Assortative Mating in Brazilian Populations

Isabela Alvim, Hanaisa de Plá, Eduardo Martin Tarazona

Human Genetic Diversity lab, UFMG

The admixed Brazilian population shows a trend of marriages between individuals with similar socioeconomic status, educational level and genomic ancestry. This behavior violates the assumption of random marriages of several statistical models in genetics populations and medical genetics, which can lead to misleading results in quantitative genetics analysis and demographic inferences. Therefore, the knowledge about the trends of marriages in Brasil is crucial to the progress of Brazilian population genomic studies. In previous studies our group verified the occurrence of ancestry-assortment for the three Brazilian populations of the EPIGEN-Brasil (The Latin American initiative in population genomics and genetic epidemiology): Salvador – BA, Pelotas – RS, Bambuí – MG. In this project we investigated the patterns of ancestry-assortment for different socioeconomic status and educational levels in these populations. The 6.487 samples from the three population-based cohorts were genotyped for 2.5 million SNPs (single nucleotide polymorphism) and each cohort was stratified in 3 categories based on individual educational level. Homozygosity excess in SNPs that are highly differentiated between ancestral populations indicate ancestry-assortment, therefore we estimated the Spearman's rank correlation (rho) for the ancestry informativeness and the homozygosity excess estimated for each SNP. Our study finds that the ancestry-assortment is significantly affected by educational level. Salvador and Pelotas showed a crescent pattern with the most remarkable correlation in the high educational level (rho = 0.07, p-value < 2.2e-16; rho = 0.34, p-value < 2.2e-16 respectively) compared to the low (rho=0.01 , p-value < 2.2e-16; rho=0.11, p-value < 2.2e-16) and middle levels (rho = 0.05 p-value< 2.2e-16; rho = 0.23, p-value < 2.2e-16). The opposite is seen in Bambuí, that shows a more expressive evidence of ancestry-assortment at the lower educational level (rho = 0.14, p-value < 2.2e-16) compared to the middle and high levels (rho = 0.05, p-value<2.2e-16; rho = 0.03, p-value < 2.2e-16). Thus, our results enables that the adjustments of mathematical models for Brazilian population genomic studies consider those regional differences in the assortative mating.