International Congress Series 1239 (2003) 541–545

Mitochondrial DNA variability patterns in Southeast Africa and forensic implications

A. Salas, T. De La Fé, B. Sobrino, M.V. Lareu, A. Carracedo*

Faculty of Medicine, Institute of Legal Medicine, Unit of Forensic Genetics, University of Santiago de Compostela, c/San Francisco s/n, Santiago de Compostela 15786, Spain

1. Introduction

The first hypervariable region and several restriction fragment polymorphisms in mitochondrial DNA were investigated in 308 individuals belonging to 16 different populations from Southeast Africa, all of them Bantu speaking (see Fig. 1). A total of 30 available African populations (or populations with an important African influence) as well as other non-African ones were used for comparison. Preliminary results are shown in the present work. High diversity values were found for the samples analysed in this work, in comparison with other African samples. Nucleotide mismatch distributions are rugged and multimodal, which could reflect the scenario of stationary populations rather than one of expansion. Phylogenetic reconstruction allowed us to infer that the Southeast African populations are distant from other population groups, with the ancient ones (Pygmies and !Kung) more closely related.

The results obtained are of special forensic interest because they produced new mtDNA patterns from human populations completely unknown (from a genetic point of view) until now, and to increment our mtDNA databases for forensic purposes.

2. Nucleotide and sequence diversity

Sequence diversity and nucleotide diversity were computed for all the populations analysed and compared with other African ones. Almost all the populations analysed here showed similar values for these indices, the lowest values were in the Makonde population ($\pi = 0.0187$ and $d = 6.830$) and the highest in the Shona ($\pi = 0.0277$ and $d = 10.177$), and they were, on average, slightly higher than those found for other African populations. When

* Corresponding author. Tel.: +34-981-582327; fax: +34-981-580336.
E-mail address: apimlang@usc.es (A. Carracedo).

PII: S0531-5131(02)00396-5

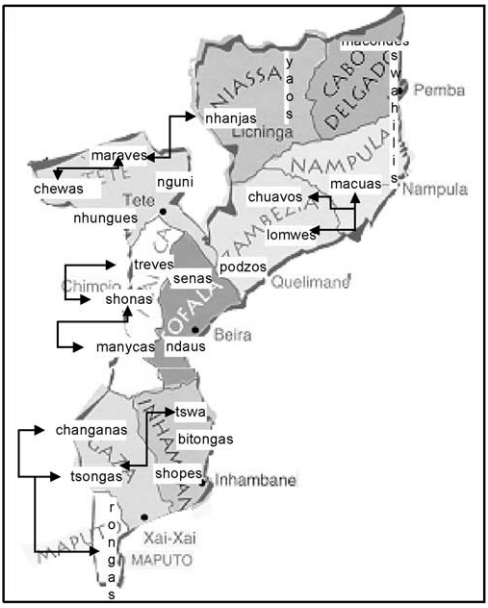

Fig. 1. Map of Mozambique indicating the origin of the samples used in the present work.

these populations are considered as a group, sequence diversity is 0.974, while nucleotide diversity is 0.0247. Higher values are found only in the Fang, Shongai, Turkana, and African Brazilian populations (this last, only for its nucleotide diversity value). In addition, these values do not show a geographical pattern of variability for our samples.

3. mtDNA haplogroup distribution

All the sequences analysed in the present work could be included in a typical African haplogroup. The most salient features of the haplogroup distribution in the 308 assigned sequences are the relatively high frequencies of L2 haplogroup (41%; high frequency of L2a), and haplogroup L1a (28%). Other haplogroups were found, some of them with significant frequencies: L3*, L1c, L3a1 and L1d.

4. Nucleotide pairwise difference distribution

According to the model proposed by Roger and Harpending (1992), populations with a constant population size during their history should show bumpy and irregular distributions while expanding populations should show unimodal and bell-shaped ones. All the populations analyzed in the present work (except Ronga) showed clear multimodal patterns, which probably means that these populations have maintained constant popula-

tion sizes for centuries. Analysis performed on haplogroups yielded similar results but with a higher tendency towards unimodal distributions.

5. Inter- and intrapopulation variability through AMOVA analysis

AMOVA results for 30 African populations show interpopulation variability of 20% ($p < 0.001$), while European populations show lower values (1%; [1,2]) and Asian

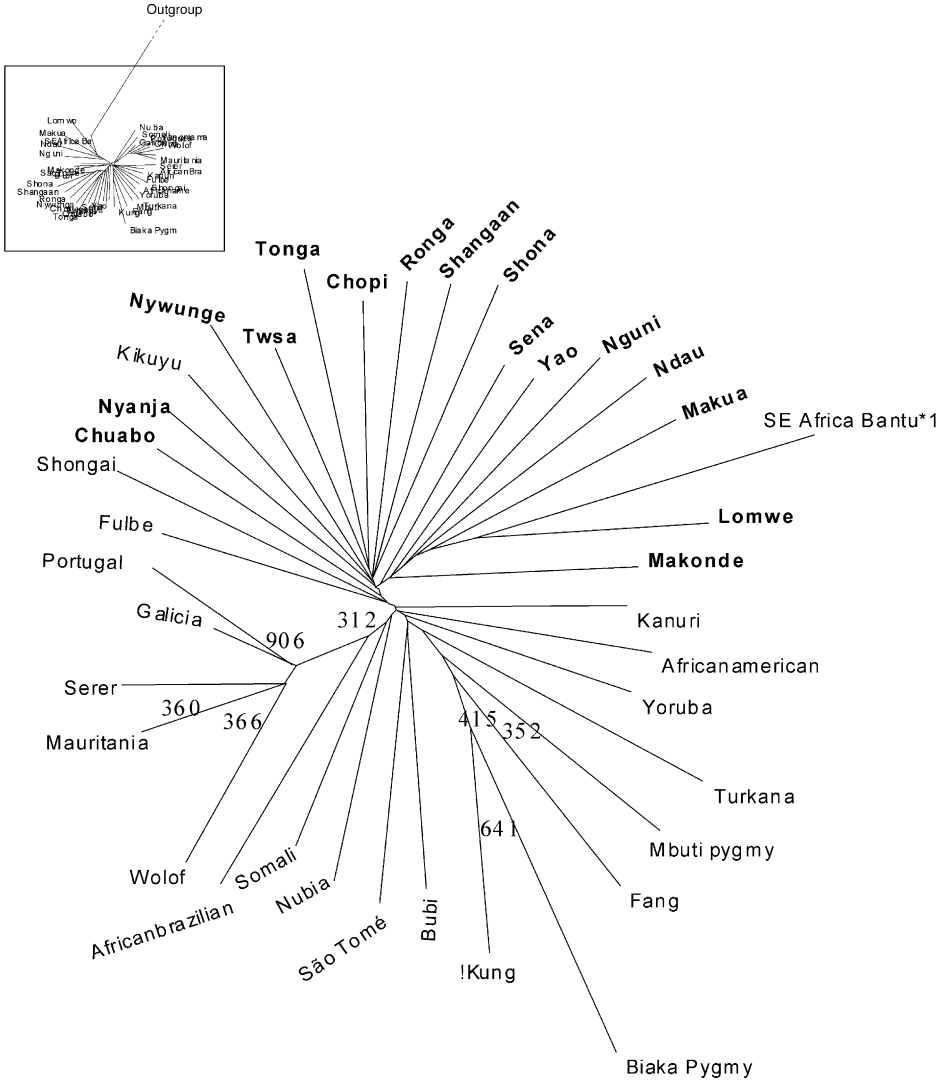

Fig. 2. Neigbour joining tree of African populations.

populations give intermediate values of approximately 5%. These results show that African populations are a heterogeneous group with a level of variability much higher than other populations from other continents.

It is interesting to note that the particularly long and robust branches in one edge separating the two southeast Bantu samples from the other African populations (Fig. 2), both for those analyzed in the present work and those analysed by Soodyall et al. [3]. The long distance separating both populations can be explain if we consider that Soodyall's sample is a subset of a general population containing a specific polymorphism.

What is more remarkable is the fact that the two populations are separated from the main African group by a long distance. This joins the ancient African populations of the Pygmies and the !Kung but they are completely separated from the tight Central and West African cluster. The two Caucasoid samples (Galicians and Portuguese) are placed in the opposite end of the tree linked by a high bootstrap support closely related with the samples from North and West Africa, while the populations from Central Africa are also closely related in a clade.

6. Discussion

The results obtained in the present work show that Southeast Africa is within the mtDNA African genetic pool. However, Bantu-speaking groups must be considered as having their own genetic identity, clearly separately from the rest of the African populations, and probably immediately related to the ancient groups of Pygmies and !Kung, while clear differences can be established with West African populations.

Although linguistic studies have provided vital evidence for the reconstruction of the African story, such as the putative spread of the Bantu people and their link with their origin in Central Africa (overall in a continent where in many regions exist a complete lack of written records), genetic analysis can be considered as an important tool to improve our knowledge about these ancient populations.

Finally, African slavery implied the movement of more than 28 million Africans, mainly between the XVII and XVIII centuries, from their native continent to America, almost all from West and Central Africa; however, a significant quantity of slaves came from the southeast of Africa and other continental regions. Knowledge of the African mtDNA genetic patterns will be of special interest to reconstruct the origin of millions of African descendants living in present-day America and the Caribbean, and to recover their cultural identity lost by one of the greatest human tragedies of recent human history.

Acknowledgements

This work has been possible thanks to the help of the Xunta de Galicia (XUGA PGIDT01PX120806PR) and the Ministerio de Educación y Ciencia (BIOT 2000/PC092). We thank the collaboration of the University of Eduardo Mondlane and the Hospital Militar of Maputo (Mozambique) and also to E. Valverde y J. Cobreiro. A.S. has a research contract with the University of Santiago de Compostela.

References

[1] T. Melton, M. Wilson, M. Batzer, M. Stoneking, Extent of heterogeneity in mitochondrial DNA of European populations, J. Forensic Sci. 42 (1997) 437–446.
[2] A. Salas, V. Lareu, F. Calafell, J. Bertranpetit, A. Carracedo, mtDNA hypervariable region II (HVII) sequences in human evolution studies, Eur. J. Hum. Genet. 8 (2000) 964–974.
[3] H. Soodyall, L. Vigilant, A.V. Hill, M. Stoneking, T. Jenkins, mtDNA control-region sequence variation suggests multiple independent origins of an "Asian-specific" 9-bp deletion in sub-Saharan Africans, Am. J. Hum. Genet. 58 (1996) 595–608.