Genomic Assortative Mating in Human Marriages

Guang Guo[1,2] , Lin Wang[3], Hexuan Liu[1] and Yilan Fu[1],

[1] *Department of Sociology and Carolina Population Center, the University of North Carolina, Chapel Hill*
[2]*Carolina Center for Genome Sciences, the University of North Carolina, Chapel Hill*
[3] *Center for Child and Family Policy, Duke University*

Correspondence author: Guang Guo, guang_guo@unc.edu

ABSTRACT

Using GWAS data and about 1,000 married couples from the Framingham Heart study, this analysis investigates genomic assortative mating in human marriages. The study calculates (1) a SNP-specific correlation for each of the 320,195 SNPs, averaging 1,023 married couples and (2) a married couple-specific correlation for each couple, averaging 287,295 SNPs. To avoid having the positive and negative assortments cancel each other out, we calculate two additional correlations for each of our couples. Permutation tests are performed to determine the statistical significance of the correlations. The same analysis is performed on full sibling pairs and parent-child pairs for quality control. Of the 320,195 SNP correlations, eight have a p-value $5x10^{-8}$ or smaller. These SNPs are all positively correlated for married couples, with a range of 0.16-0.27. Of these correlations, those within the genes that are found associated with BMI, height, risk behavior, or the human leukocyte antigen (HLA) system are further examined. For the half of all SNPs having a more positive assortment, married couples average a statistically significant correlation that is 0.001 higher than randomly paired individuals. Our simulation shows that this difference in correlation can be explained by married couples assorting on about 200 SNPs. Overall, our data suggest a small degree of genetic assortative mating at the allelic level in married couples of European origin who were born between 1910-1950 in the United States. Future studies need to consider a more general form of genomic assortment, in which different allelic forms cause the same phenotype within the same gene or in different genes.