

Original Research Article

Y-Chromosome Lineages in São Tomé e Príncipe Islands: Evidence of European Influence

RITA GONÇALVES,* HÉLDER SPÍNOLA, AND ANTÓNIO BREHM

Human Genetics Laboratory, University of Madeira, Campus of Pentecada, Funchal 9000-390, Portugal

ABSTRACT The Y-chromosome haplogroup composition of the population of São Tomé e Príncipe (STP) archipelago was analyzed using 25 biallelic markers and compared with populations of different origins from Europe, Africa, and the Middle East. Two main Y-chromosome haplogroups were found: E3a, very common among sub-Saharan accounts for 84.2% of the paternal lineages and R1b, typical of West Eurasia, represents 8.7% of the overall male population. Nevertheless, we detected in the population of STP a significant heterogeneous distribution of R1b among the two main ethnic groups of the archipelago: *Forros* (10.3%) and *Angolares* (6.6%). Together, haplogroups known to be prevalent in West Eurasia reach 12.5% of the chromosomes analyzed unequally distributed among the two groups: *Forros* present 17.7% while *Angolares* display only 8.2% of west Eurasian haplogroups. Our findings suggest that, despite its sub-Saharan genetic background, a relevant contribution of European paternal lineages is present in nowadays STP population. This influence has shown to be stronger in *Forros* than in *Angolares*, which could be explained by the social isolation that these have last experienced through their history. *Am. J. Hum. Biol.* 19:422–428, 2007. © 2007 Wiley-Liss, Inc.

Polymorphisms in the human Y-chromosome, namely base substitutions and insertion/deletions, are especially useful on the detection of male migration and admixture patterns (Bosch et al., 2001; Shen et al., 2000; Underhill et al., 1997, 2000). Y-chromosomal biallelic markers have been extensively used in order to provide a paternal historic view of relationships between populations (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). São Tomé e Príncipe (STP) islands, located in the Gulf of Guinea, 300 km from the West coast of Africa, were discovered uninhabited by Portuguese sailors in 1470 (Peres, 1960). These islands were settled by a myriad of people from different origins of Africa and Europe. The first populace was composed mainly by sub-Saharan African slaves from the Gulf of Guinea, Congo and Angola, recruited to work in local plantations, and in a minor extent Portuguese. In the first centuries after their discovery STP acted as an outpost for slave trade between Africa and the Americas. Besides Portuguese, other Europeans were involved on this trade along the coast of Africa, namely French, Spanish, Dutch, and English, which could have contributed in a minor scale to the present-day genetic pool of islanders (Neves,

1989). Since the beginning of the settlement process, Portuguese males were allowed to blend with female slaves (Garcia, 1966) and slaves from several geographic and ethnic origins could mix together as well (Tenreiro, 1961). In the 19th century, because of a new economic cycle supported by coffee and cacao plantations, the islands received a new wave of sub-Saharan African people recruited from Cabo Verde archipelago, Angola, and Mozambique (Barata, 1966). The majority of São Tomé population speaks *Forros*, a mixed dialect between Portuguese and Bantu languages used by liberated slaves, known as *Forros*, considered the most ancient African inhabitants of the archipelago (Henriques, 2000; Tenreiro, 1961). *Mancó* and *Tonga* are two other ethnic groups that use Portuguese and African mixed languages; the former are mostly from Príncipe island and the latter descends from people arrived in the 19th century after slave abolition. The *Angolares* com-

*Correspondence to: Rita Gonçalves, Laboratório de Genética Humana, Universidade da Madeira, Campus da Pentecada, Funchal 9000-390, Madeira, Portugal. E-mail: ritaAG@uma.pt

Received 6 July 2006; Revision received 13 October 2006; Accepted 19 October 2006

Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/ajhb.20604

munity that inhabits São Tomé Island has resisted to miscegenation and still maintains their one own Bantu language. The origin of the *Angolares* people remains unknown and popular belief tells they are descends of the survivors of a slave shipwreck in the middle of the 16th century roaming from the West coast of Africa (Henriques, 2000; Romana, 1997). Most probably *Angolares* are just fugitive slaves from the sugar-mills who escaped colonial control by taking refuge in the most inaccessible forest of the South-Eastern region of São Tomé (Seibert, 1998).

The Archipelago of STP is thus an interesting model to study admixture of populations from genetically different backgrounds, especially Africa and Europe. Previous studies on mtDNA of STP population revealed that the maternal influence was almost completely of sub-Saharan origin, clearly belonging to a West African cluster (Mateu et al., 1997; Trovoada et al., 2004). A study on β -globin haplotypes and eight autosomal markers (APOA1, AT3, FY, LPL, OCA2, RB1, Sb19.3, and GC) indicated that the peopling of São Tomé provided the combination of diverse African contributions and European admixture (10.7%) that emerged from the overseas population relocations promoted by the Atlantic slave trade (Tomas, 2002). A previous study on seven Y-chromosome STR *loci* detected haplotypes, most likely of European ancestry on São Tomé e Príncipe, showed statistically significant differences between *Angolares* and *Forros* (Trovoada et al., 2001).

The main aim of the present work was to analyze the Y-chromosome gene pool of the present-day STP population and to quantify the relative paternal input of European and African origin. We have intended to search for genetic differentiation between *Angolares* and *Forros*, the two main and most ancient ethnic groups of this archipelago.

MATERIALS AND METHODS

Population sample

The population included in this study consisted of 150 unrelated males from the archipelago of STP (Fig. 1) (*Forros*: 68 samples; *Angolares*: 61 samples; *Mancó*: 14 samples; and *Tonga*: 7 samples). Blood samples were collected with informed consent from volunteer males, which could unambiguously certify that all relatives back to three generations were from this archipelago. For comparisons of the Y-chromosome profiles prevailing in

these islands we used published data on European (Capelli et al., 2003; Cruciani et al., 2004; DiGiacomo et al., 2004; Hammer et al., 2000; Nebel et al., 2001; Semino et al., 2000, 2004; Underhill et al., 2000), North African (Arredi et al., 2004; Bosch et al., 2001; DiGiacomo et al., 2004; Luis et al., 2004), Middle Eastern (Cinniogulu et al., 2004; Cruciani et al., 2004; DiGiacomo et al., 2004; Semino et al., 2002; Underhill et al., 2000), and sub-Saharan African populations (Gonçalves et al., 2003; Semino et al., 2000, 2002).

DNA extraction and Y-chromosome typing

Genomic DNA was isolated from whole blood containing EDTA using the Chelex standard method (Lareu et al., 1994) and the sequence tagged-site (STS) containing Y-biallelic markers were amplified with primers described in Underhill et al. (2000, 2001) and Cinniogulu et al., (2004). Genotyping was done by restriction fragment length polymorphism analysis or direct sequencing. The following SNPs were screened: M1, M2, M9, M14, M45, M89, M91, M94, M96, M168, M170, M173, M201, M207, P2, P25, and SRY-1532. The following binary markers were assayed but derived alleles not observed: M31, M32, M33, M35, M69, M75, M304, and P36.

Data analysis

In this study we follow the haplogroup nomenclature proposed by the Y Chromosome Consortium (2002) for Y-chromosome typing. Frequencies of Y haplogroups and gene diversity measures were obtained using Arlequin v2.000 (Schneider et al., 2002). These frequencies were used for comparisons with other populations and employed in an analysis of molecular variance (AMOVA) using Euclidean distances between all pairs of haplogroups (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 using 10,000 permutations. Principal component analysis (PCA) of Y haplogroup frequencies from STP 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 other populations to the present day population of STP, taking as "parental" the populations of West sub-Sahara, North West Africa, East Africa, Middle East, and Iberia,

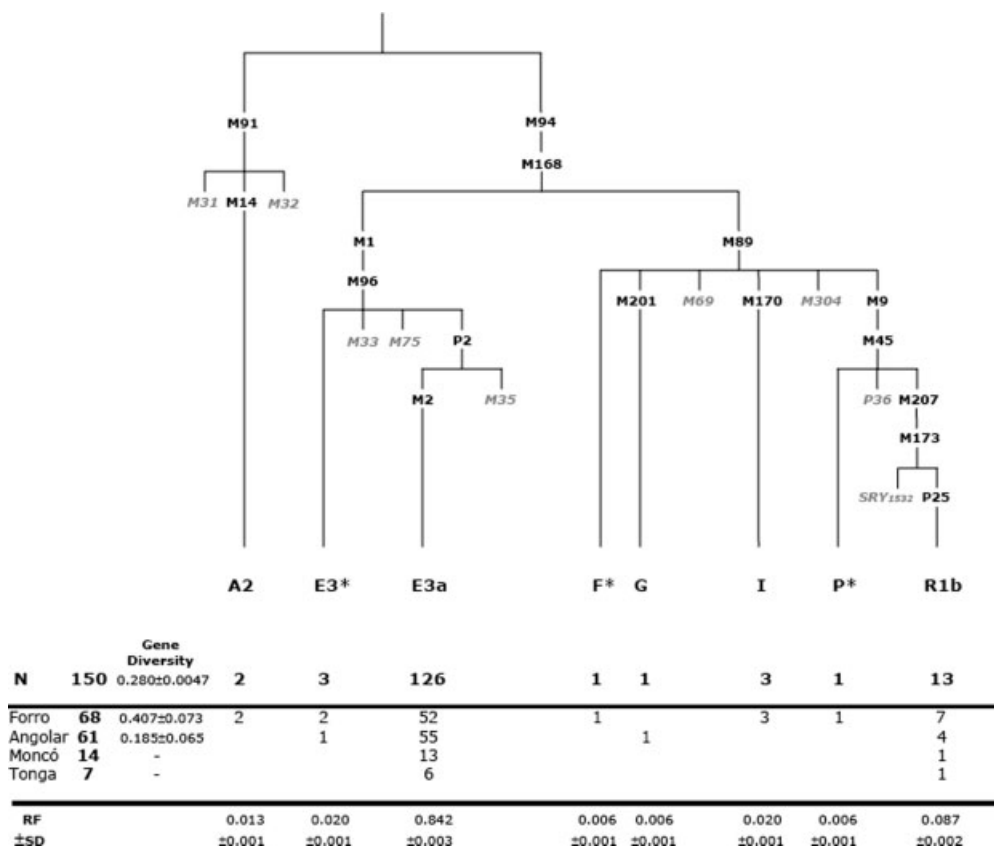


Fig. 1. Phylogenetic tree of Y-chromosome haplogroups found in São Tomé e Príncipe population. Haplogroup defining mutations assayed in this study are shown along branches. Gray indicates mutations assayed but not found. Table shows absolute frequencies. RF, Relative frequencies; SD, Standard deviation.

was evaluated by employing the program ADMIX 2.0 (Dupanloup and Bertorelle, 2001) that compute the admixture coefficient ($m\gamma$) described by Bertorelle and Excoffier (1998). All runs of ADMIX 2.0 were carried out fitting 10,000 random bootstrap samples.

RESULTS AND DISCUSSION
Y-chromosome variation

The biallelic markers used in this study identified eight haplogroups among the population of STP (Fig. 1). The most frequent haplogroup found is E3a (84.2%) also the most widespread clade in sub-Saharan Africa and by far the most frequent one in West African populations (Semino et al., 2000, 2002; Underhill et al., 2001) emphasizing the origin of the main settlers of these islands. The particular spatial distribution pattern of E3a has been

associated with the agricultural expansion of Bantu speakers (Underhill et al., 2001). Haplogroup E3a, defined by M2 mutation, is most common among Angolares (90%) than Forros (76.5%) ethnic groups. This clade reaches 80% among Senegalese (Semino et al., 2002), 71% of the Guinean paternal lineages, but only 15.9% in Cabo Verdeans (Gonçalves et al., 2003). 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; Gonçalves et al., 2005). The second most common haplogroup in São Tomé e Príncipe is R1b (8.7%), a typical haplogroup from Western Europe, most likely carried by Portuguese settlers suggesting their genetic contribution to this population. R1b is more predominant among Forros (10.3%) than Angolares (6.6%) ethnic groups. Clade R1 is the most frequent and

widespread Y-chromosomal haplogroup in Europe (~50%), probably having a Eurasian origin that traces back to the earliest colonization of Europe and West Asia (Semino et al., 2000). Two major sub-clades of this haplogroup, R1a and R1b, encapsulate all R1 haplogroups in Europe (Cruciani et al., 2002; Rosser et al., 2000; Semino 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). Haplogroup R1b was found to be the most dominant Y chromosomal lineage in Portugal, including the North Atlantic archipelagos of Azores and Madeira, covering more than half (55%) of the Y chromosomal lineages in each population (Gonçalves et al., 2005). More than 17% of Caboverdean Y-chromosomes are R1b, a West European influence in the archipelago settlement process (Gonçalves et al., 2003).

Other haplogroups found in STP show only marginal frequencies (Fig. 1). Haplogroup I, also a characteristic clade for many different European populations, constitute 2% of STP population. Among *Forros* this percentage reaches 4.4%. Haplogroups P* and G, with a Eurasian origin, and F*, Middle East origin, were found in only one individual each in STP. Considering haplogroups of West Eurasian origin (R1b, F*, G, I, and P*) found in STP, *Forros* shows a higher frequency (17.6%) than *Angolares* (8.2%).

Haplogroup A2, one of the basal clades in the Y-human phylogenetic tree, typical in sub-Saharan Africans at modest frequencies (Semino et al., 2002; Underhill et al., 2000, 2001; Y Chromosome Consortium, 2002), constitutes 1.3% of our samples.

Population structure and PCA

An AMOVA including the total population of STP as well as both main ethnic groups in separate (*Forros* and *Angolares*) with published data from populations pooled by different geographic regions yield the results presented in Table 1. The lower percentage of variance is clearly between STP and the West sub-Saharan populations, revealing once more the main origin of the settlers of these islands. Group population pair wise differences were all statistically significant ($P < 0.0001$). An AMOVA between *Forros* and *Angolares* show that the main variance found is attributed to differences within the two groups (97.7%) and

TABLE 1. AMOVA (FST) results of variation among populations grouped according to geography and STP overall, STP *Forros*, and STP *Angolares*

	West sub-Sahara	East Africa	North Africa	Europe
STP overall	0.13	0.43	0.56	0.57
STP <i>Forros</i>	0.098	0.37	0.496	0.501
STP <i>Angolares</i>	0.15	0.44	0.57	0.59

$P < 0.00001$ in all FST values.

Populations used for comparison were recovered from Bosch et al. (2001) and Semino et al. (2000). West Sub-Sahara: Guinea, Senegal; East Africa: Ethiopia, Sudan; North Africa: West Sahara, Morocco, Lybia; Europe: Iberia 1, Andalusia Iberia, Catalonia, France, Italy.

only 2.3% of variance were attributed to differences among them (FST value of 0.023, $P = 0.066$). Differentiation among *Forros* and *Angolares* was not statistically significant ($P = 0.26$) in opposition to the results obtained in a previous study on Y-Chromosome STR's (Trovoada et al., 2001).

The PCA (Fig. 2) join STP with sub-Saharanans but with a clear tendency of proximity to West Eurasian populations. A PCA constructed with *Forros* and *Angolares* plot together these groups with STP overall (*data not shown*).

Admixture estimates

The admixture coefficient ($m\gamma$) between the hybrid population of STP and some presumed parental populations was estimated using ADMIX 2.0 (Table 2). The program takes into account molecular information from any number of parental populations to compute the admixture coefficient ($m\gamma$), which estimates the proportion of each parental population to a "hybrid" population. High values of $m\gamma$ may be considered as an indicator of previous gene flow between populations. In this sense, the high value of $m\gamma$ when West sub-Saharan is taking as parental shows that these populations were the major genetic source for the STP present population. North West Africa, and in a low degree Middle East, appear also as another candidate region for gene flow to STP. East Africa and Iberia admixture coefficient were negative to STP, which means no admixture. When considering STP population groups, *Forros* and *Angolares*, we come across different degrees of gene flow. *Forros* show a positive value of admixture with Iberia (despite with a higher standard deviation value), even higher than North West Africa, in opposition to *Angolares* who reveal no admixture with Iberia and East Africa. *Angolares* present also the highest admixture coefficient value with West sub-Saharanans (Table 2).

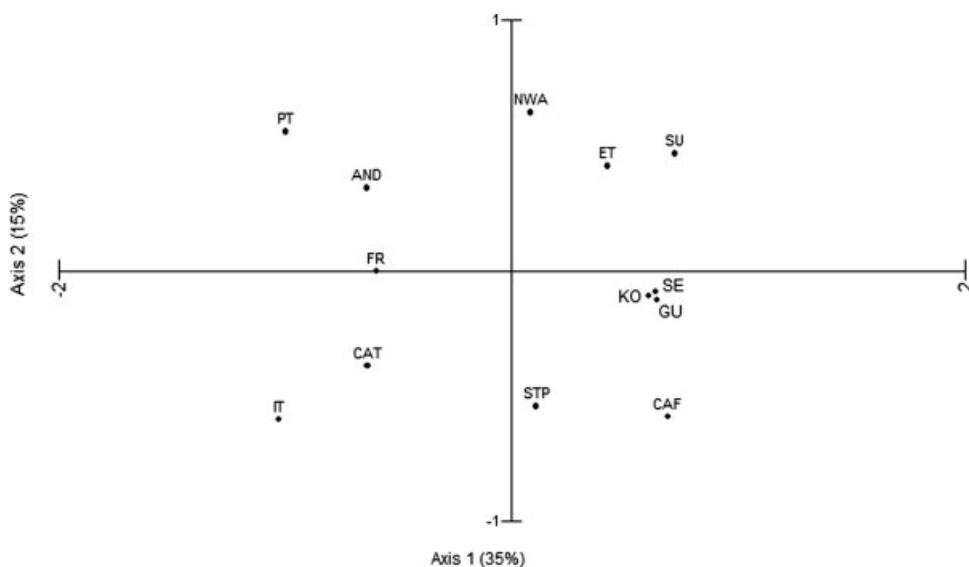


Fig. 2. Principal component analysis of the relative frequencies of Y-chromosome haplogroups from São Tomé e Príncipe and several other populations retrieved from the literature and used for comparison. Axis 1 extracted 35% and axis 2 extracted 15% of the total variation. Populations are as follows: SE, Senegal; GU, Guinea; CAF, Central Africa; KO, Khoisan; SU, Sudan; ET, Ethiopia; NWA, Northwest Africa; PT, Portugal; FR, France; IT, Italy; AND, Andalusia (Spain); CAT, Catalonia (Spain); and STP, São Tomé e Príncipe (for references see material and methods).

TABLE 2. Admixture coefficient ($m\gamma$) and Standard Deviation (SD), using the program ADMIX 2.0, between the hybrid populations STP overall, Forros or Angolares and some group populations thanking as parental

Parental	Hybrid		
	STP overall	Forros	Angolares
West Sub-Saharan	1.14 (0.05)	1.04 (0.07)	1.23 (0.057)
North West Africa	0.11 (0.08)	0.0002 (0.11)	0.21 (0.076)
Middle East	0.10 (0.04)	0.05 (0.046)	0.14 (0.037)
East Africa	-0.32 (0.11)	0.12 (0.17)	-0.49 (0.103)
Iberia	-0.03 (0.04)	0.03 (0.06)	-0.087 (0.055)

10,000 bootstrap.

The present admixture analysis has an inherent lack of power since we used only a single locus. Only a simultaneous analysis of several loci could provide more reliable admixture estimates (Dupanloup and Bertorelle, 2001). However this analysis suggests the main parental populations involved in STP hybrid population.

Despite the present day STP paternal legacy is mainly of West sub-Saharan origin, as suggested by AMOVA, PCA, admixture coefficient and haplogroup frequencies, the West Eurasian genetic contribution is far to be considered marginal. Haplogroup R1b is the clearest evidence of this influence in STP islands. R1b, with an occurrence of 55% in the

Portuguese population (Gonçalves et al., 2005), was also found in Cabo Verde as a result of West European influence, where it reaches a higher frequency (17%) (Gonçalves et al., 2003), than in STP (8.7%). Together, haplogroups known to be of West Eurasian origin (R1b, I, P*, G, and F*) reach a value of 12.5% in STP. This frequency is not similar between the two group populations we are considering: Forros present 17.7% whereas Angolares only 8.2%.

Data obtained on SNP Y-Chromosome suggests a relevant contribution of West Eurasian paternal lineages in STP today's population. This influence appears to be stronger on Forros than Angolares, which could be explained

by the isolation that these latter maintained through centuries. The lower level of gene diversity on *Angolares* (0.185 ± 0.065) corroborates this hypothesis.

In opposition to previous results on Y-chromosome STR *loci* (Trovoada et al., 2001) our data on Y-Chromosome SNP's did not reveal statistically significant differences between *Forros* and *Angolares*. The percentage of European admixture found in a study on β -globin haplotypes and eight autosomal markers (10.7%; Tomas et al., 2002) is quite similar to the present study (12.5%), despite *Forros* show a higher European influence (17.7%).

Considering previous studies (Mateu et al., 1997; Trovoada et al., 2001, 2004) and the present data, we conclude that STP population, despite its sub-Saharan African main input lineages, presents a relevant European genetic component, most evident on *Forros*.

LITERATURE CITED

- Arredi B, Poloni ES, Paracchini S, Zerjal T, Fathallah DM, Makrelouf M, Pascali VL, Novelletto A, Tyler-Smith C. 2004. A predominantly neolithic origin for Y chromosomal DNA variation in North Africa. *Am J Hum Genet* 75:338–345.
- Barata OS. 1966. O povoamento de Cabo Verde, Guiné e São Tomé e Príncipe. In: Cabo Verde, Guiné e São Tomé e Príncipe. Curso de extenSão universitária. Ano lectivo de 1965–1966. ISCSPU. Lisboa. p 923–958.
- 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 E, Calafell F, Comas D, Oefner PJ, Underhill PA, 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.
- Capelli C, Redhead N, Abernethy J, Gratix F, Wilson J, Moen T, Hervig T, Richards M, Stumpf M, Underhill P, Bradshaw P, Shaha A, Thomas M, Bradman N, Goldstein D. 2003. A Y chromosome census of the British Isles. *Curr Biol* 13:979–984.
- Cinnioglu C, King R, Kivisild T, Kalfoglu E, Atasoy S, Cavalleri GL, Lillie AS, Roseman CC, Lin AA, Prince K, Oefner PJ, Shen P, Semino O, Cavalli-Sforza LL, Underhill PA. 2004. Excavating Y-chromosome haplotype strata in Anatolia. *Hum Genet* 114:27–148.
- Cruciani F, La Fratta R, Santolamazza P, Daniele Sellito, Pascone R, Moral P, Watson E, Guida V, Colomb EB, Zaharova B, Lavinha J, Vona G, Aman R, Cal F, Akar N, Richards M, Torroni A, Novelletto A, Scozzari R. 2004. Phylogeographic analysis of haplogroups E3b (E-M215) Y chromosomes reveals multiple migratory events within and out of Africa. *Am J Hum Genet* 74:1014–1022.
- Cruciani F, Santolamazza P, Shen P, Macaulay V, Moral P, Olekers 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.
- Di Giacomo F, Luca F, Popa LO, Akar N, Anagnou N, Banyko J, Brdicka R, Barbuiani G, Papola F, Ciavarella G, Cucci F, Di Stasi L, Gavrilu L, Kerimova MG, Kovatchev D, Kozlov AI, Loutradis A, Mandarinov V, Mammi C, Michalodimitrakis EN, Paoli G, Pappa KI, Pedicini G, Terrenato L, Tofaneli S, Malaspina P, Novelletto A. 2004. Y chromosomal haplogroup J as a signature of the post-neolithic colonization of Europe. *Hum Genet* 115:357–371.
- Dupanloup I, Bertorelle G. 2001. Inferring admixture proportions from molecular data: Extension to any number of parental populations. *Mol Biol Evol* 18:672–675.
- 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.
- 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.
- Garcia A. 1966. A ilha de S. Tomé como centro experimental do comportamento do luso nos trópicos. *Separata de DTVDIA* 19. Lisboa.
- Goncalves R, Freitas A, Branco M, Rosa A, Fernandes AT, Zhivotovsky LA, Underhill PA, Kivisild T, Brehm A. 2005. Y-chromosome lineages from Portugal, Madeira and Açores record elements of Sephardim and Berber ancestry. *Ann Hum Genet* 69:443–454.
- Goncalves R, Rosa A, Freitas A, Fernandes A, Kivisild T, Villems R, Brehm A. 2003. Y-chromosome lineages in Cabo Verde Islands witness the diverse geographic origin of its first male settlers. *Hum Genet* 113:467–478.
- Hammer M, Redd A, Wood E, Bonner M, Jarjanazi H, Karafet T, Santachiara-Benerecetti S, Oppenheim A, Jobling M, Jenkins T, Ostrer H, Bonn-Tamir B. 2000. Jewish and Middle Eastern non-Jewish populations share a common pool of Y-chromosome biallelic haplotypes. *Proc Natl Acad Sci USA* 97:6769–6774.
- Henriques IC. 2000. São Tomé e Príncipe. A invenção de uma sociedade. Vega e Autor.
- Hurles M, Veitia R, Arroyo E, Armenteros M, Bertranpetit J, Perez-Lazeau A, Bosch E, Shumukova M, Cambon-Thomson 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.
- Luis JR, Rowold DJ, Regueiro M, Caeiro B, Cinnioglu C, Roseman C, Underhill PA, Cavalli-Sforza LL, Herrera RJ. 2004. The Levant versus the Horn of Africa: Evidence for bidirectional corridors of human migrations. *Am J Hum Genet* 74:532–544.
- Malaspina P, Tsopanomalou M, Duman T, Stefan M, Silvestri A, Rinaldi B, Garcia O, Gíparaki M, Plata E, Kozlov A, Barbuiani G, Vernesi C, Papola F, Ciavarella G, Kovatchev D, Kerimova M, Anagnou N, Gavrilu L, Veneziano L, Akar N, Loutradis A, Michalodimitrakis E, Terrenato L, Novelletto A. 2001. A multistep process for the dispersal of a Y chromosomal lineage in the Mediterranean area. *Ann Hum Genet* 65:339–349.
- Mateu E, Comas D, Calafell F, Pe A, Rez-Lezaun, Abade A, Bertranpetit J. 1997. A tale of two islands: Population history and mitochondrial DNA sequence variation of Bioko and São Tomé, Gulf of Guinea. *Ann Hum Genet* 61:507–518.

- Nebel A, Filon D, Brinkmann B, Majumder P, Faerman M, Oppenheim A. 2001. The Y chromosome pool of Jews as part of the genetic landscape of the Middle East. *Hum Genet* 107:630–641.
- Neves CA. 1989. S. Tomé e príncipe na segunda metade do séc, 1st ed. XVIII. Coleção Memórias 2. Funchal: Centro de Estudos de História do Atlântico.
- Peres D. 1960. História dos Descobrimentos Portugueses. Coimbra. 2nd ed. 211p.
- Romana H. 1997. São Tomé e Príncipe. Elementos para uma análise antropológica das suas vulnerabilidades e potencialidades. Lisboa: Instituto Superior de Ciência Sociais e Políticas, Universidade Técnica de Lisboa.
- Rosser Z, Zerjal T, Hurler M, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbuiani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley D, Brede G, Cooper G, Corte-Real H, Knijff P, Decorte R, Dubrova Y, Evgrafov O, Gilissen A, Glisic S, Golge M, Hill E, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, Kravchenko S, Krumina A, Kucinskas V, Lavinha J, Livshits J, Malaspina P, Maria S, McElreavey L, Meitinger T, Mikelsaar A, Mitchell R, Nafa K, Nicholson J, Norby S, Pandya A, Parik J, Patsalis P, Pereira L, Peterlin B, Pielberg G, Prata M, Previdere C, Roewer L, Rootsi S, Rubinsztein D, Saillard J, Santos F, Stefanescu G, Sykes B, Tolun A, Villems R, Tyler-Smith C, Jobling M. 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* 67:1526–1543.
- Schneider S, Kueffer JM, Roessli D, Excoffier L. 2002. Arlequin: A software for population genetic data analysis. Switzerland: Genetics and Biometry Laboratory, University of Geneva.
- Scozzari R, Cruciani F, Santolamazza P, Malaspina P, Torroni A, Sellitto D, Arredi B, Destro-Bisoli 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.
- Seibert G. 1998. A Questão da Origem dos Angolares de São Tomé. *Brief Papers* n°5/98. Lisboa: CESA.
- Semino O, Magri C, Benuzzi G, Lin A, Al-Zahery N, Battaglia V, Maccioni L, Triantaphyllidis C, Shen P, Oefner P, Zhivotovskiy L, King R, Torroni T, Cavalli-Sforza L, Underhill P, Santachiara-Benerecetti S. 2004. Origin, diffusion and differentiation of Y-chromosome haplogroups E and J: Inferences on the Neolithization of Europe and later migratory events in the Mediterranean area. *Am J Hum Genet* 74:1023–1034.
- 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* 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.
- Shen P, Wang F, Underhill PA, Franco C, Yang WH, Roxas A, Sung R, Lin AA, Hyman RW, Vollrath D, Davis RW, Cavalli-Sforza LL, Oefner PJ. 2000. Population genetic implications from sequence variation in four Y chromosome genes. *Proc Natl Acad Sci USA* 97:7354–7359.
- Tenreiro F. 1961. A ilha de São Tomé. Memórias da Junta de Investigação do Ultramar. 2nd ed. 24. Lisboa.
- Tomas G, Seco L, Seixas S, Faustino P, Lavinha J, Rocha J. 2002. The peopling of Sao Tome (Gulf of Guinea): origins of slave settlers and admixture with the Portuguese. *Hum Biol* 74:397–411.
- Trovoada MJ, Alves C, Gusmão L, Abade A, Amorim A, Prata MJ. 2001. Evidence for population sub-structuring in São Tomé e Príncipe as inferred from Y-chromosome STR analysis. *Ann Hum Genet* 65:271–283.
- Trovoada MJ, Pereira L, Gusmão L, Abade A, Amorim A, Prata MJ. 2004. Pattern of mtDNA variation in three populations from Sao Tome e Principe. *Ann Hum Genet* 68:40–54.
- Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-Sforza LL, Oefner PJ. 1997. Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res* 7:996–1005.
- Underhill PA, 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.
- Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, Kauffman E, Bonne-Tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, Kidd JR, Mehdi SQ, Seielstad MT, Wells RS, Piazza A, Davis RW, Feldman MW, Cavalli-Sforza LL, Oefner PJ. 2000. Y chromosome sequence variation and the history of human populations. *Nat Genet* 26:358–361.
- Y Chromosome Consortium. 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* 12:339–348.