

C. Xavier¹, M. Bodner¹, S. Nagl¹, G. Huber¹, D. Silva², A. Casas-Vargas³, W. Usaquén³, L. Gusmão^{2,4} and W. Parson^{1,5}

¹ Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria

² DNA Diagnostic Laboratory (DDL), State University of Rio de Janeiro (UERJ), Rio de Janeiro, Brazil

³ Group of Population Genetics and Identification, Genetics Institute, National University of Colombia, Bogotá, Colombia

⁴ Institute of Pathology and Immunology of the University of Porto (IPATIMUP), Porto, Portugal

⁵ Forensic Science Program, The Pennsylvania State University, PA, USA

catarina.gomes@i-med.ac.at

Introduction

The colonization of South America remains an unveiled topic in what regards the number and the migration routes taken within the continent. Even though the scientific community agrees in the existence of two major routes (one following the Andean chain and another into the amazonian forest), there might have been other smaller routes that contributed to the actual dispersion of Native Americans.

The Arawakan Diaspora is considered the South American version of a Neolithic expansion that has ultimately resulted in the contemporary wide distribution of Arawak speaking groups not only within the continent but also on circum-Caribbean islands. The Arawakan share several distinct characteristics such as their agricultural practices, the establishment of highly complex societies and socio-political alliances [1,2]. Two Arawakan populations have been studied here.

1) The Terena inhabit Mato Grosso do Sul state in Brazil (Fig.1). Before the 18th century, the Terena lived in the Paraguayan Chaco where they forged an alliance with Mbayá-Guaicuru tribes for protection during the Paraguayan war [3].

2) The Wayúu are settled in the La Guajira peninsula in northern Colombia. They are known for their agricultural and livestock practices that allowed their survival in a very arid environment and to escape from the Spanish control [1].

The study of poorly characterized isolated groups from remote areas can enlighten the minor human colonization routes in South America and contributes to enlarge the size and quality of forensic databases such as EMPOP [4].

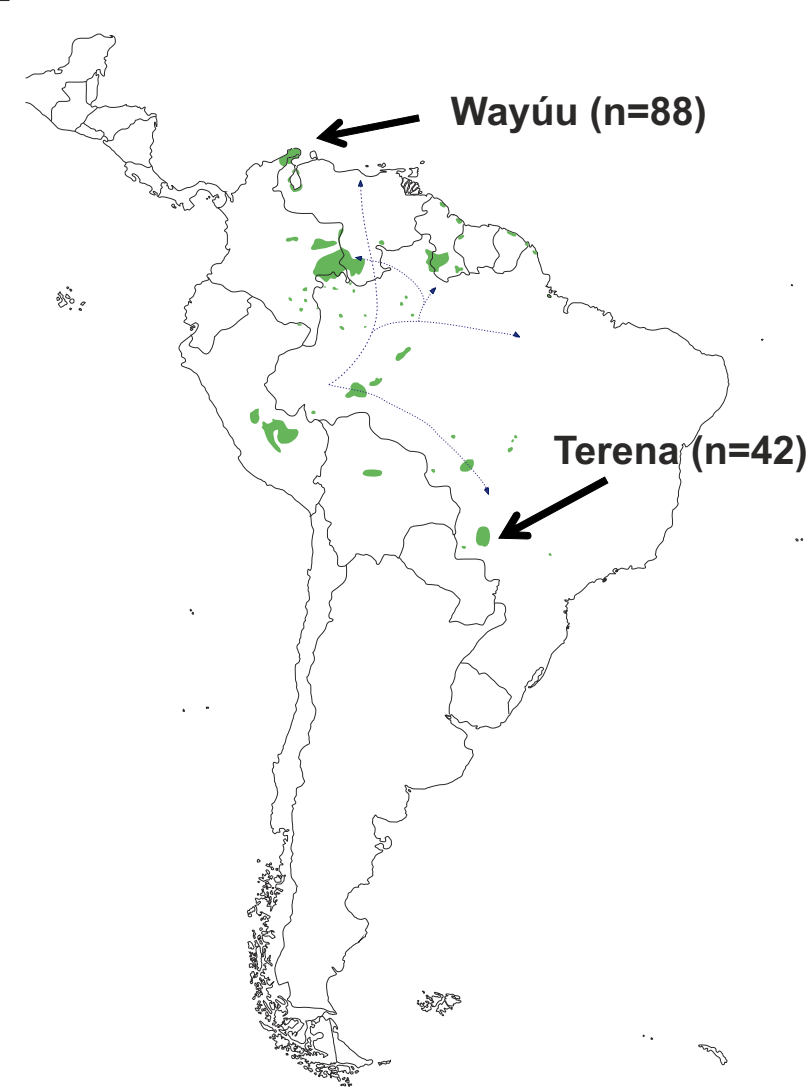


Fig.1 Present day distribution of Arawak speaking populations in South America (green), the most accepted Arawakan migratory routes (dark blue) and the two regions of sampling (black arrows).

Materials & Methods

The study was approved by Ethics Commissions. Sampling was performed under informed consent. The Terena samples were extracted using a phenol-chloroform protocol and the Wayúu samples were extracted using a Chelex based protocol. All samples were quantified using a qPCR method [5] and Sanger-sequenced for the full control region of the mtDNA following forensic highest quality protocols [6].

The sequences were compared to the revised Cambridge Reference Sequence (rCRS) [7] using the Sequencher 5.1® software (Gene Codes); haplotypes were determined by two separate analysts and haplogroups were assigned according to Phylotree (build 16) [8]. Haplogroup frequencies were calculated by direct counting, mtDNA diversities were calculated using Arlequin v.3.5 software [9] and median joining networks were designed using Network v.4.6.1.2 [10].

In order to perform a comparison study, we assembled other populations from literature that have a linguistic or geographical affiliation in common with our samples [11-13].

Discussion

The distribution of the pan-American mtDNA haplogroups (A, B, C and D) is not concordant within linguistic groups (Fig.2A), but rather indicates a geographical relation in the Wayúu group and other northwestern groups. In the southern populations, a pattern is seen in the groups from Argentinian Chaco but not in Brazil. However this might be a consequence of the limited amount of populations gathered in these regions or an effect of genetic drift acting in small groups (eg. Ache, Fig.2)

The results show that both groups have high levels of mtDNA diversity, high numbers of haplotypes and unique haplotypes (Table 1). No haplotypes are shared between the two studied populations.

A clear distinction within haplogroup B is visible in Fig.3B, that shows a separation of the haplotypes carrying the 498del polymorphism. Haplotypes presenting this polymorphism seem to be restricted to northeast South America and lower central American groups, independent of linguistic affiliation. A high frequency of sub-haplogroup B2d in the Wayúu and not in the other Arawakan groups analysed suggests the occurrence of gene flow between the first and neighboring populations. Therefore, a deeper insight into this lineage and the full mtGenome typing would be interesting for further understanding small scale migratory routes within this specific region. On the other hand, haplogroup A (Fig.3A) does not present a clear pattern of distribution.

Our results show the importance of mtDNA comparisons between tribal and admixed populations to enlighten the human migration history in South America. Nevertheless, high resolution mtDNA typing is of utmost relevance for improving the knowledge on the variation of mtDNA lineages at a regional scale.

Results

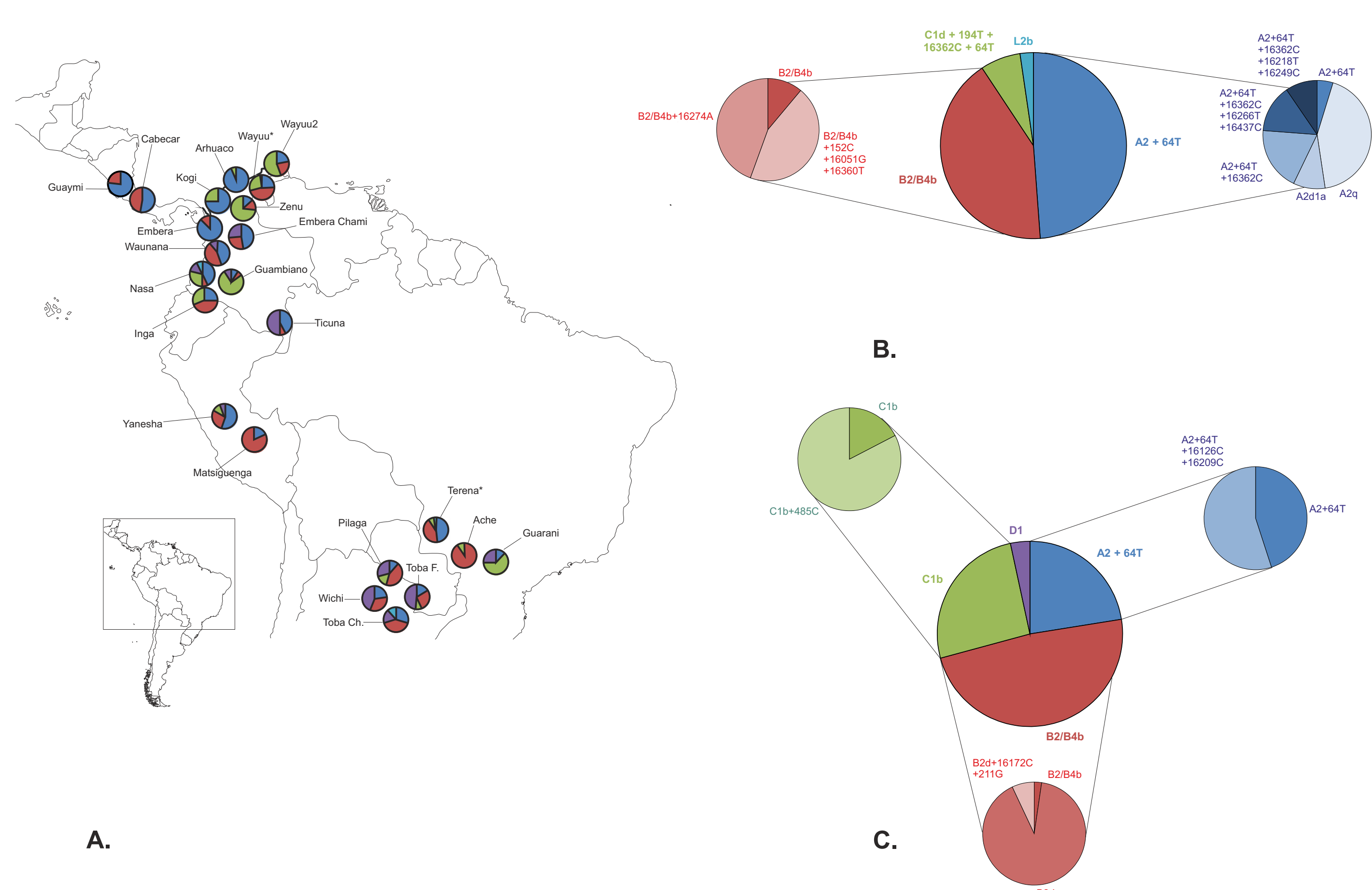


Fig.2A Overview of the pan-American haplogroup frequencies distribution in South America for the studied groups and for other relevant groups assembled from literature [11-13]. In detail, frequency distribution of the main haplogroups and sublineages found in Terena (2B) and Wayúu (2C). Color key as follows: blue - A2 lineages, red - B2/B4b lineages, green - C1 lineages, purple - D lineages and turquoise blue - others.

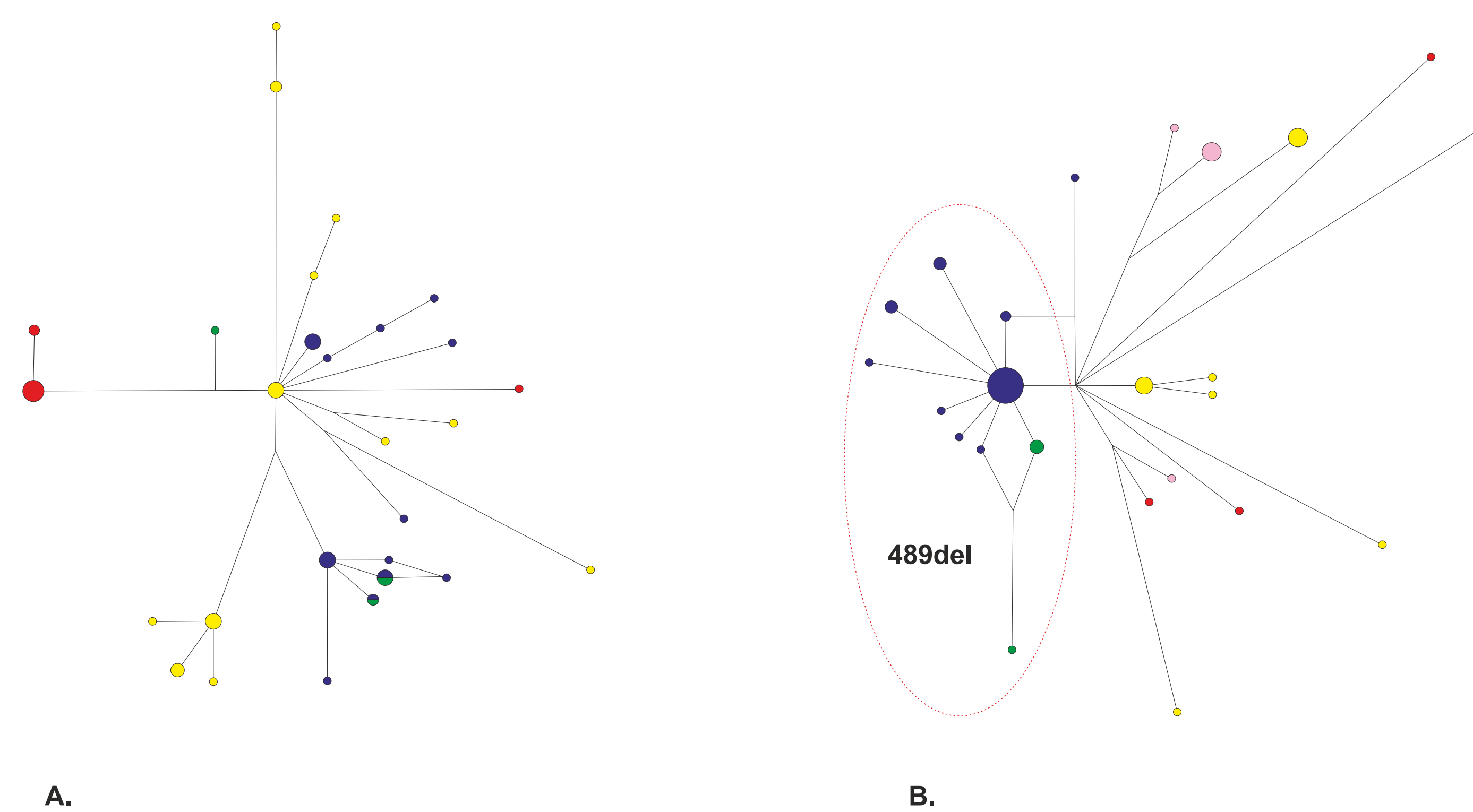


Fig. 3 Median Joining Network of the A2 (A) and B2/B4b (B) lineages found in the study groups and other arawakan groups assembled from the literature [12-13]. Each circle represents one haplotype and the size of the circle determines the number of samples with the same haplotype. Sizes of the branches vary with the number of differences present from one haplotype to the other. Color key: yellow - Terena (PS), dark blue - Wayúu (PS), green - Wayúu2, red - Yaneshá and pink - Matsiguenga.

	TERENA	WAYUU
Number of samples	42	83
Number of haplotypes	26 (61.9%)	39 (47%)
Number of unique haplotypes	18 (42.9%)	26 (31.3%)
Number of polymorphic sites	57	66
Haplotypic diversity	0.9698	0.9577
Nucleotide diversity	0.0127	0.0127

Table 1. Diversity indices calculated for the studied populations.

References & Acknowledgements

- [1] Hill JD & Santos-Granero F (2002) University of Illinois Press
- [2] Walker RS & Ribeiro LA (2011) Proc. of the Royal Society (online pub.)
- [3] Oberg K (1949), Brazil, Smithsonian Institution
- [4] Parson W & Dür A (2007) FSI:Genetics 1(2):88-92 empop.org
- [5] Niederstätter H et al (2007) FSI:Genetics 1(1):29-34.
- [6] Parson W & Bandelt HJ (2007) FSI:Genetics 1(1):13-19
- [7] Andrews RM et al (1999) Nature Genetics 23:147
- [8] Van Oven M et al (2009) Hum Mutat 30(2):386-394 www.phylotree.org
- [9] Excoffier L. et al (2010) Molecular Ecology 10:564-567
- [10] Network v. 4.6.1.0 www.fluxus-engineering.com
- [11] Sala A et al (2009) FSI: Genetics Supp Series 2:331-333
- [12] Yang et al (2010) Ann of Hum Gen 74:525-538
- [13] Sandoval et al (2013) PlosONE 8(9)

C. Xavier has a PhD grant (SFRH/BD/90873/2012) from Fundação para a Ciência e Tecnologia (FCT). M. Bodner was supported by Theodor Körner Fonds zur Förderung von Wissenschaft und Kunst.