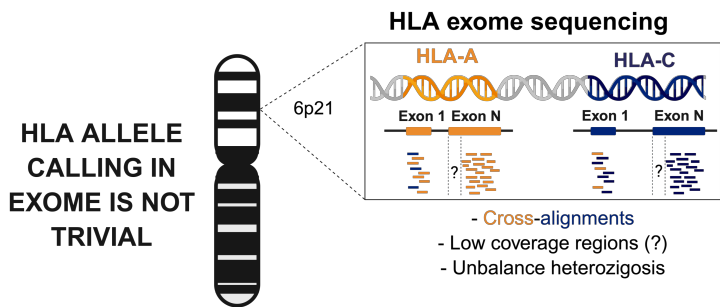


HLA-C typing departing from exomes combining HLA alignment and SNP detection, and imputation using multi-ethnic reference panels

Rafaela M. Barbosa¹, Erick E. Castelli², Leandro M. Colli¹; Diogo Meyer¹; Michel S. Naslavsky^{1,3}; Nicolas Vince⁴; Eduardo A. Donadi (eadonadi@fmrp.usp.br)¹.

¹University of São Paulo; ²University Estadual Paulista; ³Hospital Israelita Albert Einstein; ⁴Inserm, Nantes Université.

Introduction



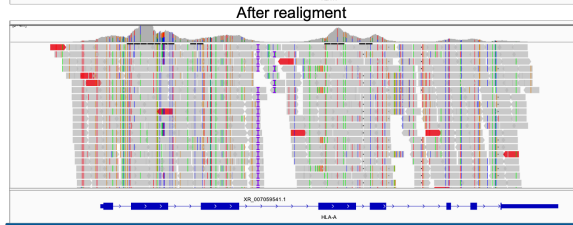
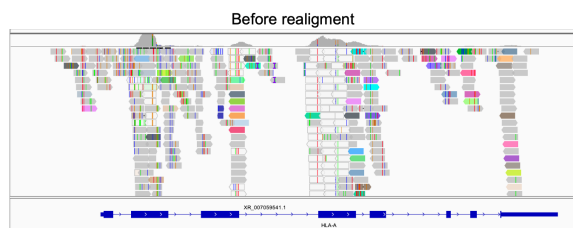
Aim

To develop and evaluate a bioinformatics pipeline for calling *HLA-C* alleles from exomes.

Methods

hla-mapper v.5 (in development)

- 233 exomes
- Read realignment
- Variant calling (MHC region)
- Filtering out low confidence variants
- Calling *HLA-C* alleles



Imputation model

- Variants across all HLA genes (*hla-mapper* file)
- Multiallelic decomposition (*bcftools*)
- Multiethnic reference panel ($n = 3674$)¹
- Filtering common variants between samples and the reference panel
- Create imputation models with *HIBAG* (HLA genotype imputation with attribute bagging)
- w 5,000,000 - b 100

¹Silva, et al., HLA, <https://doi.org/10.1111/tan.15543>

Results

HLA-C Imputation Outcomes (Proportions)

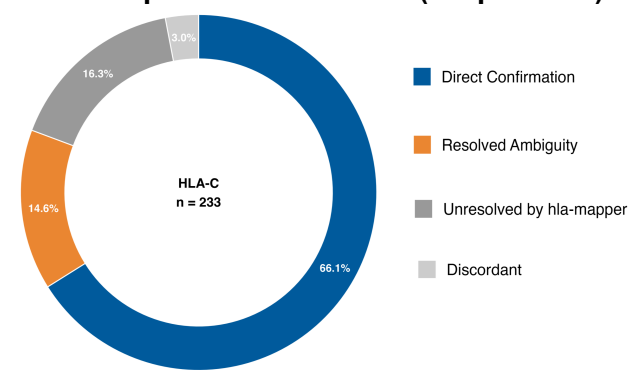


Fig 1: Comparison of *HLA-C* allele calls from *hla-mapper* and imputation ($n=233$). Discordant = different allele calls; Unresolved by *hla-mapper* = no confident call, imputation provided one; Resolved Ambiguity = multiple options resolved by imputation; Direct Confirmation = identical allele calls by both methods.

Posterior Probability of *HLA-C* Imputation

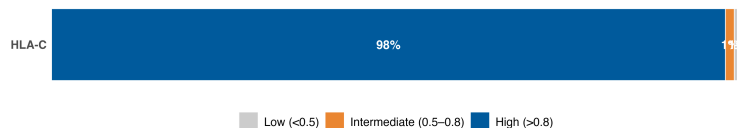


Fig 2: Posterior probability metrics for *HLA-C* alleles imputation.

HLA-C Allele Frequency Correlation

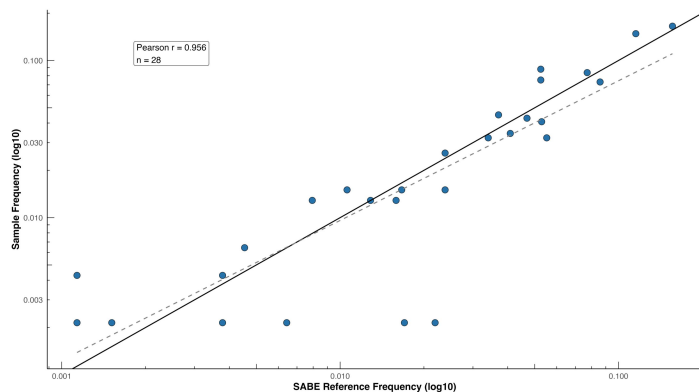


Fig 3: Correlation of *HLA-C* allele frequencies between 233 imputed Brazilian exomes and SABE reference. Points = alleles (blue). Solid line = identity; dashed = regression. r = Pearson; n = alleles. Log scale.

Conclusion

Using HLA-specific alignment, SNP detection and imputation with multi-ethnic panels proved to be a useful approach for predicting *HLA-C* alleles departing from exome data.

ACKNOWLEDGMENTS