

# Y-chromosome diversity in Bantu and Pygmy populations from Central Africa

G. Berniell-Lee, E. Bosch, J. Bertranpetit, D. Comas \*

*Unitat de Biologia Evolutiva, Facultat de Ciències de la Salut i de la Vida Universitat Pompeu Fabra, Barcelona, Spain*

---

**Abstract.** Linguistic and archaeological data have placed the Bantu homeland near the border between Cameroon and Nigeria, where the Bantu farmers are thought to have migrated from approximately 5000 years ago, spreading out over most of sub-Saharan Africa. In order to provide genetic insights into the so-called Bantu expansion, we have typed 18 well-characterized Y-chromosome short tandem repeats (STRs) in 24 Bantu and 3 Pygmy populations from Gabon and Cameroon. The results have shown not only that Bantu populations are very homogenous, but also that they are statistically very different from Pygmies and other African and non-African populations. High frequencies of the modal Bantu haplotype and of its one-step neighbours previously described have been found in all the Bantu populations sampled. Some traces of these haplotypes have also been found in Pygmies, indicating a possible Bantu-to-Pygmy flow of paternal lineages. No evidence has been found of the Fang being non-Bantu, despite their claim of having a semitic origin in Egypt. They do, however, show an intermediate allele (13.2) for DYS 385, whose date of appearance (approximately 5500 years ago) happens to coincide with the Bantu expansion. © 2005 Published by Elsevier B.V.

*Keywords:* Y chromosome; STR; Bantu expansion; Pygmy population; Human diversity

---

## 1. Introduction

Linguistic and archaeological data have placed the Bantu homeland near the border between Cameroon and Nigeria where the Bantu farmers are thought to have migrated from approximately 5000 years ago, spreading out over most of sub-Saharan Africa. Their expansion is thought to have followed two main waves: an eastern wave and a western

---

\* Corresponding author.

*E-mail addresses:* gemma.berniell@upf.edu (G. Berniell-Lee), david.comas@upf.edu (D. Comas).

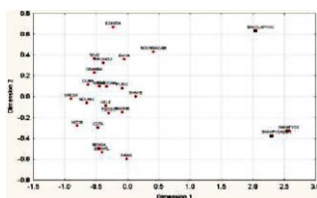


Fig. 1. Multidimensional scaling (MDS) plot based on the genetic distance matrix of 17 Y-chromosome STRs.

wave. In order to provide insights into the western migration wave from a genetic point of view, we have typed 18 well-characterized Y-chromosome STRs in a total of 873 samples from the Central African locations of Gabon and Cameroon.

## 2. Materials and methods

The 18 Y-STRs were amplified in the form of three multiplex reactions: MSI multiplex [2], CTS multiplex [1] and EBF multiplex [2]. Amplicon size was analyzed using GeneMapper Software® v3.7. A series of statistical analyses were performed using Arlequin® software v2.0 [4]. The total number of haplotypes and the frequencies of the modal Bantu haplotype described by Thomas et al. [5] and of its one-step neighbours described by Pereira et al. [3] was calculated by a simple counting scheme. Genetic distances between populations were represented in a multidimensional scaling plot (MDS) using STATISTICA 6.0. The genetic relationship between haplotypes (17Y-STR haplotypes not including DYS 385) was analyzed by a series of minimum spanning networks constructed using Network 4.1.1.1.

## 3. Results

A total of 728 different haplotypes were found, 46 of these being present in two or more populations. High frequencies of the modal Bantu haplotype described by Thomas et al. [5] and of its one-step neighbours described by Pereira et al. [3] were found in all 24 Bantu populations. The modal Bantu haplotype was also present in two Pygmy samples from Gabon and one of its one-step neighbours in one Pygmy sample from Cameroon. The MDS plot in Fig. 1 shows all Bantu populations clustering together and clearly separated from the Pygmy populations (Fig. 1).

The other MDS plot constructed using previously published data shows Bantu populations clustering together with the rest of the populations scattered in the plot (data not shown). The results obtained for the AMOVA analyses can be seen in Table 1.

Table 1  
AMOVA results

|                 | Among groups         | Among populations   | Within groups |
|-----------------|----------------------|---------------------|---------------|
| All populations |                      | 7% ( $p < 0.0001$ ) | 93%           |
| Bantu           |                      | 4% ( $p = 0.0694$ ) | 96%           |
| Pygmy           |                      | 5% ( $p < 0.0001$ ) | 95%           |
| Bantu–Pygmy     | 21% ( $p < 0.0001$ ) | 3% ( $p < 0.0001$ ) | 76%           |

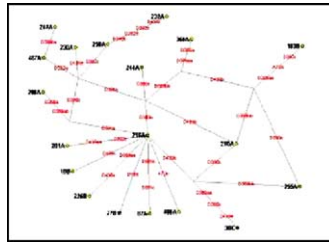


Fig. 2. Phylogenetic network of Y-chromosome haplotypes carrying the allele 13.2 of the DYS 385 STR.

The network representing the relationship between the haplotypes showing the intermediate repeat 13.2 for DYS 385 is starlike, every haplotype being different (Fig. 2).

The estimated date of appearance for the 13.2 allele (using STR mutation rate reported by Zhivotovsky et al. [6]) was found to be 5500 years ago, coinciding with the hypothesized date for the Bantu expansion.

#### 4. Discussion and conclusions

The results obtained indicate that Bantu populations from Central Africa are very homogenous, while Pygmy populations seem to be quite diverse. They also suggest that Bantu populations are statistically different from Pygmy populations. The fact that the modal Bantu haplotype and one of its one-step neighbours were found in some Pygmy samples suggests some admixture events between Bantus and Pygmies must have taken place. We found no evidence of any of the populations studied having any origin other than Bantu. However, the presence in the Fang population of the intermediate allele 13.2 for locus DYS 385 suggests their origin should be further explored.

#### References

- [1] Q. Ayub, et al., Identification and characterisation of novel human Y-chromosomal microsatellites from sequence database information, *Nucleic Acids Res.* 28 (2000) e8.
- [2] E. Bosch, et al., High resolution Y chromosome typing: 19 STRs amplified in three multiplex reactions, *Forensic Sci. Int.* 125 (2002) 42–51.
- [3] L. Pereira, et al., Bantu and European Y-lineages in sub-Saharan Africa, *Ann. Hum. Genet.* 66 (2002) 369–378.
- [4] S. Schneider, D. Roessli, L. Excoffier, Arlequin ver. 2.0: A software for population genetic data analysis, 2.0 edn. Genetics and Biometry Laboratory, University of Geneva, Switzerland, 2000.
- [5] M.G. Thomas, et al., Y chromosomes traveling south: the Cohen modal haplotype and the origins of the Lemba—the “Black Jews of Southern Africa”, *Am. J. Hum. Genet.* 66 (2000) 674–686.
- [6] L.A. Zhivotovsky, P.A. Underhill, C. Cinnioglu, The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time, *Am. J. Hum. Genet.* 74 (2004) 50–61.