

## Insights from pattern of mtDNA variation into the genetic history of São Tomé e Príncipe

Maria Jesus Trovoada<sup>a,b</sup>, Luísa Pereira<sup>a</sup>, Leonor Gusmão<sup>a</sup>,  
Augusto Abade<sup>b</sup>, António Amorim<sup>a,c</sup>, Maria João Prata<sup>a,c,\*</sup>

<sup>a</sup>*IPATIMUP, Institute of Pathology and Immunology of University of Porto, Rua Dr. Roberto Frias, s/n. 4200-465, Oporto, Portugal*

<sup>b</sup>*Department of Anthropology, University of Coimbra, Coimbra, Portugal*

<sup>c</sup>*Faculty of Sciences, University of Porto, Oporto, Portugal*

---

**Abstract.** Sequence data from the hypervariable segments I (HVS-I) and II (HVS-II) was obtained for 30 Angolares, 35 Forros and 38 Tongas, three self-reported ethnic groups from São Tomé e Príncipe, an African archipelago (Gulf of Guinea) whose settlement begun in the late 15th century. The repertory of mtDNA lineages denoted a fully African maternal pool primarily arisen from a Central/Southwestern substratum. The absence of any lineages of putative European descent means that the European impact at the mitochondrial pool was virtually nil. Angolares showed a clear reduction of mtDNA diversity and a slight genetic differentiation relatively to Tongas or Forros, whereas the two last groups did not present any signs of genetic boundaries between each other. The data now obtained reinforce the depiction of genetic substructuring in São Tomé e Príncipe previously derived from Y-chromosome STRs. © 2003 Elsevier B.V. All rights reserved.

*Keywords:* MtDNA; Haplogroups; São Tomé e Príncipe; Population groups

---

### 1. Introduction

São Tomé e Príncipe is a small archipelago located on the Equator (Gulf of Guinea) that was uninhabited when first visited by Portuguese navigators in the early 1470s. Shortly afterward, the settlement was initiated with European colonist, mainly Portuguese, and an overwhelming number of slaves brought from the African Western mainland.

Among present day São Tomé e Príncipe inhabitants, the existence of three distinct population groups, Angolares, Forros and Tongas, is a common perception. Forros are referred to be “the sons of the land”, the more ancient African inhabitants of the archipelago [1]. Tongas are the descendents from African contract labourers who, from the middle of the 19th century on with the development of the coffee and cocoa economy, massively entered the archipelago. The origin of Angolares is less clear notwithstanding the fact that the identification as a distinct community has the stronger

---

\* Corresponding author. IPATIMUP, Institute of Pathology and Immunology of University of Porto, Rua Dr. Roberto Frias, s/n. 4200-465, Oporto, Portugal. Tel.: +351-225570700; fax: +351-225570799.

*E-mail address:* mprata@ipatimup.pt (M.J. Prata).

support from the geographical confinement to the southeastern tip of São Tomé, from cultural and livelihood traditions or from the own language.

Here we present sequence data from mtDNA detected by surveying HVS-I and HVS-II in individuals from the three São Tomean groups.

## 2. Material and methods

Blood samples were collected from 30 self-reported Angolares, 35 Forros and 38 Tongas. DNA was extracted from whole blood by standard phenol chloroform or Chelex methodologies. Primers, PCR and sequencing strategy as described in Ref. [2]. PCR products purified with Microspin™ S-300 HR columns (Pharmacia). Sequence reactions carried out with the kit Big-Dye™ Terminator Cycle Sequencing Ready Reaction (AB Applied Biosystems) and then a protocol based on MgCl<sub>2</sub>/ethanol precipitation for purification of samples. Automatic sequencer: ABI Prism 3100 Genetic Analyzer. Sequence analysis: DNA Sequencing Analysis 3.7 software (AB Applied Biosystems).

## 3. Results and discussion

Among the 103 São Tomeans analysed, 71 different sequences were found (Fig. 1) belonging to 32 distinct haplogroups all specific from sub-Saharan, which testifies a full matrilineal African background. Previously we have detected some Y-STR haplotypes of likely European ancestry [3]. The combined results indicate that the European component of present São Tomé e Príncipe population was essentially mediated by males, signalling therefore the biased mating pattern that has prevailed in the archipelago.

Within the overall sample, four haplogroups, L1b, L1c, L2a1 and L3e1\*, attained considerable high frequencies representing 56% of the sequences. The four haplogroups are common in West/Central Africa. Forros and Tongas, basically share the same mtDNA profile characterised by the presence of many different haplogroups. In contrast, the Angolares show a reduced number of different haplogroups and a particularly high frequency of L3e1\* and L1c lineages.

Diversity levels in Forros, Tongas and in the overall São Tomean sample (Table 1) are very similar to those registered in mainland populations from the cluster of Central/Western ones at which São Tomé e Príncipe geographically belongs. Distinctly, Angolares show lower diversity, particularly concerning mean heterozygosity and proportion of different haplotypes, resembling what happens in Western African islands such as Cabo Verde NW or Bioko.

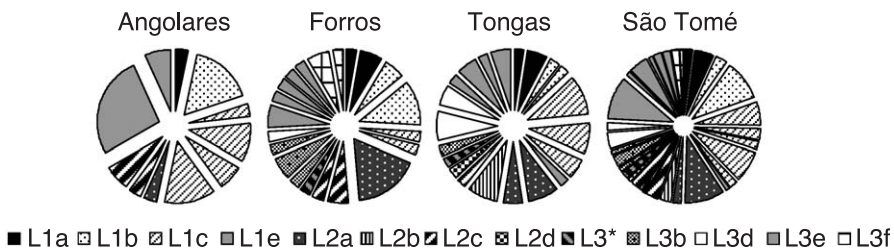


Fig 1. mtDNA haplogroup profiles in Angolares, Forros, Tongas and in overall São Tomé.

Table 1  
Diversity and neutrality measures for HVS-I in African populations

Population/group	$N^a$	$K (K/N)^b$	$H (S.E.)^c$	$M^d$	Fu's $F_s$
Morocco [4]	32	29 (90.6)	$0.988 \pm 0.014$	5.84	-25.01*
Niger [5]	34	30 (88.3)	$0.991 \pm 0.010$	6.19	-24.70*
Nigeria [5]	60	38 (63.3)	$0.972 \pm 0.010$	6.82	-23.15*
Senegal [6]	110	46 (41.8)	$0.963 \pm 0.008$	6.23	-24.53*
C.Verde NW [7]	108	28 (25.9)	$0.908 \pm 0.013$	5.84	-5.58
C.Verde SE [7]	184	101 (54.9)	$0.984 \pm 0.003$	6.29	-24.85*
Mbuti [8]	13	5 (38.5)	$0.756 \pm 0.097$	7.13	3.76
Biaka [8]	17	8 (47.1)	$0.890 \pm 0.043$	7.81	1.67
Bioko [9]	45	16 (35.6)	$0.910 \pm 0.020$	7.09	-0.55
Angolares	30	16 (53.5)	$0.913 \pm 0.037$	8.74	-1.34
Forros	35	29 (82.9)	$0.988 \pm 0.010$	7.52	-18.66*
Tongas	38	28 (73.7)	$0.984 \pm 0.009$	8.55	-12.61*
São Tomé	103	61 (59.2)	$0.985 \pm 0.004$	8.36	-24.68*
Kenya [4]	36	32 (88.9)	$0.991 \pm 0.010$	9.66	-20.75*
Mozambique [2]	109	49 (45.0)	$0.960 \pm 0.008$	7.78	-23.62*

<sup>a</sup>  $N$  = sample size. <sup>b</sup>  $K$  = number of different sequences and percentage of sample size ( $N$ ). <sup>c</sup>  $H$  = sequence diversity  $\pm$  standard error. <sup>d</sup>  $M$  = average number of pairwise differences.

\*  $P < 0.05$ .

Analysis of population pairwise  $F_{ST}$ s revealed statistically significant differences between Angolares and Forros ( $F_{ST} = 0.037$ ,  $P = 0.04$ ) or Angolares and Tongas ( $F_{ST} = 0.030$ ,  $P = 0.04$ ); while Tongas and Forros did not showed statistical differences ( $F_{ST} = 0.01$ ,  $P = 0.14$ ). The small level of genetic distinctiveness of Angolares can be attributable to genetic drift since, until recently, they lived in relative isolation in the southeastern tip of São Tomé e Príncipe, which is also reflected in their cultural and livelihood traditions or even in their own language.

For a small archipelago with a so recent population history, this is really a quite remarkable finding. In São Tomé e Príncipe, a complex net of social and historical forces must have interacted in order that a self-reported community like Angolares, with a presumable distinct origin from other São Tomé e Príncipe inhabitants, has retained and still preserves signs of genetic distinctiveness, despite the sharing of a so restricted territory with other people.

## Acknowledgements

Financial support was granted by *Fundação para a Ciência e a Tecnologia, Programa Operacional Ciência, Tecnologia e Inovação (POCTI)*.

## References

- [1] I.C. Henriques, São Tomé e Príncipe. A invenção de uma sociedade, Vega e Autor, 2000.
- [2] L. Pereira, et al., *Ann. Hum. Genet.* 65 (2001) 439–458.
- [3] M.J. Trovoada, et al., *Ann. Hum. Genet.* 65 (2001) 271–283.
- [4] J.C. Rando, et al., *Ann. Hum. Genet.* 62 (1998) 531–550.
- [5] E. Watson, et al., *Am. J. Hum. Genet.* 61 (1997) 691–704.
- [6] L. Graven, et al., *Mol. Biol. Evol.* 12 (1995) 334–345.
- [7] A. Brehm, et al., *Ann. Hum. Genet.* 66 (2002) 49–60.
- [8] L. Vigilant, et al., *Science* 253 (1991) 1503–1507.
- [9] E. Mateu, et al., *Ann. Hum. Genet.* 61 (1997) 507–518.