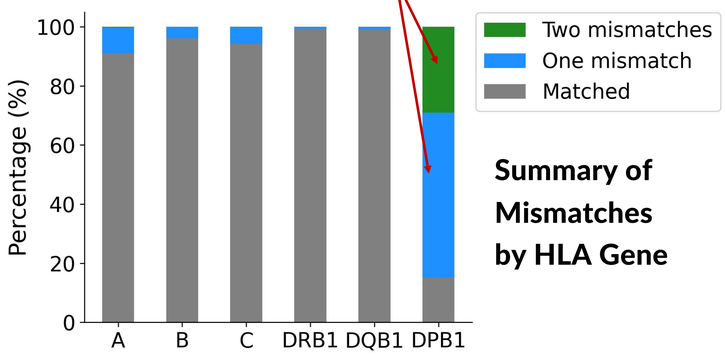
Machine Learning Insights into HLA Noncoding Sequence Mismatches and Their Impact on DPB1 Matching in Hematopoietic Cell Transplantation

Medhat Askar, Timothy L. Mosbruger, Grace Tzun-Wen Shaw, Haedong Kim, Yuncheng Duan, Andrew S. Allen, Jamie L. Duke, Timothy S. Olson, Dimitri S. Monos, Tristan J. Hayeck

Questions?
hayeckt@upenn.edu

Introduction

- Hematopoietic cell transplantation (HCT) aka bone marrow transplantation is the only curative treatment for leukemia, immunodeficiencies, and other disorders.
- However, HCT significant increases risk acute graft-versus-host disease and treatment-related mortality
- Often non-coding mismatches not collected
- Finding perfect HLA patient donor matches is often difficult, especially for HLA-DPB1



Summary of Mismatches by HLA Gene

Cohort and defining mismatches

- Center for International Blood and Marrow Transplant Research (CIBMTR) 5106 patient–donor pairs from with ultra high resolution typing of HLA class I genes split into:
  - 2030 Deplete (given ATG or Alemtuzumab)
  - 3076 Replete patient–donor pairs

Methods

- Testing defining matches with and without non-coding variants (ie 2 Field vs 4 Field)
- Cox proportional hazards with inverse probability treatment weighting to control for confounding

Results

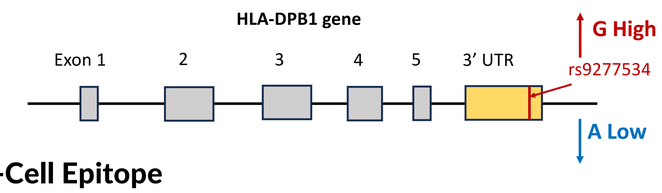
- When running the same analysis but only allowing for coding mismatches both upward bias and inflation of significance is observed
- For deplete cohort significant hazards of 2 DPB1 (DP2vsDP5 group) & ≥ 1 noncoding class I mismatch versus Baseline Matched HLA except + 2 DPB1 Mismatches
  - **Survival hazard ratio 1.67 p-value = 1.3×10^{-5}**
  - **TRM hazard ratio 1.94 p-value = 8.9×10^{-7}**

Discussion

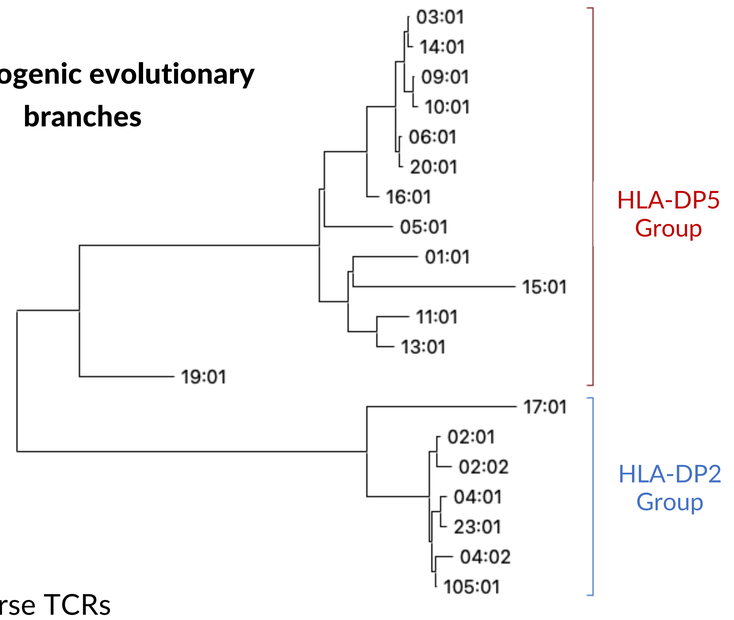
- **Non-coding mismatches** in HLA class I genes along with DP2 vs DP5 evolutionary branch mismatches show significant hazards of treatment related mortality and overall survival
- Previously overlooked non-coding mismatches may have both **upwardly biased hazards and inflated the significance** of the impact of DPB1 mismatches alone on TRM and survival

3 Ways of Defining DPB1 Mismatches

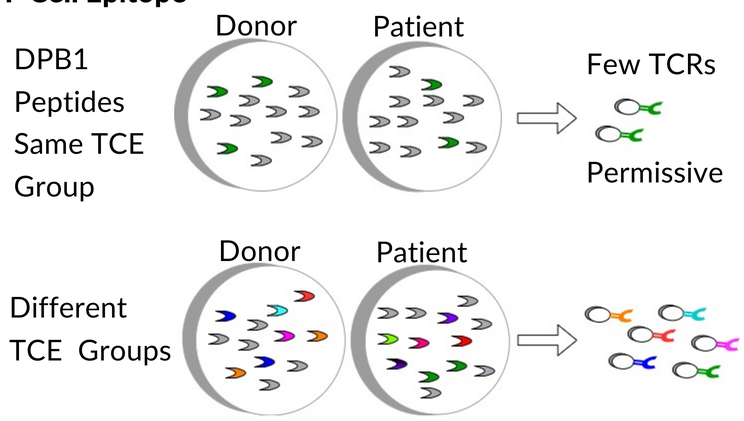
A) Expression Profile Marker



B) Phylogenic evolutionary branches

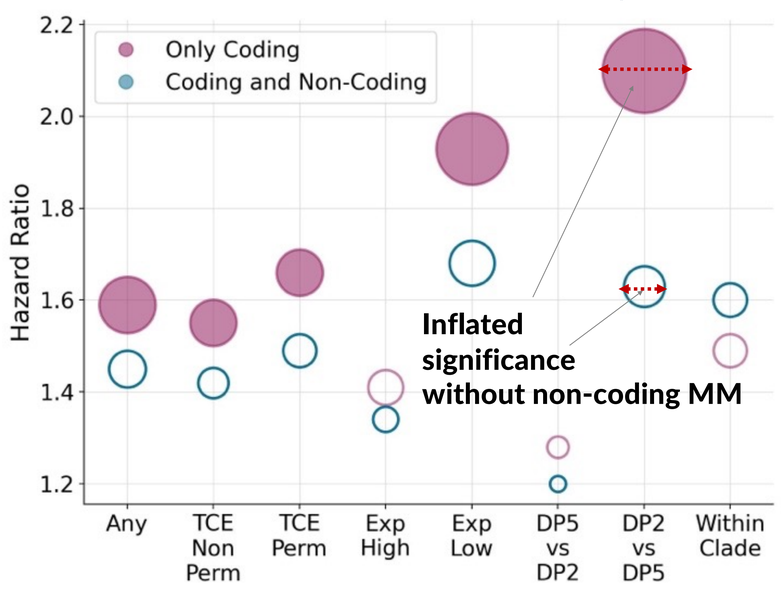


C) T-Cell Epitope

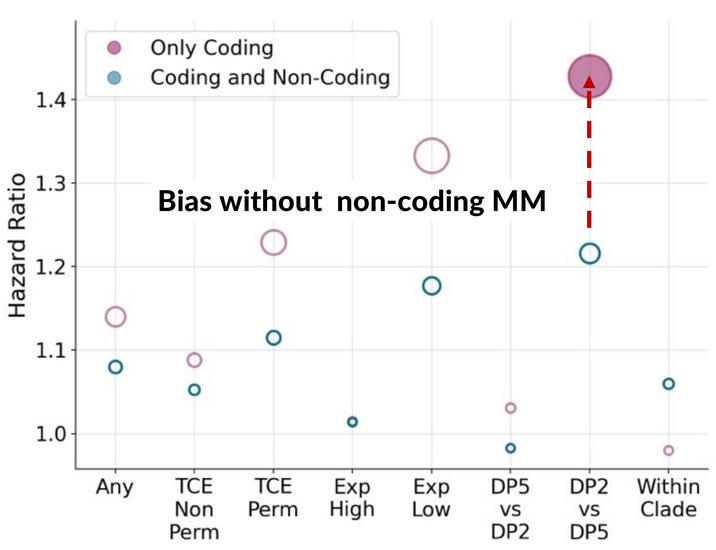


Comparison of different ways of classifying permissiveness with and without HLA class I non-coding mismatches included. Filled in dots indicate significance at an alpha level of 1% and the sized of the dots corresponds to the $-\log_{10}(p)$ (bigger means more significant).Consistent trends show across mismatch methods, hazard ratios are biased upwards and significance is inflated if HLA class I non-coding are missing or ignored.

Treatment Related Mortality

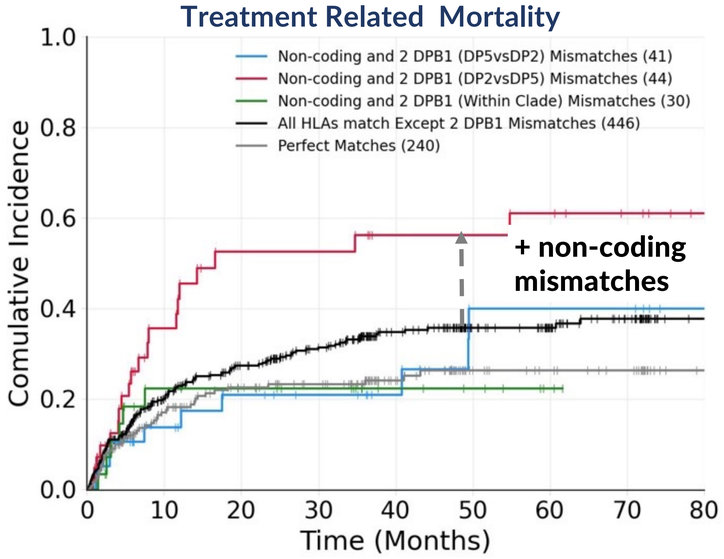


Overall Survival



Hazards of treatment related mortality and overall survival in the context of two DPB1 evolutionary branch (Patient DP2 vs Donor DP5) mismatches and non-coding mismatches Class I HLA. Addition of non-coding mismatches worsens treatment related mortality and overall survival.

Treatment Related Mortality



Overall Survival