

Inferring the genetic structure and the history of interaction between Andean and Amazonian human populations using genome-wide data

Victor Borda^{1,2}, Marilia Scliar¹, Mateus Gouvêa¹, Thiago Leal¹, Gilderlanio Araújo¹, Giordano Soares-Souza¹, Robert H Gilman^{3,4}, Heinner Guio², Eduardo Tarazona-Santos¹

¹ *Laboratorio de Diversidade Genetica Humana, Instituto de Ciencias Biologicas, Universidade Federal de Minas Gerais.* ² *Peruvian National Institute of Health, Ministry of Health, Perú.* ³ *Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205.* ⁴ *Laboratorio de Investigación de Enfermedades Infecciosas, Universidad Peruana Cayetano Heredia, 15102, Perú*

The peopling of the Americas implied a complex process in which human population entered into North America and migrated until the tip of Chile. In this process, the settlement of Amazon and Andean regions involved many demographic processes that have a strong influence on cultural and genetic flow. We use genome-wide information of several native populations from the Coast, Andes and Amazon Peruvian regions in order to disentangling the genetic structure and the historical relationships among them, as well as to reveal the genetic structure of biomedically relevant variants in these populations. We analyzed a dataset of 153 unrelated samples for 2.5 million of SNPs from 10 Natives Peruvian populations. This study involves collaboration between our group and the Peruvian National Institute of Health (INS). We use two different approaches to detect genetic structure: SNP and haplotype-based analyses. For SNP-based analysis we used PCA EIGENSOFT and ADMIXTURE and for haplotype-based we used CHROMOPAINTER and fineSTRUCTURE. SNP-based methods showed two groups. A first group included Andean populations (South-Central Andean Quechuas [SCAQ], Chopccas, Qeros, Aimaras and Uros) and Coastal population (Moches). The second group included the Amazon populations (Ashaninkas, Matsigenkas, Matses and Nahuas). The haplotype-based methods showed a fine differentiation within the Andean and Amazon populations. On the Andean region, we observed the separation of Quechua-speaking populations (SCQA and Chopccas) from the Aymaran-speaking (Aimaras and Uros) except for Qeros (Quechua-speaking), that showed more affinities with the Aymaran group. Also, both approaches showed some affinities of Moches (Shore) with Matses (North Amazon) and between High Amazon (Matsigenkas and Ashaninkas) with Aymaran speaking populations (Aimaras and Uros). We are also identifying variants that are highly differentiating between the Andes and Amazon that may be involved in adaptations to these environments, and using a network approach to define the genetic structure of GWAS-hits in these populations.

Funding support: The first author is part of the Postgraduate Program in Bioinformatics –UFMG and of the Programa Estudante Convênio Pós-Graduação of CAPES (PEC-PG). This study received financial support from CNPq and INS.