The impact of HLA genotype imputation on allele determination, eplet mismatch and alloimmune risk categories in a Southeast Asian cohort of kidney transplant recipients

ETY Wong[1,2], A Ho[2], HK Sran[1,2], MR D'Costa[1,2], ZY Chang[1,2], APY Ng[1], A Lim[1], W Koh[1], A Vathsala[1,2]

[1]National University Centre for Organ Transplantation, National University Hospital, Singapore
[2]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

Background:

HLA-DR/DQ molecular mismatch (mMM) categories had been published and validated as a prognostic biomarker of primary alloimmunity in several cohorts of kidney transplant recipients. However, eplet analysis requires high-resolution genotype data, which may often be unavailable, especially in older cohorts or under-resourced centers. . The need for high-resolution genotyping is a barrier for centers in low-resource settings or retrospective analyses of historical cohorts where low-resolution typing was the standard. We examined the accuracy of imputed 2-field alleles, eplet mismatches, and mMM categories in a cohort whose donor and recipient HLA genotypes were obtained by Next-Generation Sequencing (NGS).

Methods:

We analyzed 51 kidney transplant recipient and donor pairs. NGS HLA genotyping data were transformed to low resolution by removing the second field and considering serological splits. The low-resolution genotyping was imputed using HaploStats to generate high-resolution genotyping for HLA-A, -B, -C, DRB1, -DRB345, and -DQB1, guided by each individual's self-identified race. HLA-DQA1 was imputed using published haplotype frequency reference standards describing HLA-DRB1-DQB1-DQA1. HLAMatchmaker (ABC v4.0 and DRDQDP v2.2) was used to determine the single-molecule eplet mismatch. Recipients were categorized into mMM categories using previously published thresholds.

Table 1. Concordance of imputed and NGS alleles in each race

HLA loci	Chinese (n=52)	Malay (n=19)	Indian (n=17)	Others† (n=9)
HLA-A	85%	71%	82%	72%
HLA-B*	93%	89%	82%	72%
HLA-C	95%	87%	88%	89%
HLA-DRB1	84%	87%	82%	72%
HLA-DRB345	60%	53%	71%	50%
HLA-DQB1	88%	87%	79%	83%
HLA-DQA1	88%	92%	82%	94%

*p=0.05, all other p=NS
†Includes Caucasian, Eurasian, Thai, and Vietnamese races

NGS: Next Generation Sequencing

Table 2A. Correlation of molecular mismatch analysis between imputed and NGS genotyping

HLA molecule	Pairwise difference in single molecule mismatch between imputed and NGS genotyping		
	Mean ± SD	IQR (Max)	R^2*
HLA-A	0.2 ± 0.6	0, 0 (3)	0.97
HLA-B	0.0 ± 0.2	0, 0 (2)	1.00
HLA-C	0.1 ± 0.3	0, 0 (2)	0.99
HLA-DRβ_1	0.2 ± 0.6	0, 0 (4)	0.97
HLA-DR$\beta_{3/4/5}$	1.1 ± 2.8	0, 1 (16)	0.66
HLA-DQ$\alpha_1\beta_1$	0.8 ± 1.7	0, 1 (8)	0.95

*All p<0.0001

Table 2B. Changes in recipients' HLA-DR/DQ molecular mismatch categories between imputed and NGS genotyping

HLA-DR/DQ molecular mismatch risk category†		Number of recipients
NGS	Imputed	
Low	Low	12/51 (24%)
Low	Intermediate	Nil
Low	High	Nil
Intermediate	Low	3/51 (6%)
Intermediate	Intermediate	20/51 (39%)
Intermediate	High	2/51 (4%)
High	Low	Nil
High	Intermediate	Nil
High	High	14/51 (27%)

†Low Risk: Maximum HLA-DR$\beta_{1/3/4/5}$ <7 and Maximum HLA-DQ$\alpha_1\beta_1$ <9
Intermediate Risk: Any HLA-DR$\beta_{1/3/4/5}$ and Maximum HLA-DQ$\alpha_1\beta_1$ 9-14
High Risk: Any HLA-DR$\beta_{1/3/4/5}$ and Maximum HLA-DQ$\alpha_1\beta_1$ ≥15

IQR: Interquartile range; NGS: Next-Generation Sequencing; SD: Standard deviation

Results:

This cohort comprised 32 living- and 19 deceased-donor pairs, with 97 unique HLA samples. Most were Chinese (53%) and Malay (22%). The concordance between imputed and NGS alleles at each HLA locus was 80% (A), 89% (B), 92% (C), 83% (DRB1), 59% (DRB345), 86% (DQB1), and 88% (DQA1). The concordance at HLA-DRB345 improved to 80% when imputed DRβ_4*01:01 alleles were analyzed as DRβ_4*01:03. The concordance at each locus for each race is shown in Table 1.

Imputed and NGS single molecule eplet mismatches were highly correlated at most loci ($R_2 \geq 0.95$) except HLA-DRβ_{345} (R_2=0.66), due to null alleles that could not be imputed (Table 2A).

Using imputed genotyping, 46/51 (90%) recipients remained in the same mMM categories as NGS genotyping (Table 2B). All low-risk and high-risk pairs were identically classified by both NGS and imputation. There were no shifts from low to high risk or vice versa.

Conclusions:

Although imputation is imprecise in allele determination at high-resolution, it preserves mMM risk assessment in a Southeast Asian cohort in most cases. While this supports imputation in mMM risk categorization where NGS is unavailable, HLA databases should be enriched with Southeast Asian races to improve precision in immunological assessment.