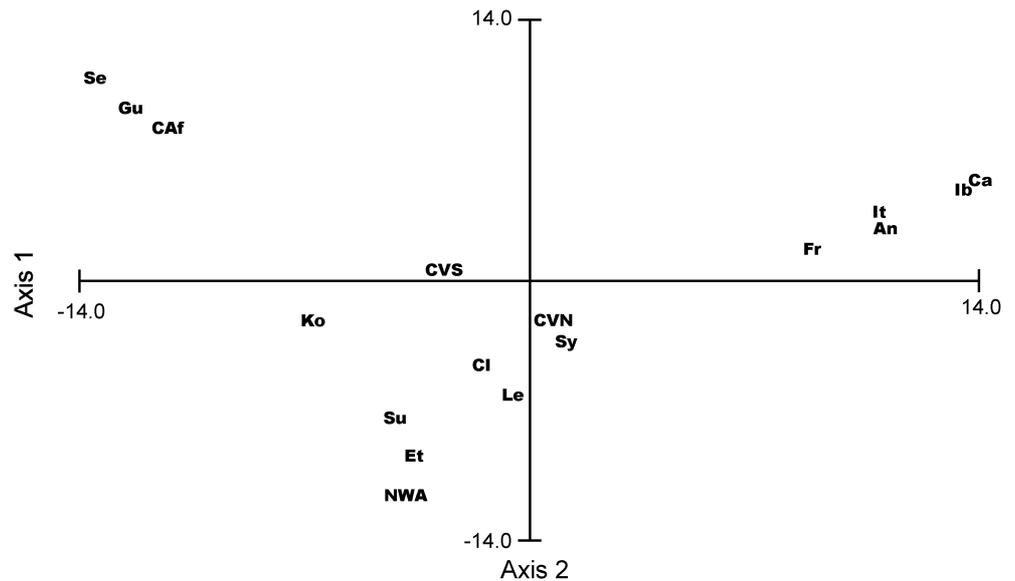
2001; Semino et al. 2002). Haplogroup A constitutes 5% of our Guinean sample, almost exclusively captured by the A1 sub-clade defined by the M31 SNP. Haplogroup E3a, defined by M2 is typically the most wide-spread clade in sub-Saharan Africa and by far the most frequent one in west African populations (Underhill et al. 2001; Semino et al. 2000, 2002). Its particular spatial pattern and high frequencies have been associated with the agricultural expansion of the Bantu speakers (Underhill et al. 2001). The M2 clade covers 71.3% of the Guinean lineages. In contrast, in Cabo Verdeans, E3a constitutes, on average, only 15.9% of the Y-chromosomes showing a significant difference between northern and southern groups (21.7% in CVS vs. 10% in CVN, $P < 0.0001$). E3a lineages also have a marginal frequency in the Canary Islands (Flores et al. 2003) and are absent in Europe (Semino et al. 2000) and Iberia (Bosch et al. 2001; own unpublished data). E3b, characterized by mutation M35, probably has an east African origin. The group occurs among Ethiopians (Semino et al. 2002) and Sudanese (Underhill et al. 2000) and appears at frequencies 12%–23% in various Jewish populations (Nebel et al. 2001). In northwest Africa, E3b is the most common cluster (reaching ~75% of Y-lineages; Bosch et al. 2001) but is characterized by an additional M81 mutation, giving rise to the sub-clade E3b2. E3b (but not E3b2) is found at 6.1% in our Guinean sample but constitutes 20.4% of Caboverdeans with an unequal distribution among the two groups of islands: 13.8% in CVS against 27% in CVN ($P < 0.0001$). The relatively high proportion of E3b lineages without the characteristic northwest African marker M81 in Cabo Verde islanders increases their affinity to populations from northeast Africa and the Middle East

(Fig. 3). The most plausible explanation for this phenomenon is the historically attested flow of Jewish immigrants to the islands in the late 15th century.

Haplogroups I, J, K, and R1, which are common in circum-Mediterranean populations of Europe, the Middle East, and north Africa but absent in sub-Saharan Africa, account for 53.5% of the Cabo Verdean sample. Clade R1 is the most frequent and widespread Y-chromosomal haplogroup in Europe (~50%), probably having an Eurasian origin that traces back to the earliest colonization of Europe and west Asia by modern humans (Semino et al. 2000). Two major sub-clades of this haplogroup, viz., R1a and R1b, encapsulate all R1 haplotypes in Europe (Cruciani et al. 2002) showing opposite frequency distributions (Semino et al. 2000; Rosser et al. 2000). West Europeans almost completely lack R1a but show the highest frequency of R1b. Iberian populations, in particular, show >77% of R1b lineages and 1% or less of R1a lineages (Bosch et al. 2001). In north Africa, only R1* (potentially R1b) lineages have been detected at a marginal frequency of 2.8% (Bosch et al. 2001). The Cabo Verde Archipelago follows a peculiar pattern: 12.8% of CVS lineages are R1b and none R1a, but CVN lineages exhibit 24% R1b and 9% R1a. In total, 22.9% of Caboverdean Y-chromosomes are R1, with a strong and significant asymmetry with respect to the distribution of the two haplogroups on both island groups (CVN and CVS).

The second most represented “circum-Mediterranean” haplogroup is the J haplotype (generally characterized by the 12f2*8 kb allele and M172 mutations). This haplogroup is present at a frequency of 2%–7% in Iberian and north African males (Bosch et al. 2001; Semino et al. 2000) and is thought to have originated in the Middle East where its frequency exceeds one third of the Y-chromosomes of Jewish, Turkish, and Arab populations (Bosch et al. 2001; Semino et al. 2001; Nebel et al. 2001). The J haplotype appears in Cabo Verde with a frequency of 19.4% and also shows a significant asymmetry between the two

Fig. 3 Principal component analysis of the relative frequencies of Y-chromosome haplogroups from Cabo Verde and several other populations retrieved from the literature and used for comparison. Axes 1 and 2 extracted 46% and 25.4% of the total variation. Populations are as follows: *Se* Senegal, *Gu* Guinea, *CAf* Central Africa, *Ko* Khoisan, *Su* Sudan, *Et* Ethiopia, *NWA* northwest Africa, *CI* Canary Islanders, *Le* Lebanon, *Sy* Syria, *Fr* French, *It* Italy, *An* Andalusia (Iberia), *Ib* Iberia, *Ca* Catalonia, *CVS* Cabo Verde southeastern group, *CVN* Cabo Verde northwestern group



groups of islands (27.7% and 11% in CVS and CVN respectively, $P < 0.001$). Of the east Atlantic island populations, CVS shows thus the highest proportion of haplogroup J lineages in comparison to only ~3% of the male population from Madeira, 10% in the Azores (unpublished), and 14% in the Canary Islands (Flores et al. 2003). Finally, the K2 and K* lineages are spread at low frequencies in the western Mediterranean region, occurring at 2%–3% in Portugal (Rosser et al. 2000), Iberia, and Morocco (Underhill et al. 2000; Semino et al. 2000; Bosch et al. 2001). Clade K2 chromosomes (M70) are present in east African populations (Ethiopians) at about ~4% (Semino et al. 2002) but have not been detected in a sample of 176 north African males (Bosch et al. 2001). Cabo Verde presents a high frequency of these lineages: ~10% are almost equally represented in both groups of islands. Because, no K lineages were found in our sample from Guiné-Bissau, we assume that those existing in Cabo Verde are probably derived from European or Middle Eastern settlers.

Population structure and PCA

An AMOVA between the two Cabo Verde groups of islands shows a large amount of variance attributed to differences among the two groups (5%) and an F_{ST} value of 0.05 ($P < 0.0001$). These values indicate a vast difference between the two populations of Cabo Verde, especially if we take into consideration that the F_{ST} value among the Iberian, French, and Italian populations studied by Bosch et al. (2001) was ~2% with $P \sim 0.08$. Additional AMOVAs with published data on populations pooled by geographic areas (Europe: Iberian, French and Italians; northwest Africa: Arabs and Berbers; western Africa: Guiné and Senegal) yield the results presented in Table 1. The highest percentage of variance is clearly found between both Cabo Verde and the sub-Saharan populations. It is impor-

Table 1 AMOVA (F_{ST}) results of variation among populations grouped according to geography (CVS Cabo Verde southeastern group, CVN Cabo Verde northwestern group). Populations used for comparison were recovered from Bosch et al. (2001) and Semino et al. (2000)

	Europe	Northwest Africa	Sub-Saharan Africa
CVS	0.2662	–	0.2739
CVN	0.1817	–	0.3435
Cabo Verde (pooled)	0.2024	0.1947	0.2730

($P < 0.00001$ all ways)

tant to note that CVN shows a significantly lower variation with European populations than does CVS. These results are also evident in the PCA performed with a large set of populations from various geographic origins. The haplogroups with most weight in the first two axes of the PCA are R1b, E3a, and E3b (Fig. 3). Three major clusters emerging from this analysis roughly correspond to the geographic spread of these haplogroups: central and west Africans share the high frequency of E3a, west European populations the high proportion of R1b lineages, and populations from the Middle East and north and east Africa show relatively the highest frequencies of haplogroup E3b. Both sub-populations of Cabo Verde cluster between these three major clusters, closest to the latter of the three, consistent with the high admixture rate of the island populations. In striking contrast to the PCA based on mitochondrial DNA haplogroups, both populations are most distant from the west sub-Saharan cluster, CVN being more closely related to a group composed by Middle East and east African populations. The position of the Canary Islands in this group is explained by the heavy weight of haplogroups highly represented in north Africa and the Middle East (especially E3b).

Table 2 Comparison of admixture estimators under different hypothesis of west Eurasian and north African parental populations to the hybrid population of Cabo Verde. m_R is based on Roberts and Hiorns (1965), and m_C was proposed by Chakraborty et al. (1992) based on a maximum-likelihood estimator from Long (1991). P1 and P2 are parental populations relative to the hybrid population H. Both estimators refer to the percentage of admixture

H	P1	P2	m_R (SD)	m_C (SD)
Cabo Verde	Iberia	Northwest Africa	0.4711 (0.0699)	0.4607 (0.0710)
Cabo Verde	Iberia	East Africa	0.3762 (0.0571)	0.7165 (0.1751)
Cabo Verde	Iberia	Middle East	0.1286 (0.0609)	0.0637 (0.0571)
Cabo Verde	Senegambia	Northwest Africa	0.4255 (0.0739)	0.4340 (0.0736)

Admixture estimates

Because of the high phylogeographic resolution of Y-chromosome haplogroups (Underhill et al. 2001), it is relatively simple to determine the lineages originating from sub-Saharan populations of west Africa and to estimate their contribution to the population of Cabo Verde. The total frequency of haplogroups A and E covering 100% of the 276 Guinean Y-chromosomes is 48% in CVN and 46% in CVS. These values represent the maximum proportion of the west African lineages in the islands. However, since E3b has a significantly higher frequency in north Africans (75%; Bosch et al. 2001) and Middle Eastern populations (12%–22% in Jews; Nebel et al. 2001) than in west Africans (6%), it seems likely that the E3b lineages arrived in Cabo Verde largely from a different source. Excluding E3b, the west African proportion drops to 21% in CVN and to 32% in CVS. The small proportion of sub-Saharan African lineages agrees, on one hand, well with previous data on HLA and STR markers (Spínola et al. 2002; Fernandes et al. 2003) but contrasts sharply with mitochondrial DNA data in which 90% of the population of Cabo Verde carries sub-Saharan mitochondrial haplogroups (Brehm et al. 2002). It is less trivial to pinpoint the exact phylogeographic origin of the of E3b, I, J, K, and R1 lineages. Here, we used estimators m_R and m_C to compare the relative proportion of Iberian, Middle Eastern, and east and north African populations to the “hybrid” population of Cabo Verde (Table 2). These estimators have been demonstrated to be reliable, even with single locus data, when both parental populations are distant to each other (Bertorelle and Excoffier 1998; Bertorelle and Dupanloup 2001). By using Iberians as the first parental group, the Middle East (and also east Africa) appears as another candidate region for significant gene flow to Cabo Verde. The affinity of populations from the Middle East and east Africa can be explained by the relatively high frequency of J* lineages in Cabo Verde and their absence in Iberia. As the J* lineages occur frequently in Jewish populations (Nebel et al. 2001), they may be attributable to the involvement of settlers of Jewish origin; the preservation of multiple toponyms in the countryside of Cabo Verde of unquestionable Jewish origin bolsters the interpretation that Jewish founders left a significant imprint in the genesis of the present-day paternal gene pool of Cabo Verde islanders.

from P1 relative to H. SD is the bootstrap standard deviation of the admixture coefficients (10,000 steps). Cabo Verde represents the pooled data. In the first three analysis, the west sub-Saharan genetic component including haplogroups A, E*, E1, E2, and E3a is excluded. The last analysis includes only the sub-Saharan component

A comparison of two African sources with the Cabo Verde islanders indicates that only haplogroup E3b is shared by the Guinean sample and north Africans (Bosch et al. 2001). The relatively high homogeneity of the west Africans here is unlikely to be a consequence of restricted sampling, because the 276 Guineans represent an extremely diverse assortment of ethnic groups, some of them, interestingly, with known historical records of trade with north African people. The two populations of Cabo Verde (CVS and CVN) differ significantly, in an opposite way, in their proportion of sub-Saharan (A and E3a) and north African (E3b) haplogroups, whereas they have a similar proportion of west Eurasian lineages, totalling about 60%. Although the populations of the separate islands of the windward group differ considerably in haplogroup E3b frequencies (data not shown), the difference between the windward and leeward groups probably reflects the effect of drift rather than differences in the source of the settlers. More than one half of the samples of the leeward (CVS) group were taken from its major island Santiago, which showed the lowest frequency of E3b (6%).

Although the present-day population of Cabo Verde is considered to be the result of admixture of mainly sub-Saharan Africans from the coast of Guinea with a comparatively small percentage of European (mainly Iberian) male colonizers (Carreira 1983), the significant differences between Y-chromosomal haplogroup profiles in the Archipelago and Guiné attests to the influence of north African, European, and Middle Eastern colonists. Given that Europeans composed less than 10% of the total population and the short time of colonization history, the differences between Cabo Verde and Guiné cannot be attributed solely to genetic drift. In contrast to mitochondrial DNA profiles, which reflect the high proportional ratio of west sub-Saharan admixture, the paternal legacy of the minority group of settlers is such that the present-day Cabo Verde population should never be simply characterized as African.

Acknowledgements We express our gratitude to the Chairmen of the Joint Chiefs of Staff from the Republic of Cabo Verde and from the Republic of Guiné-Bissau. We are also grateful to the Ministry of Public Health of Guiné-Bissau. We thank Dr. Peter A. Underhill, Dr. José Jesus, and two anonymous reviewers for useful suggestions and comments on earlier versions of the manuscript. The support of ICCTI (Ministry of Science, Lisbon) and the Regional Government of Madeira (SRE) to A.B. is acknowledged. R.G. has a research contract with the University of Madeira. A.F. is an FCT PhD student (contract SFRH/BD/8592/2002).

References

- Barry B (1998) Senegambia and the Atlantic slave trade. Cambridge University Press, Cambridge
- Bertorelle G, Dupanloup I (2001) Inferring admixture proportions from molecular data: extension to any number of parental populations. *Mol Biol Evol* 18:672–675
- Bertorelle G, Excoffier L (1998) Computing admixture coefficients from molecular data. *Mol Biol Evol* 15:1298–1311
- Bosch H, Calafell F, Comas D, Oefner P, Underhill P, Bertranpetit J (2001) High-resolution analysis of human Y-chromosome variation shows a sharp discontinuity and limited gene flow between northwestern Africa and the Iberian Peninsula. *Am J Hum Genet* 68:1019–1029
- Brehm A, Pereira L, Bandelt H-J, Prata M, Amorim A (2002) Mitochondrial portrait of the Cabo Verde archipelago: the Senegambian outpost of the Atlantic slave trade. *Ann Hum Genet* 66:49–60
- Carreira A (1983) Migrações nas ilhas de Cabo Verde, 2nd edn. Instituto Caboverdeano do Livro, Lisboa
- Chakraborty R, Kamboh MI, Nwankwo M, Ferrell RE (1992) Caucasian genes in American Blacks: new data. *Am J Hum Genet* 50:145–155
- Cruciani F, Santolamazza P, Shen P, Macaulay V, Moral P, Olckers A, Modiano D, Holmes S, Destro-Bisol G, Coia V, Wallace D, Oefner P, Torroni A, Cavalli-Sforza L, Scozzari R, Underhill P (2002) A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *Am J Hum Genet* 70:1197–1214
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491
- Fernandes A, Velosa R, Jesus J, Carracedo A, Brehm A (2003) Genetic differentiation of the Cabo Verde Archipelago population analysed by STR polymorphisms. *Ann Hum Genet* 67:(in press)
- Flores C, Maca-Meyer N, Pérez J, González A, Larruga J, Cabrera V (2003) A predominant European ancestry of paternal lineages from Canary Islanders. *Ann Hum Genet* 67:138–152
- Godinho VM (1965) Os Descobrimentos e a Economia Mundial, vol 1 and 2. Arcádia, Lisbon
- Hammer MF, Redd A, Wood E, Bonner M, Jarjanazi H, Karafet T, Santachiara-Benerecetti S, Oppenheim A, Jobling M, Jenkins T, Ostrer H, Bonnè-Tamir B (2000) Jewish and Middle Eastern non-Jewish populations share a common pool of Y-chromosome biallelic haplotypes. *Proc Natl Acad Sci USA* 10:1073–1078
- Hurles M, Veitia R, Arroyo E, Armenteros M., Bertranpetit J, Pérez-Lazeum A, Bosch E, Shumukova M, Cambon-Thomsen A, McElreavey K, Munain A, Rohl A, Wilson I, Singh L, Pandya A, Santos F, Tyler-Smith C, Jobling M (1999) Recent male-mediated gene flow over a linguistic barrier in Iberia, suggested by analysis of a Y-chromosomal DNA polymorphism. *Am J Hum Genet* 65:1437–1448
- Lareu MV, Phillips C, Carracedo A, Lincoln P, Syndercombe D, Thompson J (1994) Investigation of the STR locus HUMTH01 using PCR and two electrophoresis formats: UK and Galician Caucasian population surveys and usefulness in paternity investigations. *Forensic Sci Int* 66:41–52
- Lessa A, Ruffie J (1960) Seroantropologia das Ilhas de Cabo Verde. Mesa redonda sobre o homem Cabo-Verdiano. Estudos, Ensaios e Documentos, vol 32. Junta de Investigações do Ultramar, Lisboa
- Long JC (1991) The genetic structure of admixed populations. *Genetics* 142:417–428
- Malaspina P, Tsopanomalou M, Duman T, Stefan M, Silvestri A, Rinaldi B, Garcia O, Giparaki M, Plata E, Kozlov A, Barbujani G, Vernesi C, Papola F, Ciavarella G, Kovatchev D, Kerimova M, Anagnou N, Gavrila L, Veneziano L, Akar N, Loutradis A, Michalodimitrakis E, Terrenato L, Novelletto A (2001) A multi-step process for the dispersal of a Y chromosomal lineage in the Mediterranean area. *Ann Hum Genet* 65:339–349
- Nebel A, Filon D, Brinkmann B, Majumder PP, Faerman M, Oppenheim A (2001) The Y chromosome pool of Jews as part of the genetic landscape of the Middle East. *Am J Hum Genet* 69:1095–1112
- Roberts D, Hiorns R (1965) Methods of analysis of the genetic composition of a hybrid population. *Hum Biol* 37:38–43
- Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley DG, Brede G, Cooper G, Corte-Real HB, Knijff P de, Decorte R, Dubrova YE, Evgrafov O, Gilissen A, Glisic S, Golge M, Hill EW, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, Kravchenko SA, Krumina A, Kucinskias V, Lavinha J, Livshits LA, Malaspina P, Maria S, McElreavey K, Meitinger TA, Mikelsaar AV, Mitchell RJ, Nafa K, Nicholson J, Norby S, Pandya A, Parik J, Patsalis PC, Pereira L, Peterlin B, Pielberg G, Prata MJ, Previdere C, Roewer L, Rootsi S, Rubinsztein DC, Saillard J, Santos FR, Stefanescu G, Sykes BC, Tolun A, Villems R, Tyler-Smith C, Jobling MA (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* 67:1526–1543
- Russel-Wood AJ (1998) The Portuguese Empire, 1415–1808. A world on the move. Johns Hopkins University Press, Baltimore
- Schneider S, Kueffer J-M, Roessli D, Excoffier L (2000) Arlequin: a software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland
- Scozzari R, Cruciani F, Santolamazza P, Malaspina P, Torroni A, Sellitto D, Arredi B, Destro-Bisol G, De Stefano G, Rickards O, Martinez-Labarga C, Modiano D, Biondi G, Moral P, Olckers A, Wallace D, Novelletto A (1999) Combined use of bi-allelic and microsatellite Y-chromosome polymorphisms to infer affinities among African populations. *Am J Hum Genet* 65:829–846
- Semino O, Passarino G, Oefner P, Lin A, Arbuzova S, Beckman L, De Benedictis G, Francalacci P, Kouvatsi A, Limborska S, Marcikiae M, Mika A, Mika B, Primorac D, Santachiara-Benerecetti A, Cavalli-Sforza L, Underhill P (2000) The genetic legacy of palaeolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science* 290:1155–1159
- Semino O, Santachiara-Benerecetti A, Falaschi F, Cavalli-Sforza L, Underhill P (2002) Ethiopians and Khoisans share the deepest clades of the human Y-chromosome phylogeny. *Am J Hum Genet* 70:265–268
- Spínola H, Brehm A, Williams F, Jesus J, Middleton D (2002) Distribution of HLA alleles in Portugal and Cabo Verde. Relationships with the slave trade route. *Ann Hum Genet* 66:285–296
- Underhill P, Shen P, Lin A, Jin L, Passarino G, Yang W, Kauffman E, Bonne-Tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, Kidd J, Mehdi S, Seielstad M, Wells R, Piazza A, Davis R, Feldman M, Cavalli-Sforza L, Oefner P (2000) Y chromosome sequence variation and the history of human populations. *Nat Genet* 26:358–361
- Underhill P, Passarino G, Lin A, Shen P, Mirazon Lahr M, Foley R, Oefner P, Cavalli-Sforza L (2001) The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet* 65:43–62
- Y Chromosome Consortium (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* 12:339–348