Plano de Trabalho para Bolsa de Pós-Doutorado Associada a Projeto Temático

Thematic Project:

i) **Title**: Natural selection in the MHC region of human populations

ii) **Supervisor**: Diogo Meyer

iii) **Institution**: University of São Paulo, Institute of Biosciences

iv) **Summary**

Admixture can contribute to adaptation by introducing novel advantageous variants into a population (Barreiro and Quintana-Murci 2020) . One of the regions of the genome with strong evidence for recent admixture-mediated selection is the MHC, which shows an excess of African ancestry with respect to the rest of the genome (Meyer et al. 2018; Zhou, Zhao, and Guan 2016), consistent with the hypothesis that African alleles are advantageous in the admixed population. This is an instance of *adaptive admixture*, the process where advantageous alleles are introduced into a population through admixture and can contribute to rapid adaptation (Norris et al. 2020). We will investigate patterns of ancestry deviation in a wide array of admixed populations, identify putatively selected alleles, and use simulation and theoretical approaches to understand the selective regimes that can drive the excess of African ancestry in the MHC.

v) **Goals**

First, we will ask if populations with distinct sources of African ancestry (i.e., with migrants from different regions from Africa) share the signal of recent admixture-mediated natural selection. Secondly, we will identify the specific African HLA alleles and haplotypes that have increased in frequency in the admixed population. Our study has the general goal of identifying whether recent selection plays an important role in shaping MHC variation. We will attempt to connect the theoretical models explaining long-term balancing selection at HLA genes with the cases where recent selection (as is the case of admixture mediated selection) occurs.

vi) **Strategy**

We will implement statistical tests designed to identify deviation from background ancestry within the MHC region (Grinde et al. 2019). Local ancestry will be estimated using RFMIX (Maples et al. 2013) or other methods that may be developed during the funding period. We will use simulations to evaluate the robustness of local ancestry estimation in the MHC. To do so, we will use forward simulations to model populations evolving under neutral conditions, but with source populations with high heterozygosity (thus mimicking potential biases associated to the MHC region). Our strategy relies on identifying HLA alleles in the admixed populations and carrying out tests for recent selection to identify putatively selected variants. To identify recently selected alleles we will use tests based on linkage disequilibrium flanking the HLA loci (e.g., integrated haplotype homozygosity, iHS) (Szpiech and Hernandez 2014), which will also provide an estimate of the timescale of selection on the HLA loci, allowing us to distinguish between the role of recent selection (post-admixture) from that which took place in Africa. Finally, using a simulation-based framework (Haller and Messer 2019) , we will explore a theoretical model of selection, according to which the high polymorphism of African populations explains the selective advantage of African alleles.

vii) Expected outcome.

Our study of adaptive admixture in the MHC will provide a detailed survey of how selection on very recent timescales (i.e.,after the onset of admixture, fewer than 20 generations) has shaped the diversity of key genes of the immune system. We will be able to identify how selection has favored alleles of specific ancestries, and provide mechanistic explanations for how selection favors specific alleles in admixed populations.

viii) Justification for postdoctoral scholarship request.

This project requires a researcher with the ability to navigate between the immunogenetics literature (central to the question about recent selection on HLA genes), skills in programming (necessary in order to implement the simulations, which will be used to model selective scenarios for the admixed populations), and thorough knowledge of population genetics theory. This diverse set of abilities will be most appropriately found in a PD researcher, which in accordance with FAPESP's expectations should be a promising researcher with a distinguished track record.

ix) Timeline

Semester	Activity
1	Curation of diverse datasets for admixed populations, local ancestry estimation; experiments to validate reliability of local ancestry inference in complex region such as MHC
2 and 4	Formal statistical testing of deviation of African ancestry within the MHC, with respect to genomewide averages
4	Identification of specific alleles experiencing directional selection, use of tests for positive selection and partial sweeps
5 and 6	Modelling of selective processes, test of whether heterozygote advantage on its own, without specifically favored alleles, can result in increase of African ancestry in MHC region.

x) References

Barreiro, Luis B., and Lluis Quintana-Murci. 2020. “Evolutionary and Population (epi)genetics of Immunity to Infection.” *Human Genetics* 139 (6): 723–32.

Grinde, Kelsey E., Lisa A. Brown, Alexander P. Reiner, Timothy A. Thornton, and Sharon R. Browning. 2019. “Genome-Wide Significance Thresholds for Admixture Mapping Studies.” *American Journal of Human Genetics* 104 (3): 454–65.

Haller, Benjamin C., and Philipp W. Messer. 2019. “SLiM 3: Forward Genetic Simulations Beyond the Wright-Fisher Model.” *Molecular Biology and Evolution* 36 (3): 632–37.

Maples, Brian K., Simon Gravel, Eimear E. Kenny, and Carlos D. Bustamante. 2013. “RFMix: A Discriminative Modeling Approach for Rapid and Robust Local-Ancestry Inference.” *American Journal of Human Genetics* 93 (2): 278–88.

Meyer, Diogo, Vitor R. C Aguiar, Bárbara D. Bitarello, Débora Y. C Brandt, and Kelly Nunes. 2018. “A Genomic Perspective on HLA Evolution.” *Immunogenetics* 70 (1): 5–27.

Norris, Emily T., Lavanya Rishishwar, Aroon T. Chande, Andrew B. Conley, Kaixiong Ye, Augusto Valderrama-Aguirre, and I. King Jordan. 2020. “Admixture-Enabled Selection for Rapid Adaptive Evolution in the Americas.” *Genome Biology* 21 (1): 29.

Szpiech, Zachary A., and Ryan D. Hernandez. 2014. “Selscan: An Efficient Multithreaded Program to Perform EHH-Based Scans for Positive Selection.” *Molecular Biology and Evolution* 31 (10): 2824–27.

Zhou, Quan, Liang Zhao, and Yongtao Guan. 2016. “Strong Selection at MHC in Mexicans since Admixture.” *PLOS Genetics*. https://doi.org/10.1371/journal.pgen.1005847.