

MtDNA VARIATION IN CENTRAL AFRICA : A MICRO-EVOLUTIONARY STUDY IN BANTU-SPEAKING POPULATIONS FROM GABON

Francesca Luca (1), Lolke van der Veen (2), Emeline Andre (1), Patrick Mouguiama-Daouda (3), Lucas Sica (4), Jean-Marie Hombert (2), Lluís Quintana-Murci (1)
quintana@pasteur.fr

(1) CNRS URA1961-Institut Pasteur, Paris, France ; (2) CNRS UMR5596, ISH, Lyon, France
(3) University Omar Bongo, Libreville, Gabon ; (4) CIRMF, Franceville, Gabon

The great majority of sub-Saharan Africans speak one of > 500 closely related Bantu languages. The highest linguistic diversity of the Bantu family is found in Nigeria and western Cameroon, suggesting this area as the « core » region of the Bantu dispersals starting 5000 YBP. The Ogooué-Ivindo region, covering Gabon and parts of Congo-Brazzaville, is located in the southern lobe of the putative Bantu homeland and it has been extensively studied from a linguistic standpoint. Central Africa is poorly sampled and mtDNA variation in this region remains largely uncharacterised. We have therefore defined mtDNA variation 308 individuals from 7 Bantu-speaking populations from the Ogooué-Ivindo area in order to (i) determine the maternal variation in central African populations and compare it to the global African landscape and (ii) to contrast the linguistic classification of these populations with genetic data. All individuals, which belong to Fang, Galoa, Mitsogo, Kele, Punu, Obamba and Teke populations, were analysed for HVS-I sequence variation and a number of coding-region sites. Overall, high sequence and lineage diversity values were observed in all populations. The amount of variation due to inter-population diversity turned out to be very low (~1%), indicating that virtually all variation is explained by intra-population diversity. Standard neutrality tests suggested that all populations, except the Galoa, exhibit signals of population growth. The two most prevalent lineages in the Central African populations studied are L1c and L3e. The origin of L1c lineage remains controversial since it is found at high frequencies in African-Americans, but in Africa it has been observed in some central and west Africans. It has been suggested that it could have originated in Central Africa, since west African sequences tend to be derived from central African ones. L1c lineage is present at highest frequencies in our populations (20-40%) and therefore, strongly suggests a central African origin for it. The second most prevalent lineage in our samples is L3e. It is present at low frequencies in west and east Africans but it has been found at higher frequencies in a small central African sample and in southeastern populations. Both its distribution and internal diversity suggest a central African origin of L3e that was probably subsequently dispersed into the southeast via the Bantu-expansions. Our data have also shed light on the linguistic classification of the different Bantu-speaking groups of Gabon. Detailed linguistic comparisons have evinced two major linguistic clusters: a northern cluster, comprising the B10 and B30 groups represented by the Galoa and Mitsogo respectively, and a southern cluster, comprising the B40, B60 and B70 groups represented by the Punu, Obamba and Teke respectively. Our genetic data strongly support the linguistic grouping of the northern cluster, while the southern cluster seems to be less supported. The latter observation could be due to long-term linguistic convergence as a result of prolonged contact with few genetic exchanges among these populations. In addition, the origin of the Fang population, who claims a Semitic origin in Egypt and not to be related with their Bantu-speaking neighbours, has been clarified. No detectable traces of an Egyptian origin were observed in this population, who otherwise exhibits strong genetic affinities with neighbouring Bantu-speaking groups, in contrast with Fang cultural tradition. However, Y-chromosome data is needed to conclusively clarify the origins of this population.