HLA Amino Acid Polymorphisms and Kidney Allograft Survival

Supplemental Digital Content

Malek Kamoun, MD, PhD[a], Keith P. McCullough, MS[b], Martin Maiers, MS[c], Marcelo A. Fernandez Vina, MD, PhD[d], Hongzhe Li, PhD[e], Valerie Teal, MS[e], Alan B. Leichtman, MD[b], Robert M. Merion, MD[b]

[a]Department of Pathology and Laboratory Medicine and [e]Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA; [b]Arbor Research Collaborative for Health, Ann Arbor, MI; [c]National Marrow Donor Program, Minneapolis, MN; and [d]Department of Pathology, Stanford University, Palo Alto, CA

Corresponding Author: Malek Kamoun (malekkam@mail.med.upenn.edu)

Table of Contents

Supplemental Digital Content: Materials and Methods

Imputation of HLA alleles and HLA amino acid mismatches

We used a multistep process to impute HLA-A, -B and -DRB1 alleles from the SRTR dataset covering the period of 1987-2009. In the first step, HLA-A, -B, and -DRB1 typing (first field, classically two-digit specificity) of recipient and donors were obtained from the SRTR files which contains data from the Recipient Histocompatibility Forms and Donor Histocompatibility Forms. Because these Forms list a few HLA broad specificities instead of split antigens, in these cases, HLA two-digit specificities were imputed using an algorithm developed by the National Marrow Donor Program [1]. HLA specificities of recipient and donors (first field, classically two-digit) were then mapped to a list of HLA alleles corresponding to the two-digit HLA specificities using an algorithm developed by the National Marrow Donor Program [1]. In the next step, we used a previously described computational method for estimating HLA alleles [2]; through the application of a set of statistical inference rules, this method utilizes haplotype frequency information for each self-reported racial/ethnic group (Whites, Asian/Pacific Islanders, Blacks, and Hispanics) [3]. Based on genotype likelihoods computed from 3-locus haplotype frequencies, a statistical assignment was made at the allele level for each HLA-A, -B and –DRB1 specificity (first field, classically two-digit). The algorithm used frequency thresholds at the haplotype and genotype level [2]. The threshold value for HLA haplotypes and genotypes was set at 99% and 98%, respectively. HLA-A, -B, and -DRB1 alleles were imputed on over 240,024 pairs; less than 3% of cases were rejected due to missing or incomplete HLA typing data (HLA-A, B or DRB1) for either the recipient or the donor. The output of the imputation provided a listing of all possible

HLA alleles for each HLA-A, -B and -DRB1 specificity present in each donor-recipient pair, and their relative likelihoods. Imputation based on information from 3-locus haplotype improved prediction probability, compared to imputation based on single-locus allele frequencies; imputation from 3-locus haplotypes also provided a relatively accurate estimation of most HLA-A, -B and, -DRB1 alleles inferred from low resolution (first field, classically two-digit) HLA assignments [2].

The assignment of HLA amino acid MMs was achieved as follow. The statistically imputed HLA-A, -B and -DRB1 alleles were used to determine the corresponding assignment of amino acid polymorphic sites of these HLA alleles by converting each allele to its constituent amino acid variants using the allele alignment data from the IMGT/HLA database (https://www.ebi.ac.uk/ipd/imgt/hla/; Release 2.28.0, January 2010). This analysis was restricted to polymorphic residues localized in the alpha1 and alpha2 domains of HLA-A and –B and beta1 domain of HLA-DRB1 molecules where peptide binding and T cell receptor-based immune recognition occurs [4,5]. The relationship between the polymorphic sites and the function of these domains in the context of allorecognition were previously described [5]. We applied this method in a pairwise fashion to the full cohort of kidney transplant donor-recipient pairs to obtain a comprehensive listing of a) all polymorphic amino acid positions in the alpha1 and alpha2 domains of HLA-A and –B and beta1 domain of HLA-DRB1 molecules for each transplant pair, b) the likelihood of amino acid MM and c) the exact residues involved (e.g., valine vs. glycine). The likelihood of a MM at one position for each donor-recipient pair was computed as the sum of the likelihoods of a MM at each polymorphic amino

acid position for each HLA-A, -B or -DRB1 locus counting mismatches in the direction of donor into recipient only. Thus, donor-recipient MMs were characterized at each residue across the length of the HLA amino acid sequences with a relative likelihood of 0 MM, 1 MM or 2 MMs. The final imputed datasets contained the amino acid MM variables combined with patient and donor demographic and clinical variables known to influence recipient outcomes.

The uncertainty in the estimation of HLA amino acid MM sites was relatively low for most HLA-A, -B, and -DRB1 sites as illustrated by the percentage of cases with a probability of a zero MM in the 0.05-0.95 range (Fig. S2). For HLA-A, out of 49 amino acid MM sites, only MMs at position 156 had a probability of a zero MM that fell in this range in a significant percentage of cases (17%). For HLA-B, out of 55 amino acid MM sites, only MMs at positions 114, 116, and 156 had a probability of a zero MM that fell in the 0.05-0.95 range in more than 15% of the cases (18%, 17%, and 25% respectively). For HLA-DRB1, out of the 30 MM sites evaluated, the following positions had a probability of a zero MM in the 0.05-0.95 range in more than 20% of the cases: position 47 (19%), 57 (27%), 67 (32%), 70 (25%), 71 (47%), 74 (24%), and 86 (70%) [Fig. S2]. For all DRB1 MM sites the percentage of cases with a probability of a zero MM in the 0.25-0.75 range were less than 10% except for MM at positions 71 and 86 (20% and 38% respectively) [data not shown]. We have also evaluated the error rate per amino acid MM site due to the imputation by using datasets consisting of HLA genotypes with phased haplotype pairs and no allelic ambiguity, then artificially introduced ambiguities in this datasets by converting the four-digit specificity to serologic specificity (two-digit

specificity). Datasets of over 200,000 individuals obtained from the National Marrow Donor Program registry, each were constructed for the four primary race groups in the registry: White (CAU), Asian (API), Black (AFA) and Hispanic (HIS). Each individual genotype was created by randomly sampling two haplotypes from previously published US high-resolution A~B~DRB1 haplotype frequencies [3]. The original unambiguous genotypes were used as truth for the evaluation of the error rates of amino acid MMs; the error rates were computed by comparing the amino acid MM of the simulated subject to the imputed amino acid MM after conversion to serologic specificity (two-digit specificity). The error rate per amino acid MM site for HLA-A, B, and DRB1 loci for each race was equal or less than 10% for all amino acid MM sites except for MMs at position 57, 71, and 86 of DRB1 (Fig. S3).

Racial Distribution

The racial distribution in this population was as follows: 144,183 Whites; 10,367 Asian/Pacific Islanders; 56,327 Blacks; 26,913 Hispanics; and 2,234 in other categories.

Exclusion Criteria

Patients were excluded from all analyses if they received a multi-organ transplant, or if they were missing information about the HLA-A, -B, or -DRB1 loci, age race or ethnicity. Less than 1% of patients were excluded for these reasons. White, Asian/Pacific Islander, Black, and Hispanic recipients were evaluated separately since the degree and characteristics of amino acid polymorphism varies by race and ethnicity. Since many donor-recipient race combinations had insufficient data for separate analyses, we

analyzed data for all transplants, adjusting for donor and recipient race, and also conducted separate analyses on specific donor-recipient race combinations, namely those where both the donor and recipient had the same race, or where the donor-recipient pair was either Blacks and Whites or Whites-Blacks.

HLA amino acid covariates, including HLA-A, -B, and -DRB1 amino acid variable sites which had expected mismatch levels less than 1% of the time for every racial group, were eliminated from analysis. This approach decreased the number of HLA-A, -B and -DRB1 variable sites from over 400 to 125.

Statistical Analysis

Continuous variables were modeled by a series of linear splines in order to reduce the need for assumptions about functional form for donor age (spline knot at 18 years), recipient age (spline knots at 18, 50, and 65 years), and BMI (spline knot at 20). These splines were chosen after statistical review of the significance and clinical review of their plausibility. Interactions were included in the survival models if they had clinical face validity, were statistically significant; these interaction terms were donor ECD with recipient age, recipient age with recipient diabetes (including a spline knot in the interaction at age 50 years), recipient BMI with recipient diabetes, and recipient prior transplant with recipient zero interval years on dialysis (i.e. preemptive subsequent kidney transplant). We used log-likelihood tests, which measure variables' contributions to the improvement in the fit of a model, to determine the additional predictive power that each new variable brought to each iterative version of the model. Since we were testing a large number of variables and interactions on an extremely large dataset, we

used Bonferroni-style correction on the many HLA amino acid mismatch indicators to avoid the multiple-comparisons problem (Table S3). We did not just rely on statistical significance, including only variables that made clinically important differences in the predictive value of the model on outcomes [6,7]

Adjusted model results accounted for the following confounding patient and donor variables.

- HLA MM, using classical two-digit specificity (Indicators of 0, 1, and 2 MMs at HLA-A, -B, and -DRB1, along with an overall 0 ABDRB1 MM term)
- Whether the kidney was shared between donation service areas (shipped out of the local area in which it was recovered)
- Donor after cardiac death
- Donor age (linear, with spline knot at 18)
- Donor cause of death (anoxia, cerebrovascular accident (CVA), central nervous system tumor, head trauma, other)[‖]
- Donor cytomegalovirus (CMV) serological status (negative or positive on screening for existing anti-CMV antibody)
- History of donor hypertensive
- Donor weight (log, with spline knot at ln(55 kg))
- Meets definition of an Expanded Criteria Donor (ECD) (age 60 years or greater, or age 50 to 59 years with two of the following criteria; history of hypertension, death by CVA, or terminal creatinine (creatinine at allocation) greater than or equal to 1.5 mb/dL[‖].

- ECD by age interaction‖
- Year of transplant (to account for trends over time in outcomes)
- Recipient age
- Recipient diabetes status
- Recipient age by diabetes interaction
- Cause of recipient End stage Renal Disease (hypertension, polycystic kidney disease, glomerulonephritis, other)
- History of recipient previous kidney transplant
- Recipient history of previous transplant by diabetes interaction)
- Recipient Body Mass Index (BMI)
- Recipient BMI by diabetes interaction
- Recipient peak Panel Reactive Antibody (PRA) at transplant (categories <10, 10-79, 80+); these PRA represent the maxima among the reported PRAs for each patient
- Recipient years on dialysis at time of transplant (log, with indicator for zero)
- Interaction between no dialysis prior to transplant and prior transplant)
- Relationship (related, unrelated, missing)**

‖applicable only to deceased donor (DD) transplants

**applicable only to living donor (LD) transplants

Supplemental Digital Content: Results

Mean frequency of expected HLA-A, -B, and -DRB1 amino acid mismatched sites by race

The mean frequency of expected HLA-A, -B, and -DRB1 amino acid mismatched sites by race is shown in Table S1. Overall, HLA-A, -B, and –DRB1 amino acid MM sites included polymorphic solvent exposed residues as well as unexposed peptide binding sites, including peptide-binding pockets of HLA class I and class II molecules at the antigen recognition site. In HLA-DRB1, overall, the most frequent mismatched sites included (listed in numerical order): positions 11 [pocket 6, (65%)], 13 [pocket 6, (74%), 37 [peptide-contact site], and 67 [pocket 7, (54%)], 70 [pocket 4, (43%)], 71 [pocket 4, (61%)], and 74 [pocket 4, (46%)].

Positional specificity and number of HLA-DRB1 amino acid mismatches

Comparison of the amino acid sequence alignment between HLA-DRB1 mismatched specificities (two-digit specificity) using the allele alignment data from the IMGT/HLA database (https://www.ebi.ac.uk/ipd/imgt/hla/) shows, a wide range of variation in the number and positions of amino acid MMs depending on the mismatched HLA-DRB1 antigen (two-digit specificity). As an example, a comparison of amino acid MMs between a DR13 recipient versus a DR14 or a DR9 donor is shown in Table S2; there are only 7-9 amino acid MM sites between HLA-DR13 and DR14; in contrast, there are 17-19 amino acid MM sites between HLA-DR13 and DR9. In these cases, the number of amino acid MMs that can be attributed to the variations in the DRB1*13 alleles (four-digit septicity) is limited to 1-3 sites. These findings are consistent with a previous study [8], as well as our

previous prospective cohort study of 697 renal transplant recipients of deceased donors in which we have performed allele level HLA-A, -B, and -DRB1 (four-digit-specificity) typing to characterize HLA matching [9,10]. In addition, the examples above illustrate that many of the amino acid substitutions associated with HLA mismatches at the antigen level (two-digit specificity) are peptide-binding sites including pockets of the binding groove (Table S2). Thus, the number and position of amino acid MMs vary significantly depending on the mismatched HLA specificity (two-digit specificity). Yet, the current standard approach based on HLA antigen matching (1 or 2 DRB1 antigen MM) does not take into account these biologically significant structural differences in the HLA molecules among various HLA specificities.

Analyses of Cox model by expected number of amino acid mismatch adjusted for donor type, race and ethnicity

In order to determine whether amino acid MM might have different effects in different races, we constructed separate models for deceased and living (related and unrelated) donors, and race combinations. These were fully adjusted for the same donor and recipient factors as used throughout these analyses. Each model is a survival model of graft failure using the continuous sum of amino acid MMs and splines at 5, 10, and 15 MMs at HLA-A, -B and –DRB1. Results showed the likelihood ratio test for interaction between the amino acid mismatch variables and race showed an overall p-value of 0.10 among various race and ethnic groups (data not shown); however in the Blacks, Hispanic, and Asian American groups, fewer results had P-values < 0.05 perhaps reflecting a higher rate of uncertainty associated with the imputation of HLA alleles and

the estimated amino acid MMs due to limited data on haplotype frequencies in these racial groups as compared to that of White populations [2]. We have also tested whether the sum of expected amino acid MMs have different effects by race using a log-likelihood test to evaluate interaction terms between race and the adjusted graft survival using the amino acid mismatch variables (data not shown); the P-values for these interactions were not significant (P=1.0 for deceased donors and P=0.097 for living donors).

Adjusted hazard ratio of graft failure by sum of amino acid mismatch in HLA-A, -B, and -DRB1 loci: Model based on 10 imputations

In this second analysis, a Cox model using 10 separate analyses of imputed amino acid variables was used to evaluate the adjusted HR of graft failure by sum of amino acid MMs in HLA-A, -B, and -DRB1 loci based on 10 imputations using a continuous sum of amino acid mismatches and splines at 5, and 10 MMs, at HLA-A, B and DRB1. Each imputed dataset was analyzed separately in a Cox model and the data generated from the 10 imputation set was combined as previously described [10]. The effect of the imputation on the beta in terms of the variability of that estimate (i.e. the between-iteration SD) was between 0.001 and 0.005 for all these runs, which is at least one log lower compared to the variance within imputations. As in Figure 1 of the main manuscript, a statistically significant linear relationship was observed between the number of HLA-B and HLA-DRB1 amino acid MMs and graft failure for deceased donors (P=0.04 for HLA-B and -DRB1); for living related donors the initial DRB1 slope

was borderline significant (P=0.07). No other slopes or splines were significant (Fig. S1).

Additionally, a Cox model adjusted for each donor type showed that the interaction between HLA amino acid MMs versus deceased donor and living donor transplants was highly significant in all 10 imputations. The effect of the number of amino acid MM on graft failure in DRB1 was significant and similar across the donor types; in HLA-A, this effect was significant in deceased donor transplants but was only marginally significant in living related donor transplants and not significant in living unrelated donor transplants; HLA-B amino acid MMs appear to have a marginally significant association with increased risk among deceased donor recipients, and was not significant among living unrelated and living related donor transplant recipients (Table S4).

Joint effect of HLA amino acid mismatches at individual sites on graft failure

Stepwise Cox proportional-Hazards regression models using individual amino acid MM sites were run separately for deceased donor and living donor transplants. A forward selection was applied starting with the demographics, co-morbidity factors, and HLA antigen MM variables described in the Methods section. Two different analyses were used in the stepwise model. In the first analysis, the amino acid MM variables that were found to be statistically significant by themselves, listed in Table S3 were allowed to enter the model one at a time. Increased risk of graft failure was found to be associated with amino acid MMs at several functionally important positions of the antigen recognition sites of HLA-A, -B, and -DRB1 molecules. Notably, in HLA-DRB1, MMs at

positions 10, 13 (peptide-binding site, pocket 6), 26 (peptide-binding site), 28 (peptide-binding, pocket 4), 71 (peptide-binding, pocket 4), and 74 (peptide-binding site, pocket 4) were all found to be associated with increased risk of graft failure (Table S7). As a consequence of the well-known complex linkage disequilibrium patterns of polymorphic amino acids in the peptide binding site of HLA class I and class II molecules [11], several HLA-DRB1 amino acid MMs that are associated with graft failure were found to be highly correlated with other MMs that were not selected in this step-wise model. This finding is not unexpected since the stepwise process is very sensitive to small variations; for example, if there are two correlated positions, the inclusion of one over the other can be influenced by which variables is retained in the model during prior steps in the selection process. For example, a MM at position 10 in DRB1 molecule, is highly correlated with a few MMs located in pocket 7 at the peptide peptide-binding sites, these MMs include position 9 (pocket 6 and 9) and position 11 (pocket 6). Similarly, a mismatch MM at position 13 (pocket 6) is highly correlated with a MM at position 11 (pocket 6), and position 37 (peptide-binding site) [Table 2 of the main manuscript]. In addition, increased risk of graft failure was found to be associated with amino acid MMs at several functionally important positions of the antigen recognition sites of HLA-A, and -B, molecules. In HLA-A molecule, these sites include amino acid substitutions at positions 116 (peptide-binding site, pocket E and F), 62 (TCR contact site), 150 (TCR contact site), and 158 (TCR contact site); in the HLA-B molecule, substitutions at position 63 (peptide-binding site, pocket A), and 143 (pocket F). The apparent disparity in the assortment of MM sites associated with graft failure observed between deceased donors and living donor transplant could also be influenced by the

imputation itself or by which variables were retained in the model during prior steps in the selection.

These findings were confirmed using 10 separate analyses of the imputed HLA alleles with their corresponding amino acid MM variables; stepwise Cox proportional-hazards regression models using individual amino acid MM sites were fitted separately for deceased donor and living donor transplants (Table S6). In HLA-DRB1, amino acid MM at 5 positions, including 4 peptide-binding sites, were associated with a significantly higher risk of graft failure in all 10 imputations, with $P<0.05$; these sites include amino acid substitutions at positions 10; 13 (in pocket 6); 26; 28 (pocket 7); and 74 (pocket 4). Two additional DRB1 MMs sites positions 32 and 71 (pocket 4) were also found to be associated with graft failure with $P<0.05$; these sites were selected in 60% and 50% of the imputations respectively (Table S6). HLA MM sites that were selected in all 10 imputations provide evidence that its relationship with graft failure is robust in the imputation process. Similarly, in HLA-A, amino acid MM located at position 116 (pocket E and F) was associated with a significantly higher risk of graft failure in all 10 imputations, confirming the finding described in the above stepwise logistic regression model. A few other functionally important HLA-A amino acid MMs were also found to be associated with graft failure with $P<0.05$ but were selected only in 40-70% of the imputations. These sites include two TCR contact sites (position 150, and 158); position 77 (pocket F); and position 144 (highly correlated with MM at 151, a TCR contact site) (Table S6). In HLA-B, amino acid MM at position 158 (TCR contact site) was associated with a significantly higher risk of graft failure in all 10 imputations. Five other

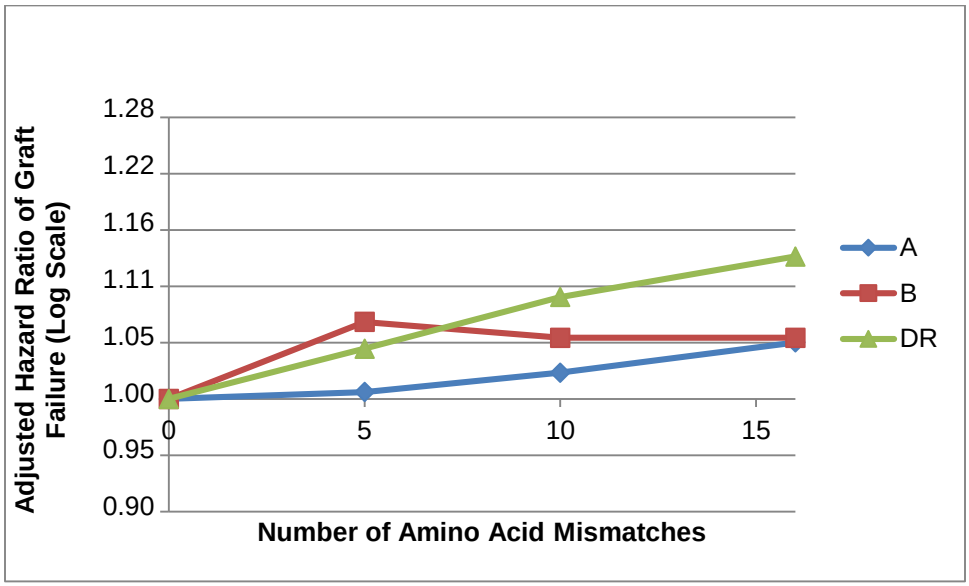
additional HLA-B amino acid MMs which were selected in 70-90% of the imputations were also found to be associated with graft failure with $P<0.05$. These sites include positions 63 (pocket A), 67 (pocket B), and 147 (pocket E/F) [Table S6].

References

1. Holdsworth R, Hurley C, Marsh S, et al. The HLA dictionary 2008: a summary of HLA-A, -B, -C, -DRB1/3/4/5, and -DQB1 alleles and their association with serologically defined HLA-A, -B, -C, -DR, and -DQ antigens. *Tissue antigens.* 2009;73(2):95-170.
2. Madbouly A, Gragert L, Freeman J, et al. Validation of statistical imputation of allele-level multilocus phased genotypes from ambiguous HLA assignments. *Tissue antigens.* 2014;84(3):285-292.
3. Maiers M, Gragert L, Klitz W. High-resolution HLA alleles and haplotypes in the United States population. *Human immunology.* 2007;68(9):779-788.
4. Bjorkman P, Saper M, Samraoui B, Bennett W, Strominger J, Wiley D. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature.* 1987;329(6139):512-518.
5. Reche P, Reinherz E. Sequence variability analysis of human class I and class II MHC molecules: functional and structural correlates of amino acid polymorphisms. *Journal of molecular biology.* 2003;331(3):623-641.
6. Dickinson D, Arrington C, Fant G, et al. SRTR program-specific reports on outcomes: A guide for the new reader. *American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons.* 2008;8(4p2):1012-1026.
7. Dickinson D, Dykstra D, Levine G, Li S, Welch J, Webb R. Transplant data: sources, collection and research considerations, 2004. *American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons.* 2005;5(4p2):850-861.
8. Kamoun M, Israni A, Joffe M, et al. Assessment of differences in HLA-A, -B, and -DRB1 allele mismatches among African-American and non-African-American recipients of deceased kidney transplants. *Transplantation proceedings.* 2007;39(1):55-63.
9. Baxter-Lowe L, Maiers M, Spellman S, et al. HLA-A disparities illustrate challenges for ranking the impact of HLA mismatches on bone marrow transplant outcomes in the United States. *Biol Blood Marrow Transplant.* 2009;15(8):971-981.
10. Kamoun M, Holmes J, Israni A, et al. HLA-A amino acid polymorphism and delayed kidney allograft function. *Proceedings of the National Academy of Sciences of the United States of America.* 2008;105(48):18883-18888.
11. Thomson G, Marthandan N, Hollenbach J, et al. Sequence feature variant type (SFVT) analysis of the HLA genetic association in juvenile idiopathic arthritis. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing.* 2010:359-370.

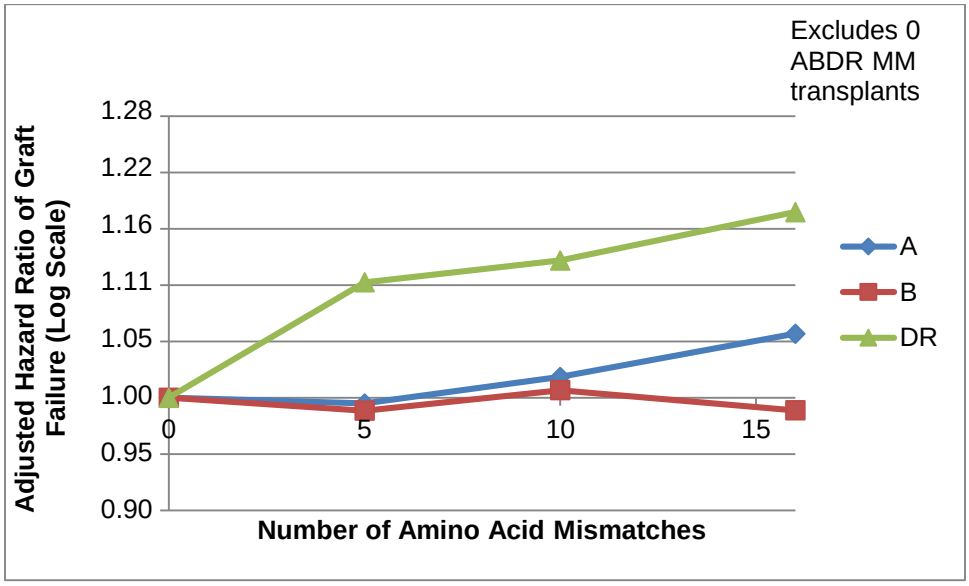
Figure S1. Adjusted hazard ratio of graft failure: Model based on 10 imputations of amino acid mismatches in HLA-A, -B, and -DRB1 loci

Figure S1A. Adjusted hazard ratio of graft failure in deceased donor transplants



Summarized data over 10 separate analyses of imputed amino acid mismatches in HLA-A, -B, and -DRB1 loci. Initial B, DRB1 slopes have p-values of 0.04. Spline of B slope at 5 has p-value of 0.04. No other slopes or splines have $p < 0.20$. 0 ABDRB1 MM transplants were not excluded from these analyzes.

Figure S1B. Adjusted hazard ratio of graft failure in living related donor transplants

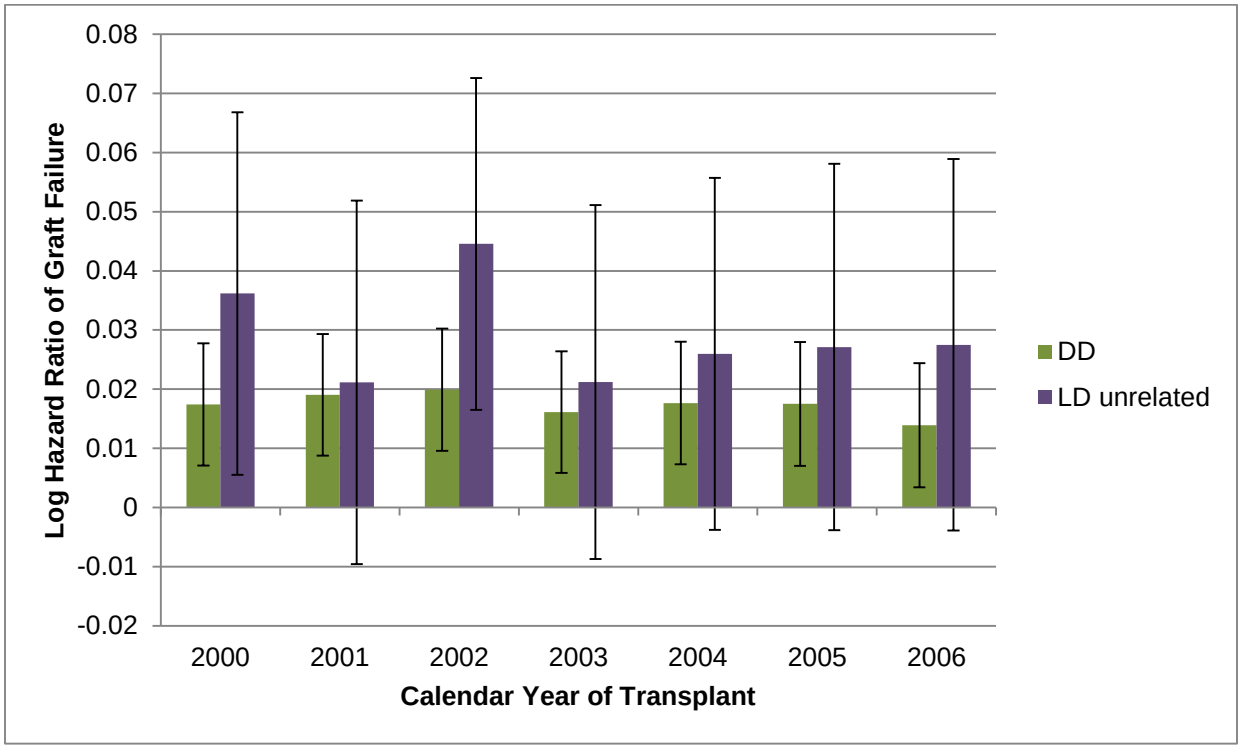


Summarized data over 10 separate analyses of imputed amino acid mismatches in HLA-A, -B, and -DRB1 loci. Initial DRB1 slope has a p-value of 0.07. No other slopes or splines have $p < 0.20$.

These figures show the adjusted, log-scale hazard ratio of graft failure by expected number of amino acid MMs by HLA locus. Zero HLA-A, -B, -DRB1 antigen (first field, classically two-digit) mismatched transplants were excluded from these analyses. Figures show results of models of all races combined. Estimated hazard ratios for all loci were calculated based on 0 MMs as the reference using the initial slopes plus spline factors to model the degree of change in slope at 5, 10, and 15 in a single model adjusted for patient risk factors, donor risk factors, and HLA antigen MMs. The effect of the imputation on the beta in terms of the variability of that estimate (i.e. the between-

iteration SD) was between 0.001 and 0.005 for all these runs, which is at least one log lower compared to the variance within imputations (data not shown).

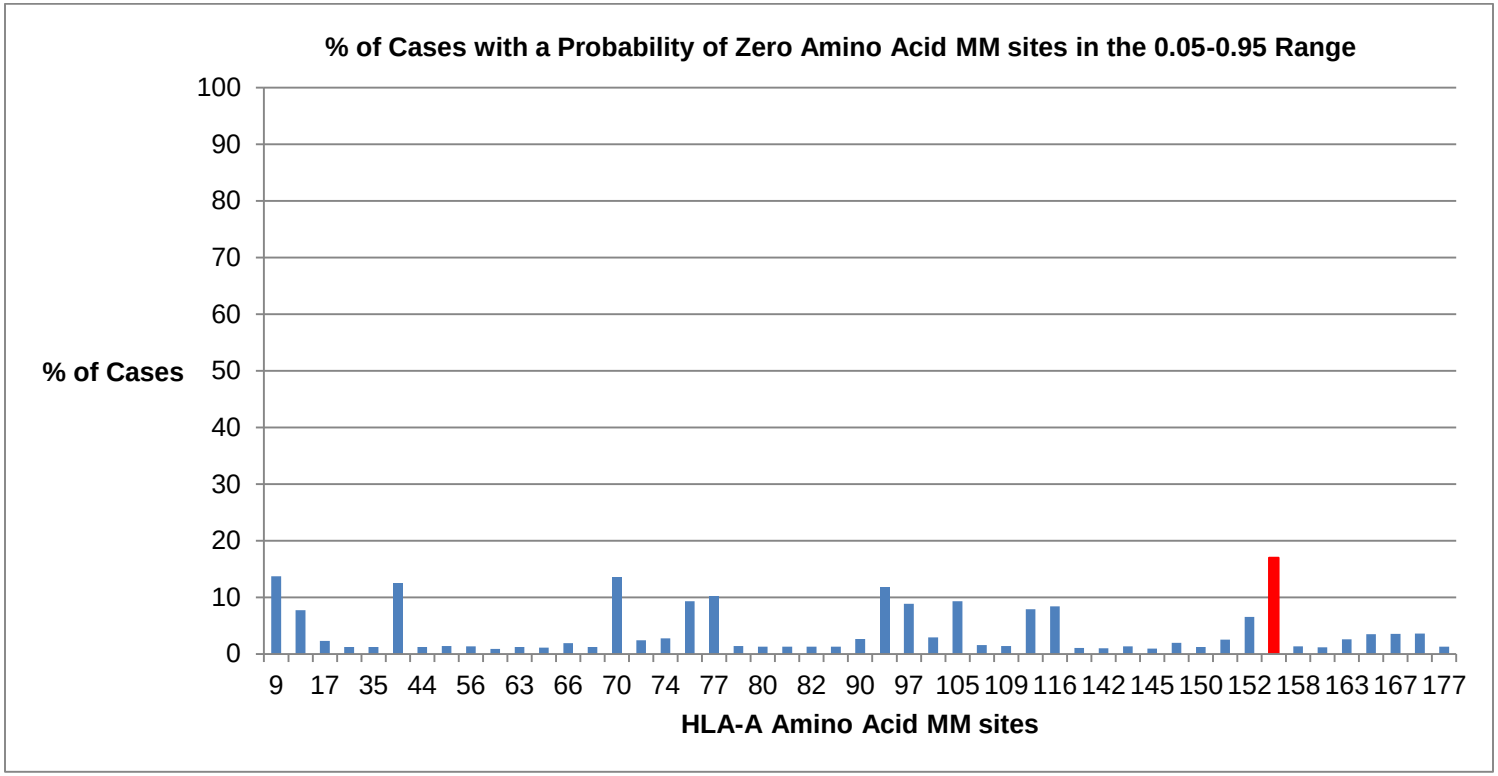
Figure S2. Graft Failure per DRB1 amino acid MM (continuous) during the first year after transplant, by calendar year of transplant



This figure shows adjusted, log-scale hazard ratios of graft failure associated with each additional imputed HLA-DRB1 amino acid MM (i.e. the linear slope of the DR AAMM variable). Models are based on 10 imputations; figure shows results of models of all races combined. Zero HLA-A, -B, -DRB1 antigen mismatched transplants were excluded from these analysis. Estimated hazard ratios were calculated based on 0 MMs as the reference, in a single model adjusted for patient risk factors, donor risk factors, and HLA-A, -B, and -DRB1 antigen MMs. *P* values were statistically significant for deceased donor transplants ($p < 0.006$) and varied between $p<0.002$ and 0.09 for living unrelated transplants.

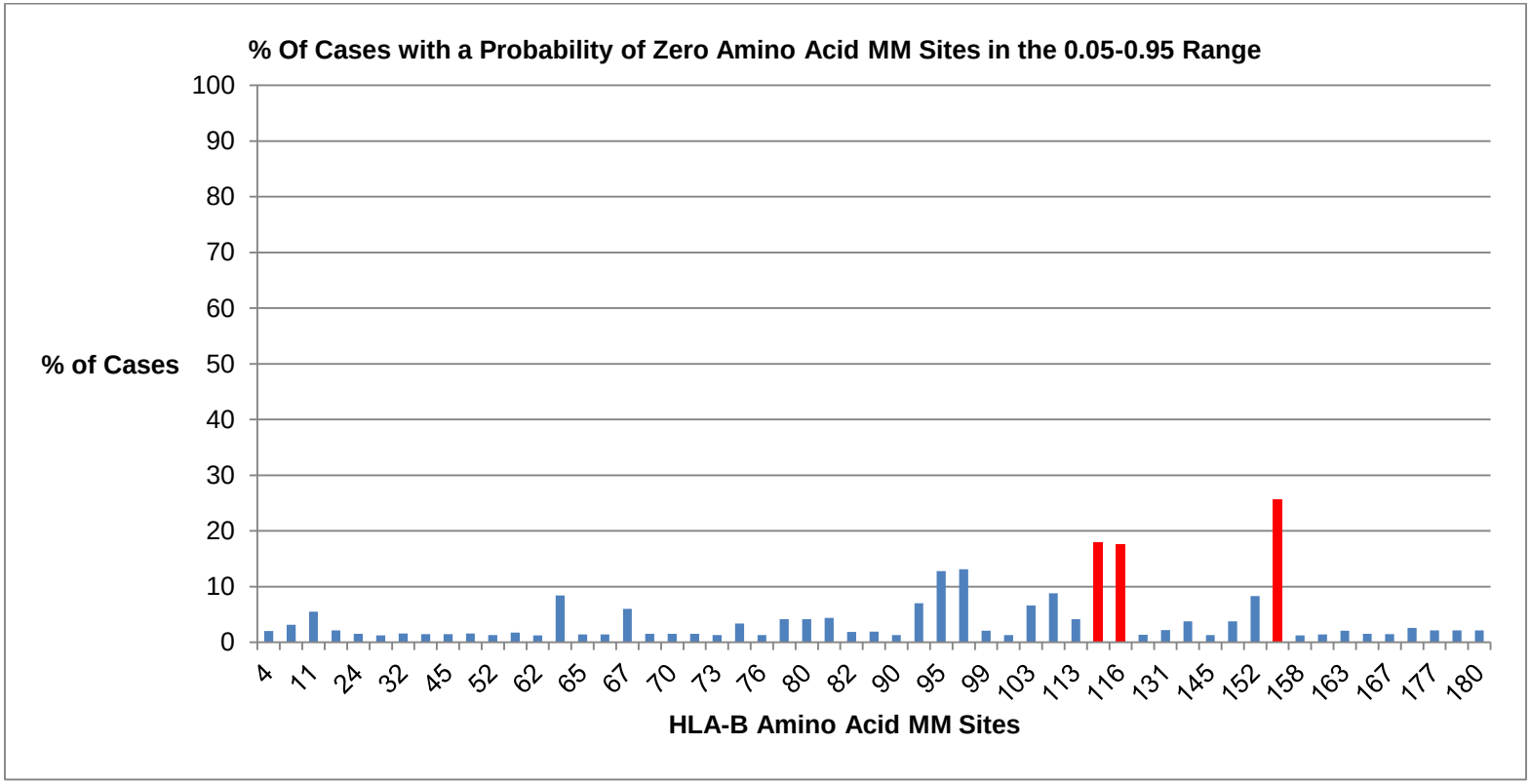
Figure S3. Percent of cases with a probability of a zero amino acid MM in the 0.05-0.95 range at each HLA-A, -B, -DRB1 variable site

Figure S3A. Percent of cases with a probability of a zero HLA-A amino acid MM in the 0.05-0.95 range



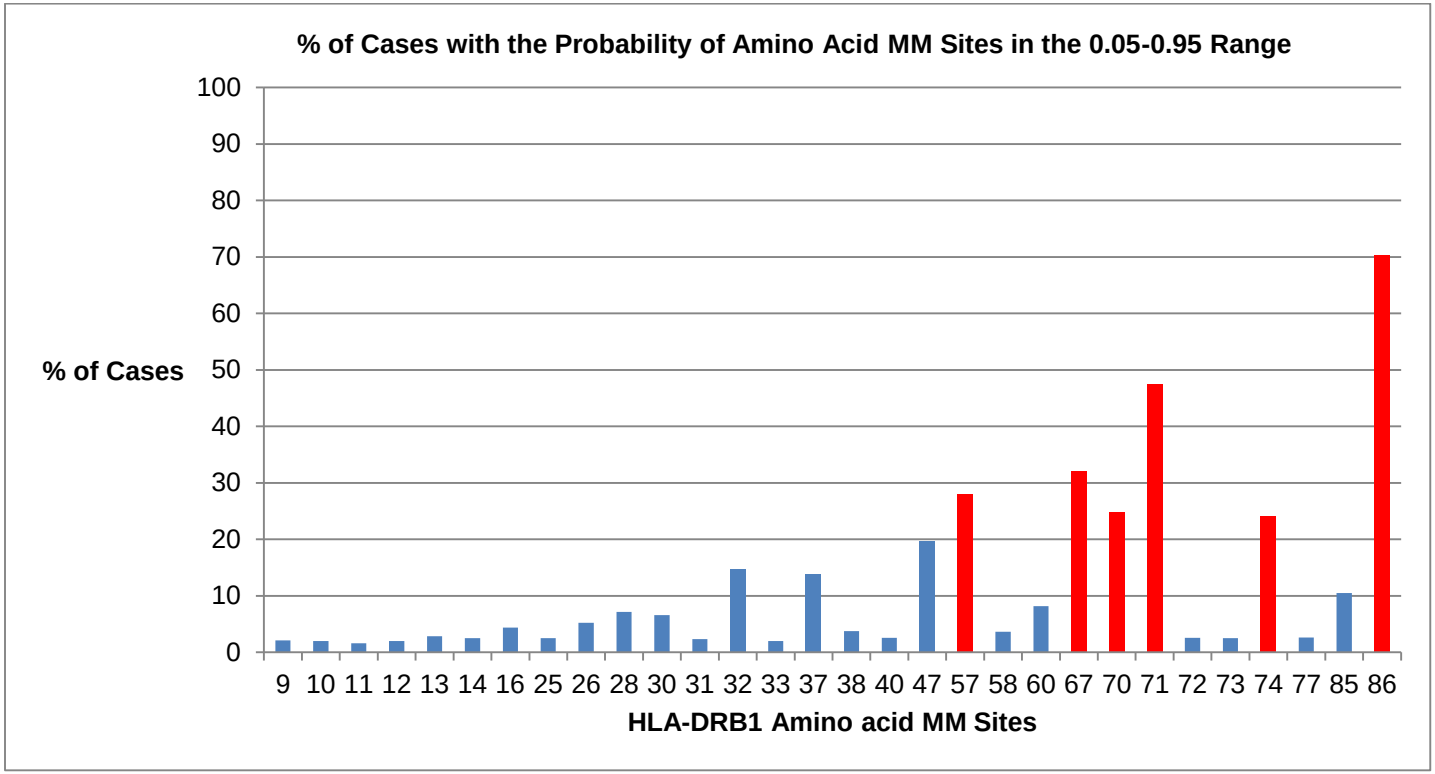
This Figure shows the percent of cases with a probability of a zero amino acid MM in the 0.05-0.95 range at each HLA-A variable site using 265,573 donor-recipient pairs from the SRTR dataset. The bar highlighted in red shows mismatch at position 156 (17% of cases).

Figure S3B. Percent of cases with a probability of a zero HLA-B amino acid MM in the 0.05-0.95 range



This Figure shows the percent of cases with a probability of a zero amino acid MM in the 0.05-0.95 range at each HLA-B variable site using 265,573 donor-recipient pairs from the SRTR dataset. Bars highlighted in red show mismatch at position 114 (18%), 116, (17%) and 156 (25%).

Figure S3C. Percent of cases with a probability of a zero HLA-DRB1 amino acid MM in the 0.05-0.95 range



This Figure shows the percent of cases with a probability of a zero amino acid MM in the 0.05-0.95 range at each HLA-DRB1 variable site using 265,573 donor-recipient pairs from the SRTR dataset. Bars highlighted in red shows mismatch at position 57 (28%), 67 (32%), 70 (25%), 71 (47%), 74 (24%), and 86 (70%).

Figure S4. Estimation of the error rate per amino acid MM site due to the imputation using simulated datasets

Figure S4A. Percent error per amino acid MM site for HLA-A

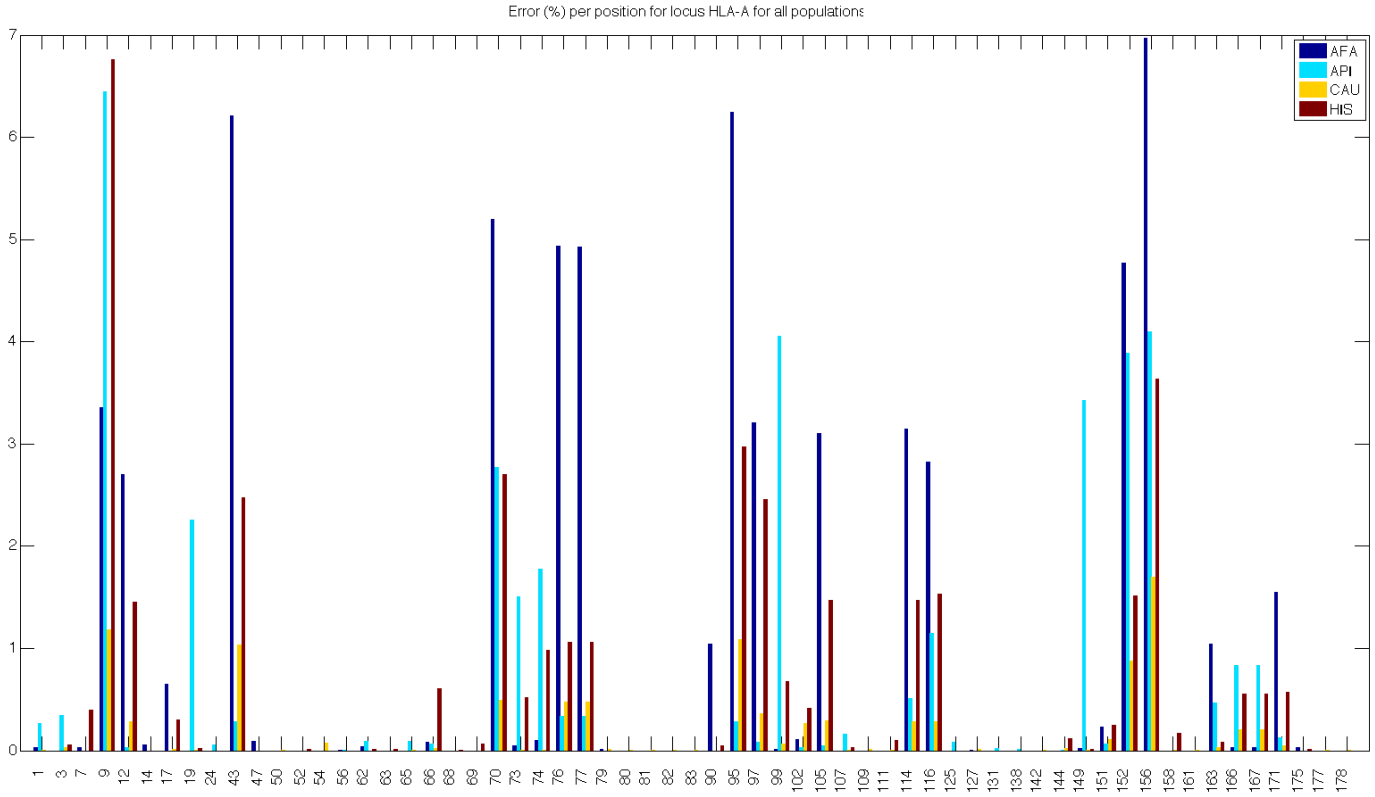


Figure S4B. Percent error per amino acid MM site for HLA-B

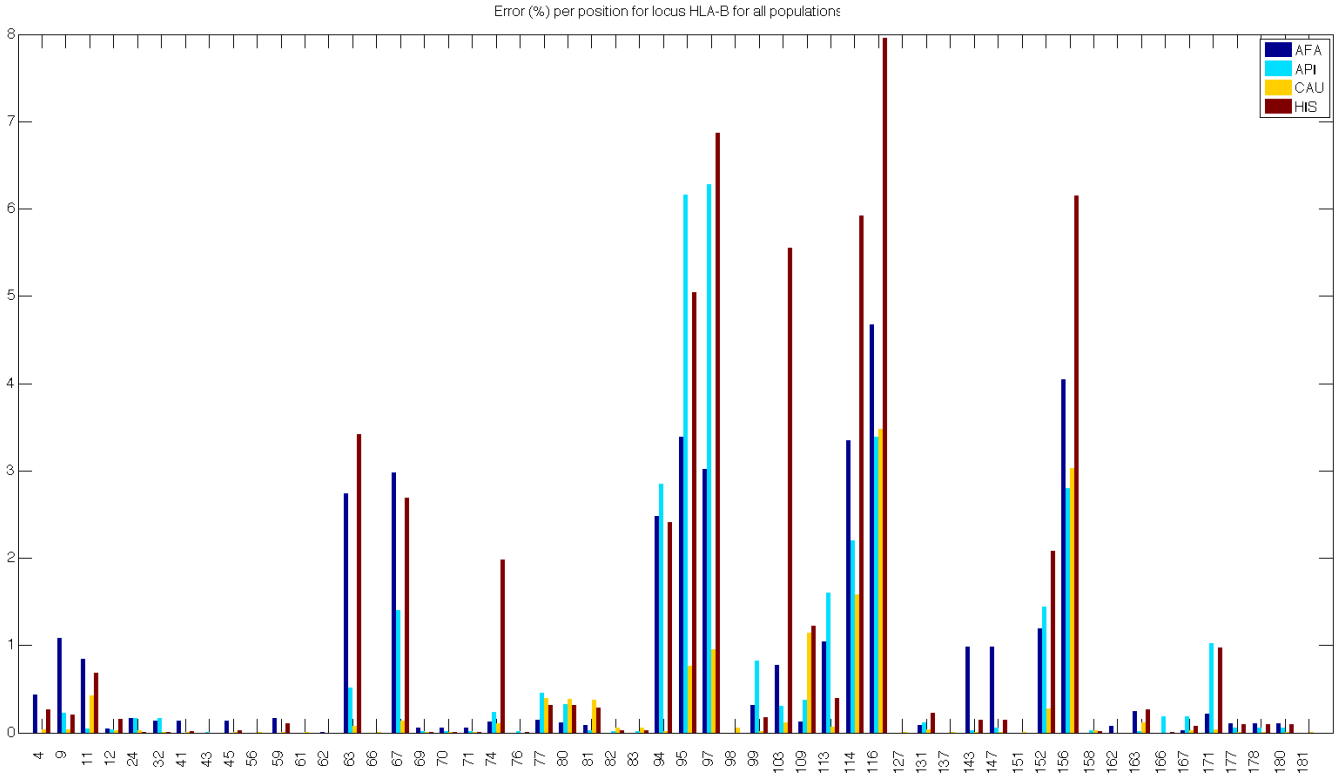
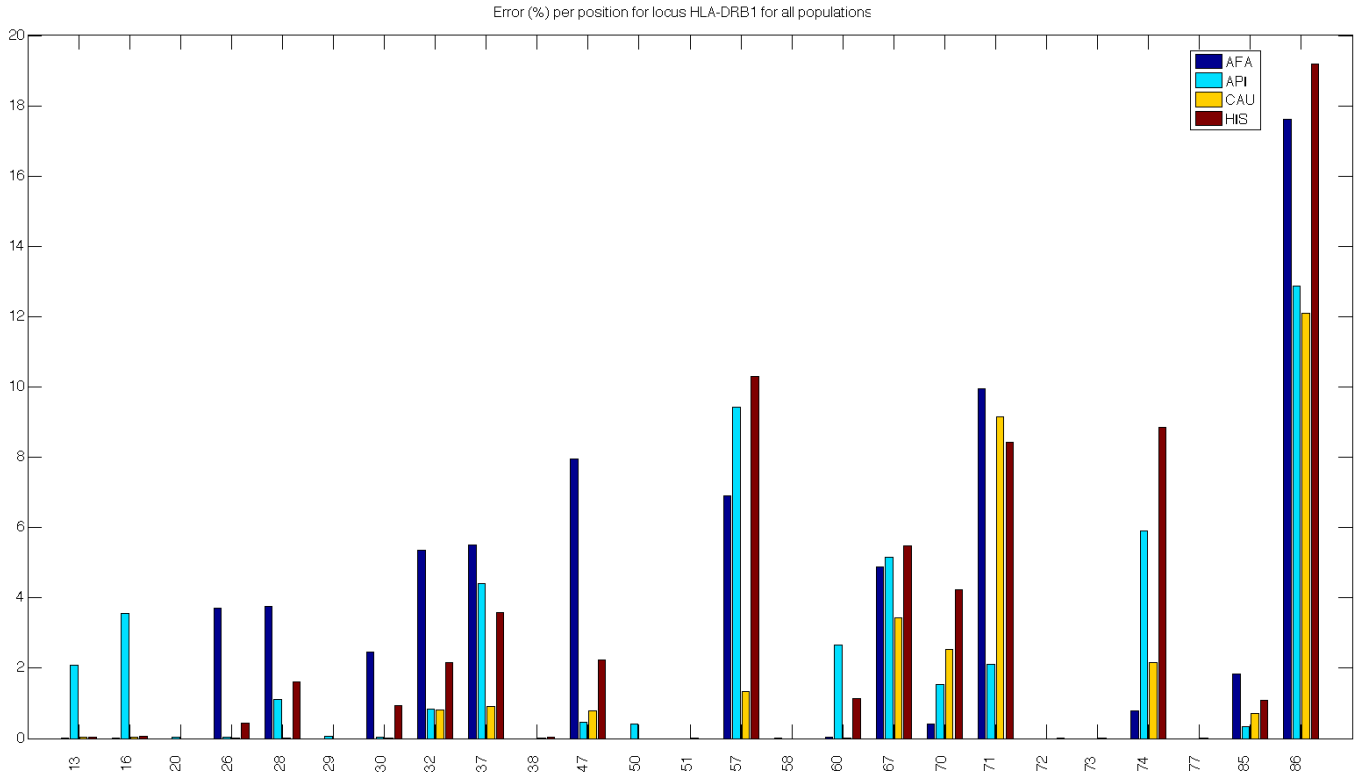


Figure S4C. Percent error per amino acid MM site for HLA-DRB1



Error rate per amino acid MM site due to the imputation of HLA-A, -B, and –DRB1 alleles were evaluated using datasets consisting of HLA genotypes with phased haplotype pairs and no allelic ambiguity, then artificially introduced ambiguities in this datasets by converting the four-digit specificity to serologic specificity (two-digit specificity). Datasets of over 200,000 individuals each were constructed for the four primary race groups in the National Marrow Donor Program

registry: White (CAU), Asian (API), Black (AFA) and Hispanic (HIS). Each individual genotype was created by randomly sampling two haplotypes from previously published US high-resolution A~B~DRB1 haplotype frequencies [3]. The original unambiguous genotypes were used as truth for the evaluation of the error rates of amino acid MMs; the error rates were computed by comparing the amino acid MM of the simulated subject to the imputed amino acid MM after conversion to serologic specificity (two-digit specificity). Error rate per amino acid MM site for HLA-A, -B, and –DRB1 alleles are shown in panel A, B, and C respectively.

Table S1. Mean frequency of expected HLA-A, -B, and -DRB1 amino acid mismatched sites by race

Locus/ position	Other	Both Asian	Don Black, Rec White	Don White, Rec Black	Both Hispanic	Both Black	Both White	All
A_1	0	0	0	0	0	0	0	0
A_3	0	0	0	0	0	0	0	0
A_7	0	0	0	0	0	0	0	0
A_9	0.82	0.64	0.87	0.83	0.68	0.74	0.51	0.64
A_12	0.03	0	0.1	0.02	0.04	0.09	0.01	0.03
A_14	0	0	0	0	0	0	0	0
A_17	0.07	0.02	0.22	0.07	0.06	0.18	0.03	0.06
A_19	0	0.02	0	0	0	0	0	0
A_24	0	0	0	0	0	0	0	0
A_31	0	0	0.01	0	0	0.01	0	0
A_35	0	0	0.01	0	0	0.01	0	0
A_43	0.03	0	0.11	0.02	0.04	0.09	0.02	0.03
A_44	0.15	0.05	0.14	0.22	0.08	0.1	0.16	0.15
A_47	0	0	0	0	0	0	0	0
A_52	0	0	0	0	0	0	0	0
A_54	0	0	0	0	0	0	0	0
A_56	0.13	0.06	0.24	0.12	0.12	0.2	0.07	0.1
A_62	0.87	0.64	0.94	0.91	0.7	0.75	0.62	0.72
A_63	0.3	0.22	0.41	0.32	0.26	0.34	0.2	0.25
A_65	0.21	0.22	0.2	0.18	0.17	0.16	0.13	0.15
A_66	0.39	0.27	0.38	0.39	0.29	0.29	0.3	0.32
A_67	0.15	0.05	0.14	0.22	0.08	0.1	0.16	0.15
A_68	0	0	0	0	0	0	0	0
A_69	0	0	0	0	0	0	0	0
A_70	0.35	0.29	0.41	0.39	0.27	0.34	0.28	0.31
A_73	0.11	0.14	0.14	0.08	0.1	0.12	0.06	0.08
A_74	0.3	0.19	0.28	0.32	0.24	0.2	0.24	0.26
A_76	0.54	0.34	0.59	0.62	0.41	0.45	0.45	0.49
A_77	0.44	0.28	0.47	0.51	0.34	0.36	0.38	0.41
A_79	0.25	0.22	0.23	0.26	0.2	0.18	0.19	0.21
A_80	0.25	0.22	0.23	0.26	0.2	0.18	0.18	0.21
A_81	0.25	0.22	0.23	0.26	0.2	0.18	0.18	0.21
A_82	0.25	0.22	0.23	0.26	0.2	0.18	0.18	0.21
A_83	0.25	0.22	0.23	0.26	0.2	0.18	0.18	0.21
A_90	0.29	0.25	0.28	0.37	0.17	0.2	0.25	0.27
A_95	0.57	0.44	0.56	0.57	0.46	0.42	0.42	0.47
A_97	0.71	0.5	0.76	0.76	0.54	0.63	0.55	0.61
A_99	0.22	0.3	0.2	0.18	0.18	0.16	0.13	0.16
A_102	0.01	0.02	0	0.01	0	0	0	0.01
A_105	0.33	0.26	0.36	0.4	0.22	0.29	0.28	0.3
A_107	0.3	0.19	0.27	0.32	0.24	0.19	0.25	0.26
A_109	0.06	0.01	0.1	0.08	0.04	0.09	0.05	0.06
A_111	0	0	0	0	0	0	0	0
A_114	0.71	0.49	0.86	0.83	0.55	0.75	0.52	0.62

Locus/ position	Other	Both Asian	Don Black, Rec White	Don White, Rec Black	Both Hispanic	Both Black	Both White	All
A_116	0.48	0.31	0.59	0.55	0.38	0.54	0.35	0.42
A_125	0	0	0	0	0	0	0	0
A_127	0.39	0.29	0.41	0.4	0.29	0.33	0.31	0.34
A_131	0	0	0	0	0	0	0	0
A_138	0	0	0	0	0	0	0	0
A_142	0.34	0.2	0.35	0.35	0.28	0.26	0.27	0.29
A_144	0.34	0.25	0.56	0.44	0.25	0.34	0.23	0.29
A_145	0.34	0.2	0.35	0.35	0.28	0.26	0.27	0.29
A_149	0.1	0.14	0.14	0.11	0.06	0.11	0.07	0.09
A_150	0.15	0.05	0.14	0.22	0.08	0.1	0.16	0.15
A_151	0.29	0.17	0.53	0.37	0.22	0.34	0.18	0.25
A_152	0.58	0.43	0.72	0.74	0.37	0.6	0.5	0.55
A_156	0.7	0.5	0.67	0.71	0.51	0.45	0.54	0.58
A_158	0.15	0.05	0.14	0.22	0.08	0.1	0.16	0.15
A_161	0.15	0.03	0.14	0.21	0.08	0.11	0.16	0.15
A_163	0.28	0.23	0.26	0.41	0.17	0.17	0.25	0.27
A_166	0.29	0.24	0.29	0.34	0.22	0.21	0.24	0.26
A_167	0.29	0.24	0.29	0.34	0.22	0.21	0.24	0.26
A_171	0.02	0	0.04	0.02	0.03	0.04	0.01	0.02
A_175	0	0	0	0	0	0	0	0
A_177	0	0	0	0	0	0	0	0
A_178	0	0	0	0	0	0	0	0
B_4	0.01	0	0.03	0	0.01	0.03	0	0.01
B_9	0.38	0.23	0.38	0.46	0.27	0.25	0.36	0.36
B_11	0.33	0.23	0.29	0.35	0.26	0.23	0.26	0.28
B_12	0.35	0.23	0.32	0.37	0.27	0.25	0.27	0.29
B_24	0.68	0.48	0.69	0.74	0.52	0.54	0.53	0.58
B_30	0.06	0.02	0.05	0.08	0.05	0.04	0.06	0.06
B_32	0.32	0.22	0.28	0.37	0.24	0.21	0.26	0.28
B_41	0.31	0.24	0.27	0.35	0.23	0.2	0.25	0.27
B_43	0	0	0	0	0	0	0	0
B_45	0.83	0.68	0.82	0.87	0.61	0.65	0.63	0.7
B_46	0.19	0.24	0.15	0.21	0.11	0.12	0.14	0.16
B_52	0	0.03	0	0	0	0	0	0
B_59	0	0	0.01	0	0	0	0	0
B_61	0	0	0	0	0	0	0	0
B_62	0.08	0.07	0.19	0.08	0.05	0.15	0.05	0.07
B_63	0.39	0.3	0.4	0.41	0.31	0.34	0.3	0.33
B_65	0.09	0.07	0.23	0.1	0.06	0.17	0.06	0.08
B_66	0.1	0.13	0.23	0.1	0.06	0.17	0.06	0.09
B_67	0.82	0.6	0.95	0.94	0.62	0.79	0.64	0.72
B_69	0.29	0.26	0.38	0.32	0.16	0.29	0.23	0.26
B_70	0.35	0.28	0.5	0.43	0.18	0.4	0.29	0.33
B_71	0.28	0.22	0.38	0.32	0.16	0.29	0.23	0.25
B_73	0	0	0	0	0	0	0	0
B_74	0.33	0.21	0.33	0.37	0.22	0.25	0.28	0.29
B_76	0.01	0.06	0	0	0	0	0	0
B_77	0.42	0.3	0.46	0.46	0.31	0.34	0.35	0.38

Locus/ position	Other	Both Asian	Don Black, Rec White	Don White, Rec Black	Both Hispanic	Both Black	Both White	All
B_80	0.51	0.4	0.59	0.6	0.37	0.43	0.41	0.45
B_81	0.34	0.26	0.39	0.37	0.27	0.31	0.27	0.3
B_82	0.36	0.28	0.39	0.38	0.28	0.31	0.29	0.31
B_83	0.36	0.28	0.39	0.38	0.28	0.31	0.29	0.31
B_90	0	0	0	0	0	0	0	0
B_94	0.32	0.25	0.37	0.34	0.26	0.3	0.25	0.28
B_95	0.62	0.49	0.65	0.65	0.52	0.54	0.46	0.53
B_97	0.7	0.5	0.7	0.76	0.57	0.55	0.56	0.61
B_98	0	0	0	0	0	0	0	0
B_99	0.03	0.03	0.02	0.04	0.01	0.02	0.02	0.03
B_103	0.29	0.23	0.44	0.34	0.25	0.32	0.19	0.25
B_109	0.02	0	0	0.02	0.02	0	0.02	0.02
B_113	0.31	0.18	0.29	0.34	0.22	0.23	0.25	0.27
B_114	0.49	0.35	0.51	0.52	0.37	0.37	0.39	0.42
B_116	0.91	0.66	0.93	0.96	0.74	0.68	0.71	0.78
B_127	0	0	0	0	0	0	0	0
B_131	0.3	0.23	0.29	0.34	0.21	0.22	0.25	0.27
B_143	0.09	0.13	0.08	0.1	0.06	0.06	0.07	0.08
B_145	0.03	0.06	0.02	0.04	0.01	0.01	0.03	0.03
B_147	0.09	0.13	0.08	0.1	0.06	0.06	0.07	0.08
B_151	0	0	0	0	0	0	0	0
B_152	0.36	0.3	0.37	0.39	0.28	0.3	0.28	0.31
B_156	0.51	0.28	0.51	0.6	0.32	0.36	0.49	0.49
B_158	0.09	0.08	0.03	0.07	0.09	0.02	0.05	0.06
B_162	0	0	0	0	0	0	0	0
B_163	0.6	0.48	0.59	0.69	0.45	0.4	0.51	0.54
B_166	0	0.01	0	0	0	0	0	0
B_167	0.17	0.09	0.18	0.2	0.12	0.14	0.15	0.16
B_171	0.21	0.12	0.17	0.23	0.18	0.14	0.16	0.17
B_177	0.28	0.19	0.29	0.32	0.16	0.22	0.25	0.25
B_178	0.2	0.17	0.18	0.24	0.11	0.13	0.18	0.18
B_180	0.28	0.19	0.29	0.32	0.16	0.22	0.25	0.25
DRB1_9	0.37	0.36	0.37	0.38	0.25	0.32	0.31	0.33
DRB1_10	0.37	0.33	0.38	0.4	0.3	0.33	0.29	0.32
DRB1_11	0.73	0.61	0.68	0.76	0.58	0.54	0.63	0.65
DRB1_12	0.34	0.29	0.35	0.37	0.28	0.3	0.27	0.3
DRB1_13	0.86	0.75	0.81	0.85	0.7	0.68	0.69	0.74
DRB1_14	0.14	0.08	0.13	0.17	0.1	0.11	0.13	0.13
DRB1_16	0.14	0.24	0.15	0.08	0.12	0.13	0.06	0.09
DRB1_20	0	0	0	0	0	0	0	0
DRB1_25	0.14	0.08	0.13	0.17	0.1	0.11	0.13	0.13
DRB1_26	0.34	0.32	0.37	0.37	0.23	0.31	0.3	0.31
DRB1_28	0.35	0.33	0.38	0.35	0.27	0.34	0.26	0.29
DRB1_29	0	0	0	0	0	0	0	0
DRB1_30	0.44	0.39	0.6	0.52	0.29	0.51	0.33	0.39
DRB1_31	0.18	0.14	0.19	0.19	0.13	0.15	0.14	0.16
DRB1_32	0.29	0.27	0.33	0.31	0.24	0.3	0.24	0.26
DRB1_33	0.19	0.14	0.1	0.2	0.2	0.06	0.16	0.16

Locus/ position	Other	Both Asian	Don Black, Rec White	Don White, Rec Black	Both Hispanic	Both Black	Both White	All
DRB1_37	0.78	0.72	0.8	0.82	0.59	0.72	0.64	0.7
DRB1_38	0.08	0.16	0.1	0.05	0.04	0.09	0.04	0.05
DRB1_40	0.03	0.03	0.03	0.02	0.02	0.03	0.01	0.02
DRB1_47	0.37	0.31	0.38	0.39	0.27	0.35	0.29	0.32
DRB1_50	0	0	0	0	0	0	0	0
DRB1_57	0.42	0.52	0.39	0.38	0.3	0.34	0.28	0.33
DRB1_58	0.13	0.08	0.17	0.13	0.09	0.14	0.11	0.12
DRB1_60	0.27	0.32	0.26	0.27	0.16	0.22	0.21	0.23
DRB1_67	0.63	0.53	0.62	0.62	0.5	0.54	0.49	0.54
DRB1_70	0.51	0.48	0.49	0.48	0.38	0.41	0.39	0.43
DRB1_71	0.64	0.37	0.7	0.7	0.42	0.65	0.6	0.61
DRB1_72	0	0	0	0	0	0	0	0
DRB1_73	0.23	0.13	0.25	0.25	0.17	0.22	0.21	0.22
DRB1_74	0.59	0.48	0.52	0.52	0.55	0.47	0.43	0.48
DRB1_77	0.11	0.06	0.16	0.13	0.08	0.14	0.12	0.12
DRB1_78	0.17	0.16	0.16	0.18	0.11	0.14	0.14	0.15
DRB1_85	0.09	0.13	0.15	0.06	0.07	0.11	0.04	0.06
DRB1_86	0.43	0.36	0.44	0.44	0.39	0.39	0.35	0.39

The data have already been limited by excluding amino acid locus/positions whose maximum probability for any individual transplant was < 2.5 %. Among the remaining 158 locus/positions, there were 125 locus/positions that had a mean of the composite (X = 0 * P(0MM) + 1 MM probability + 2 * 2 MM probability) of at least 0.01 mean within at least one race group. Highlighted rows indicate locus/positions where the mean expected mismatch is at least 1% for at least one donor/recipient race combination.

Table S2. Positional specificity of HLA-DRB1 amino acid mismatches

Example 1 - Recipient: HLA-DRB1*13:01 versus Donor: HLA-DRB1*09:01

HLA Alleles (four-digit specificity)‡	Total Number of Amino Acid MMs§ (Donor versus Recipient)	Variation in the Position of Amino Acid MMs§ depending on the HLA-DRB1*13 Allele
Recipient: HLA-DRB1*13:01 (5.6% in Caucasian) Donor: HLA-DRB1*09:01	19	DRB1*13:01 is used as the reference allele
Recipient: HLA-DRB1*13:02 (7.3% in Black) Donor: HLA-DRB1*09:01	18	86
Recipient: HLA-DRB1*13:03 (3.7% in Black) Donor: HLA-DRB1*09:01	17	47, 86
Recipient: HLA-DRB1*13:04 (1.4% in Black) Donor: HLA-DRB1*09:01	19	32
Recipient: HLA-DRB1*13:05 (0.7% in Hispanic) Donor: HLA-DRB1*09:01	16	67, 71, 86
Recipient: HLA-DRB1*13:06 (< 0.05%) Donor: HLA-DRB1*09:01	18	57
Recipient: HLA-DRB1*13:07 (< 0.05%) Donor: HLA-DRB1*09:01	15	32, 47, 67, 71
Recipient: HLA-DRB1*13:08 (< 0.05%) Donor: HLA-DRB1*09:01	19	47, 37
Recipient: HLA-DRB1*13:09 (< 0.05%) Donor: HLA-DRB1*09:01	19	None
Recipient: HLA-DRB1*13:10 (< 0.05%) Donor: HLA-DRB1*09:01	19	None

Recipient: HLA-DRB1*13:01 versus Donor: HLA-DRB1*09:01

Amino Acid Mismatch	Exon	Pockets	Position of Mismatch	Binding Sites¶
Glutamic Acid--> Lysine	2	B, C	9	PEP
Tyrosine--> Glutamine	2		10	
Serine--> Aspartic Acid	2		11	PEP
Threonine--> Lysine	2		12	
Serine--> Phenylalanine	2		13	PEP
Phenylalanine--> Tyrosine	2		26	PEP
Aspartic Acid--> Histidine	2		28	PEP
Tyrosine--> Glycine	2		30	PEP
Phenylalanine--> Isoleucine	2		31	
Histidine--> Tyrosine	2		32	PEP
Phenylalanine--> Tyrosine	2		47	PEP
Aspartic Acid--> Valine	2		57	PEP
Tyrosine--> Serine	2		60	PEP
Isoleucine--> Phenylalanine	2	B	67	PEP
Aspartic Acid--> Arginine	2	A, B, C	70	PEP
Glutamic Acid--> Arginine	2		71	PEP
Alanine--> Glutamic Acid	2	C	74	PEP
Tyrosine--> Valine	2		78	PEP
Valine--> Glycine	2		86	PEP

Example 2 - Recipient: HLA-DR13 versus Donor: HLA-DR14

HLA Alleles (four-digit specificity)‡	Total Number of Amino Acid MMs§ (Donor versus Recipient)	Variation in the Position of Amino Acid MMs§ depending on the HLA-DRB1*13 Allele
Recipient: HLA-DRB1*13:01 (5.6% in Caucasian) Donor: HLA-DRB1*14:01	8	DRB1*13:01 is used as the reference allele
Recipient: HLA-DRB1*13:02 (7.3% in Black) Donor: HLA-DRB1*14:01	9	86
Recipient: HLA-DRB1*13:03 (3.7% in Black) Donor: HLA-DRB1*14:01	9	32, 47, 86
Recipient: HLA-DRB1*13:04 (1.4% in AA) Donor: HLA-DRB1*14:01	9	
Recipient: HLA-DRB1*13:05 (0.7% in Hispanic) Donor: HLA-DRB1*14:01	8	71, 86
Recipient: HLA-DRB1*13:06 (< 0.05%) Donor: HLA-DRB1*14:01	7	71

Recipient: HLA-DRB1*13:01 versus Donor: HLA-DRB1*14:01

Amino Acid Mismatch	Exon	Pockets	Position of Mismatch	Binding Sites¶
Asparagine--> Phenylalanine	2		37	PEP
Phenylalanine--> Tyrosine	2		47	PEP
Aspartic Acid--> Alanine	2		57	PEP
Tyrosine--> Histidine	2		60	PEP
Isoleucine--> Leucine	2	B	67	PEP
Aspartic Acid--> Arginine	2	A, B, C	70	PEP
Glutamic Acid--> Arginine	2		71	PEP
Alanine--> Glutamic Acid	2	C	74	PEP

‡ Allele frequencies were obtained from http://allelefrequencies.net/default.asp

§ Comparison of the amino acid sequence alignment between HLA alleles (four-digit specificity) was performed using the allele alignment tool from https://www.mh-hannover.de/institute/transfusion/histocheck/index.html, and the IMGT/HLA database (https://www.ebi.ac.uk/ipd/imgt/hla/)

¶ PEP, indicates peptide contact site; TCR, indicates T cell receptor contact site

Table S3. Graft survival model results using mean numbers of expected amino acid mismatches

Cox proportional-hazards models are adjusted for all covariates (including HLA variables), and are run 1-at-a-time for each locus/position. Since there were 1750 models run, a Bonferroni correction was applied and only hazard ratios associated with p-values < 0.05/1750 were highlighted.

Locus/	Cadaveric donor							Living Donor						
Position	dw rw	db rb	dh rh	dw rb	db rw	da ra	All	dw rw	db rb	dh rh	dw rb	db rw	da ra	All
A_102	1.24	0.37	2.54	1.19	0.70	1.00	1.05	1.19	2.74	0.58	1.41	> 100	0.93	0.98
A_105	1.05	0.98	1.01	1.02	0.96	1.12	1.03	1.09	1.06	1.09	1.00	0.88	1.14	1.09
A_107	1.04	1.01	0.99	1.04	1.01	1.20	1.03	1.12	1.06	1.08	0.87	0.87	1.30	1.10
A_109	1.04	0.99	0.95	0.99	1.01	0.52	1.03	1.06	1.05	1.34	1.33	1.70	1.84	1.08
A_114	1.06	1.03	1.02	1.02	1.02	1.26	1.05	1.10	1.16	1.15	1.12	1.18	0.96	1.11
A_116	1.05	1.02	1.04	1.03	1.02	1.13	1.05	1.10	1.11	1.11	1.06	0.98	1.06	1.10
A_12	0.99	1.05	0.88	1.11	0.94	> 100	1.09	1.05	1.03	1.19	0.67	1.36	0.01	1.04
A_127	1.04	1.06	1.06	1.02	1.03	1.10	1.04	1.11	1.07	1.11	1.09	0.92	1.06	1.10
A_142	1.04	1.01	1.04	1.04	1.01	1.19	1.03	1.11	1.05	1.10	0.92	0.98	1.18	1.09
A_144	1.04	1.04	1.04	1.04	1.03	1.09	1.04	1.08	1.06	1.09	0.97	0.90	1.02	1.07
A_145	1.04	1.01	1.04	1.04	1.01	1.19	1.03	1.11	1.05	1.10	0.92	0.98	1.18	1.09
A_149	1.01	0.98	1.08	0.99	0.92	1.44	1.00	1.04	1.08	1.12	0.87	1.08	1.09	1.05
A_150	1.07	0.97	1.16	1.06	1.08	0.83	1.06	1.12	1.29	1.24	0.94	0.85	1.01	1.14
A_151	1.04	1.05	1.06	1.03	1.04	0.96	1.04	1.09	1.08	1.04	1.04	0.85	1.09	1.08
A_152	1.05	1.03	1.05	1.02	1.04	1.38	1.05	1.12	1.13	1.01	0.96	0.78	0.96	1.12
A_156	1.06	1.04	1.04	1.04	1.00	1.16	1.04	1.12	1.14	1.05	1.06	1.08	1.18	1.11
A_158	1.07	0.97	1.16	1.06	1.08	0.83	1.06	1.12	1.29	1.25	0.94	0.85	1.01	1.15
A_161	1.03	1.01	1.10	0.99	1.03	1.52	1.02	1.07	1.06	0.90	1.11	0.55	0.71	1.06
A_163	1.05	0.96	1.02	1.02	1.01	1.23	1.03	1.09	1.16	1.05	0.95	0.90	0.97	1.09
A_166	1.06	1.05	1.01	1.03	1.04	0.96	1.04	1.10	1.14	1.06	1.12	1.03	1.28	1.11
A_167	1.06	1.05	1.01	1.03	1.04	0.96	1.04	1.10	1.14	1.06	1.12	1.03	1.28	1.11
A_17	1.05	1.03	0.97	1.06	1.03	1.77	1.07	1.02	1.06	0.94	0.73	1.33	1.16	1.04
A_171	1.04	1.08	1.20	1.01	1.07	0.09	1.09	1.03	1.28	0.84	1.11	3.65	0.32	1.03
A_19	> 100	.	0.42	0.03	.	2.97	0.89	0.37	.	0.00	0.00	.	1.04	0.94
A_43	1.07	1.05	0.72	1.08	0.96	7.15	1.13	0.96	1.10	1.14	1.22	0.66	0.74	1.07

Locus/ Position	Cadaveric donor							Living Donor						
	dw rw	db rb	dh rh	dw rb	db rw	da ra	All	dw rw	db rb	dh rh	dw rb	db rw	da ra	All
A_44	1.07	0.97	1.16	1.06	1.08	0.83	1.06	1.12	1.29	1.24	0.94	0.85	1.01	1.14
A_56	1.06	1.02	1.03	1.05	1.01	1.39	1.06	1.02	1.05	1.04	1.10	1.41	1.44	1.04
A_62	1.05	1.02	1.06	1.04	1.03	1.14	1.05	1.13	1.14	1.13	1.00	0.93	1.18	1.13
A_63	1.03	1.00	1.13	1.03	0.95	0.92	1.03	1.07	1.12	1.06	1.00	1.19	0.95	1.07
A_65	1.04	1.05	0.96	0.99	1.06	1.02	1.03	1.05	1.12	1.04	1.11	0.89	1.37	1.07
A_66	1.04	1.05	1.01	1.02	1.04	1.09	1.04	1.11	1.07	1.11	1.04	1.01	0.97	1.09
A_67	1.07	0.97	1.16	1.06	1.08	0.83	1.06	1.12	1.29	1.24	0.94	0.85	1.01	1.14
A_70	1.03	1.06	1.08	1.03	1.01	1.11	1.04	1.08	1.12	1.09	1.06	0.65	1.13	1.09
A_73	1.06	1.06	1.02	1.00	1.01	0.95	1.05	1.01	1.17	1.02	1.29	1.67	0.88	1.05
A_74	1.04	1.00	0.98	1.04	1.00	1.27	1.03	1.12	1.06	1.09	0.86	0.81	1.30	1.10
A_76	1.05	1.00	1.02	1.04	1.01	1.04	1.04	1.09	1.16	1.08	1.02	0.85	1.30	1.10
A_77	1.04	1.00	1.01	1.02	0.95	1.10	1.03	1.07	1.09	1.09	0.96	0.87	1.34	1.07
A_79	1.03	1.04	0.96	1.01	1.03	1.03	1.03	1.06	1.12	1.10	1.24	0.87	1.43	1.08
A_80	1.04	1.04	0.96	1.01	1.03	1.03	1.03	1.06	1.12	1.10	1.26	0.79	1.43	1.08
A_81	1.04	1.04	0.96	1.01	1.03	1.03	1.03	1.06	1.12	1.10	1.26	0.79	1.43	1.08
A_82	1.04	1.04	0.96	1.01	1.03	1.03	1.03	1.06	1.12	1.10	1.26	0.79	1.43	1.08
A_83	1.04	1.04	0.96	1.01	1.03	1.03	1.03	1.06	1.12	1.10	1.26	0.79	1.43	1.08
A_9	1.03	1.04	1.07	1.02	1.01	1.25	1.04	1.08	1.14	1.05	1.03	1.05	1.32	1.09
A_90	1.05	0.97	1.05	1.02	1.00	1.10	1.03	1.09	1.16	1.01	1.00	0.90	1.08	1.09
A_95	1.04	1.04	0.96	1.02	1.05	1.15	1.04	1.11	1.03	1.11	0.96	0.89	1.27	1.09
A_97	1.05	1.00	1.01	1.03	1.04	1.22	1.04	1.13	1.13	1.07	0.83	0.84	1.17	1.12
A_99	1.04	1.05	0.96	0.99	1.06	1.00	1.03	1.05	1.11	1.05	1.11	0.89	1.32	1.07
B_103	1.07	1.03	1.09	1.01	1.00	1.14	1.04	1.03	1.06	1.17	1.20	0.93	1.04	1.05
B_109	1.07	1.67	0.89	0.91	1.76	1.64	0.97	0.80	0.51	1.24	0.53	0.00	0.22	0.82
B_11	1.05	1.04	1.08	1.01	1.04	1.12	1.03	1.08	1.06	1.10	1.20	0.53	0.77	1.07
B_113	1.06	1.04	0.94	1.00	1.05	0.98	1.04	1.08	1.08	1.02	1.02	1.02	0.78	1.07
B_114	1.05	1.03	1.11	1.01	1.02	1.33	1.04	1.11	1.05	1.03	1.24	0.92	1.02	1.09
B_116	1.08	1.06	1.06	1.02	1.05	1.25	1.05	1.12	1.08	1.11	1.00	0.96	1.14	1.11
B_12	1.05	1.04	1.07	1.01	1.05	1.12	1.03	1.08	1.07	1.13	1.11	0.46	0.77	1.07
B_131	1.05	1.06	1.17	1.02	1.07	1.07	1.04	1.07	1.00	1.15	1.07	1.13	1.01	1.07
B_143	1.06	0.95	1.11	1.07	1.04	1.04	1.06	1.11	1.10	1.16	1.50	0.76	1.24	1.11
B_145	1.07	1.02	1.41	1.03	1.16	1.21	1.04	1.11	0.94	1.26	0.70	1.04	1.41	1.11
B_147	1.06	0.95	1.11	1.07	1.04	1.04	1.06	1.11	1.10	1.16	1.50	0.76	1.24	1.11
B_152	1.04	1.05	0.96	0.99	1.02	1.28	1.02	1.07	1.06	1.06	1.20	0.72	1.04	1.07

Locus/ Position	Cadaveric donor							Living Donor						
	dw rw	db rb	dh rh	dw rb	db rw	da ra	All	dw rw	db rb	dh rh	dw rb	db rw	da ra	All
B_156	1.07	1.05	1.03	1.01	1.08	0.96	1.05	1.09	1.02	1.06	1.20	1.02	1.02	1.09
B_158	1.03	0.95	1.02	1.05	0.85	1.69	1.00	1.05	1.11	1.11	0.63	0.00	0.98	1.05
B_163	1.07	1.03	1.11	1.04	1.04	1.16	1.05	1.14	1.07	1.08	1.08	0.77	0.93	1.12
B_167	1.06	1.02	1.00	1.01	1.01	0.72	1.05	1.09	1.04	1.02	0.92	0.89	1.05	1.07
B_171	1.03	1.04	0.99	1.01	0.99	1.08	1.02	1.08	1.19	1.05	0.89	1.41	1.00	1.08
B_177	1.05	1.07	1.15	1.03	1.08	1.00	1.05	1.07	1.01	1.10	1.10	1.05	1.02	1.07
B_178	1.06	1.03	1.08	1.02	1.03	0.97	1.04	1.07	1.03	0.99	1.48	1.04	1.00	1.07
B_180	1.05	1.07	1.15	1.03	1.08	1.00	1.05	1.07	1.01	1.10	1.10	1.05	1.02	1.07
B_24	1.07	1.02	1.05	1.04	1.05	1.24	1.05	1.14	1.12	1.14	1.00	0.78	0.88	1.12
B_30	1.02	1.04	1.00	1.02	0.85	1.06	1.00	1.08	1.20	0.97	0.80	0.76	0.76	1.08
B_32	1.05	1.03	1.05	1.02	1.03	1.09	1.04	1.10	1.08	1.10	1.02	1.07	0.98	1.09
B_4	1.09	0.94	0.48	1.22	1.10	.	1.10	1.33	1.26	1.32	2.30	0.01	.	1.24
B_41	1.05	1.02	1.07	1.03	0.99	1.20	1.04	1.09	1.05	1.15	0.94	1.40	1.07	1.09
B_45	1.07	1.03	1.03	1.05	1.06	0.96	1.05	1.15	1.13	1.14	1.17	0.91	1.01	1.14
B_46	1.03	0.99	1.04	1.03	1.01	1.11	1.02	1.10	1.07	1.06	1.27	0.76	1.03	1.10
B_52	.	.	.	1.15	.	0.52	0.82	.	0.46	0.00	.	.	1.36	1.15
B_62	1.03	0.99	1.11	1.01	0.98	1.01	1.04	1.10	0.99	1.10	1.25	0.62	1.06	1.05
B_63	1.05	1.07	0.98	1.04	0.99	1.00	1.04	1.12	1.10	1.10	1.11	1.25	1.00	1.11
B_65	1.03	1.00	1.11	1.01	1.01	1.00	1.05	1.08	1.04	1.13	1.31	0.42	1.00	1.06
B_66	1.03	1.00	1.14	1.01	1.01	0.86	1.04	1.08	1.04	1.13	1.31	0.42	1.07	1.06
B_67	1.07	1.06	1.03	1.04	1.07	1.20	1.06	1.13	1.15	1.14	1.18	0.75	1.01	1.13
B_69	1.06	1.06	1.04	0.99	1.06	0.97	1.04	1.07	1.07	1.04	1.27	0.91	1.04	1.08
B_70	1.07	1.03	1.03	1.00	1.06	0.99	1.04	1.09	1.08	1.05	1.34	0.66	1.02	1.08
B_71	1.06	1.06	1.03	0.99	1.06	0.96	1.04	1.07	1.07	1.04	1.27	0.91	1.09	1.08
B_74	1.05	1.05	0.93	1.04	1.10	0.82	1.04	1.09	1.08	1.07	1.06	0.66	0.95	1.08
B_76	0.69	1.38	3.03	0.77	0.00	0.76	0.98	1.12	0.64	1.33	.	.	1.18	1.06
B_77	1.06	1.01	1.00	1.02	1.05	1.19	1.04	1.09	1.06	1.15	1.01	0.65	0.95	1.08
B_80	1.06	1.02	1.01	1.04	1.02	1.20	1.05	1.10	1.08	1.15	1.00	0.61	0.95	1.09
B_81	1.05	1.01	0.96	1.02	1.05	1.15	1.04	1.08	1.05	1.12	1.02	0.81	0.95	1.07
B_82	1.05	1.02	0.99	1.02	1.03	1.11	1.04	1.08	1.05	1.14	1.00	0.81	0.95	1.07
B_83	1.05	1.02	0.99	1.02	1.03	1.11	1.04	1.08	1.05	1.14	1.00	0.81	0.95	1.07
B_9	1.06	1.03	1.10	1.04	1.02	1.06	1.05	1.13	1.08	1.14	0.97	0.84	1.02	1.11
B_94	1.05	1.01	1.12	1.01	1.01	1.01	1.04	1.11	1.03	1.11	1.06	0.89	1.03	1.09
B_95	1.07	1.03	1.14	1.02	1.05	1.08	1.05	1.11	1.06	1.12	1.03	0.73	1.03	1.10

Locus/ Position	Cadaveric donor							Living Donor						
	dw rw	db rb	dh rh	dw rb	db rw	da ra	All	dw rw	db rb	dh rh	dw rb	db rw	da ra	All
B_97	1.07	1.06	1.12	1.03	1.08	1.01	1.05	1.14	1.08	1.10	1.07	0.89	0.96	1.12
B_99	1.08	1.06	1.06	1.05	1.11	1.79	1.08	1.11	1.10	1.40	0.95	0.33	0.83	1.12
DRB1_10	1.10	1.12	1.13	1.06	1.10	1.13	1.08	1.13	1.14	1.13	0.97	0.99	1.01	1.12
DRB1_11	1.10	1.09	1.18	1.08	1.11	1.15	1.09	1.18	1.15	1.07	1.05	0.94	1.12	1.16
DRB1_12	1.10	1.10	1.12	1.07	1.11	1.17	1.08	1.13	1.15	1.11	1.00	1.00	1.02	1.13
DRB1_13	1.10	1.09	1.20	1.08	1.08	1.25	1.09	1.20	1.17	1.10	1.08	0.92	1.19	1.18
DRB1_14	1.08	1.12	1.17	1.04	1.19	1.09	1.07	1.15	1.16	1.08	0.96	1.18	1.43	1.14
DRB1_16	1.06	0.97	1.19	1.04	0.95	1.09	1.05	1.10	1.11	1.06	1.20	0.95	1.10	1.09
DRB1_25	1.08	1.12	1.17	1.04	1.19	1.09	1.07	1.15	1.16	1.08	0.96	1.18	1.43	1.14
DRB1_26	1.09	1.07	1.13	1.07	1.07	1.13	1.08	1.11	1.12	1.01	1.16	0.78	0.88	1.10
DRB1_28	1.10	1.11	1.19	1.07	1.05	1.09	1.08	1.13	1.12	1.00	1.12	1.04	1.03	1.12
DRB1_30	1.09	1.09	1.19	1.03	1.04	1.13	1.07	1.13	1.14	1.04	1.02	1.04	1.04	1.13
DRB1_31	1.08	1.09	1.22	1.06	1.04	1.01	1.07	1.08	1.10	0.99	1.14	0.80	1.01	1.08
DRB1_32	1.08	1.09	1.07	1.05	1.07	1.07	1.07	1.15	1.08	1.08	1.18	0.86	1.09	1.13
DRB1_33	1.06	1.16	1.07	1.05	1.03	1.12	1.04	1.10	1.11	1.09	0.96	0.70	1.40	1.10
DRB1_37	1.09	1.07	1.15	1.08	1.08	1.19	1.08	1.19	1.14	1.05	1.25	0.92	1.09	1.17
DRB1_38	1.08	1.09	1.22	0.97	0.96	1.31	1.05	1.04	1.13	1.07	1.19	1.24	0.89	1.08
DRB1_40	1.17	1.17	1.34	0.93	1.03	1.21	1.10	0.96	1.09	1.28	0.69	1.16	1.00	1.03
DRB1_47	1.09	1.04	1.08	1.06	1.07	1.40	1.07	1.09	1.16	1.01	0.99	0.88	1.01	1.09
DRB1_57	1.08	1.06	1.13	1.04	1.06	1.17	1.06	1.14	1.15	1.05	1.07	0.85	1.09	1.12
DRB1_58	1.07	1.13	1.02	1.00	1.03	1.46	1.05	1.04	1.11	1.10	1.00	1.04	1.04	1.07
DRB1_60	1.07	1.07	1.16	1.04	1.09	1.09	1.06	1.13	1.16	1.03	1.05	0.89	0.99	1.12
DRB1_67	1.09	1.05	1.12	1.05	1.04	1.32	1.07	1.11	1.11	1.08	1.04	1.04	1.08	1.11
DRB1_70	1.07	1.04	1.13	1.04	1.04	1.08	1.05	1.08	1.12	1.07	1.01	1.14	1.07	1.09
DRB1_71	1.09	1.05	1.07	1.05	1.06	1.19	1.07	1.21	1.10	1.10	1.10	0.99	0.88	1.17
DRB1_73	1.08	1.05	1.14	1.07	1.07	1.19	1.08	1.10	1.10	1.07	1.10	0.71	1.13	1.09
DRB1_74	1.10	1.04	1.14	1.07	1.08	1.11	1.08	1.17	1.14	1.08	1.17	0.88	1.08	1.15
DRB1_77	1.10	1.07	1.03	1.08	1.07	1.04	1.09	1.12	1.07	1.13	1.10	0.66	0.61	1.10
DRB1_78	1.08	1.09	1.15	1.05	1.17	1.09	1.07	1.15	1.15	1.09	1.01	1.20	1.11	1.14
DRB1_85	1.02	1.04	1.15	1.01	0.94	1.32	1.03	1.07	1.15	0.85	1.61	1.73	0.85	1.08
DRB1_86	1.09	0.99	1.03	1.05	1.07	1.36	1.07	1.12	1.11	1.16	0.99	1.15	1.23	1.12
DRB1_9	1.09	1.10	1.15	1.07	1.08	1.12	1.08	1.15	1.11	1.03	1.15	1.14	1.02	1.13

Factors that were still significant after multiple-testing corrections are highlighted.

dx = Donor was race x; rx = recipient was race x.

Races: w = White, b = Black, h = Hispanic, a = Asian, All = all donor/recipient race combinations

Among the ‘significant’ results, the distribution of HRs for graft failure are:

%ile	HR
0	0.46
0.05	1.04
0.1	1.04
0.25	1.06
0.5	1.08
0.75	1.11
0.9	1.14
0.95	1.16
1	1.69

Out of the 398 ‘significant’ results, only 1 (B_12 for db, rw) had a HR for graft failure < 1.

Table S4. Interaction between living donors and deceased donors versus number of HLA amino acid mismatches[#]

	# Amino Acid MM (locus)	HR per additional MM	95% CI	p
All Donors	A	1.0041	1.002 - 1.0061	0.0001
	B	1.0070	0.9969 - 1.0173	0.1764
	B (spline at 5)	0.9925	0.9818 - 1.0033	0.1715
	DRB1	1.0088	1.0064 - 1.0112	<.0001
Deceased	A	1.0038	1.0014 - 1.0062	0.0019
	B	1.0125	1.0003 - 1.0249	0.046
	B (spline at 5)	0.9866	0.9739 - 0.9995	0.0416
	DRB1	1.0083	1.0055 - 1.0111	<.0001
Living, UnRelated	A	1.0043	0.9963 - 1.0123	0.2954
	B	0.9835	0.9275 - 1.0429	0.5787
	B (spline at 5)	1.0188	0.9582 - 1.0832	0.5523
	DRB1	1.0087	0.9991 - 1.0183	0.0776
Living, Related	A	1.0044	0.9993 - 1.0095	0.0886
	B	0.9999	0.9783 - 1.0219	0.9899
	B (spline at 5)	1.0006	0.9768 - 1.0251	0.9589
	DRB1	1.008085	1.002 - 1.0142	0.0089

[#] This table shows slopes of the adjusted hazard ratios of graft failure per additional amino acid mismatch. Model is based on 10 imputations. Zero HLA-A, B, DRB1 antigen (first field, classically two-digit) mismatch transplants were excluded from these analyses. Separate models were used for all donors, deceased, living unrelated, and living related. Each model adjusted for patient risk factors, donor risk factors, and HLA antigen mismatches. Table shows results of models of all races combined.

Table S5. Effect of HLA amino acid mismatches on graft failure in separate post-transplantation intervals

	Number of Amino Acid MM (locus)	HR (per 1 MM increase)	95% CI	p
Overall	A	1.0041	1.002 - 1.0061	0.0001
	B	1.0070	0.9969 - 1.0173	0.1764
	B (spline at 5)	0.9925	0.9818 - 1.0033	0.1715
	DRB1	1.0088	1.0064 - 1.0112	<.0001
< 1 yr	A	1.0050	1.0007 - 1.0092	0.0218
	B	1.0233	1.0008 - 1.0462	0.0426
	B (spline at 5)	0.9806	0.9578 - 1.004	0.1046
	DRB1	1.0173	1.0123 - 1.0223	<.0001
1+ yrs	A	1.0037	1.0014 - 1.0061	0.002
	B	1.0028	0.9915 - 1.0142	0.6312
	B (spline at 5)	0.9954	0.9834 - 1.0076	0.4612
	DRB1	1.0061	1.0033 - 1.0089	<.0001

The model parameter results shown in this table are represented in Figure 2 of the main manuscript. Table shows results of models based on 10 imputations for all races combined.

This table shows slopes of the adjusted hazard ratios of graft failure per additional amino acid mismatch. Zero HLA-A, B, DRB1 antigen (first field, classically two-digit) mismatch transplants were excluded from these analyses. The spline factor to model a change in slope at 5 HLA B amino acid mismatches is coded as the number of HLA B amino acid mismatches minus five, among transplants with at least five amino acid mismatches. Separate models were used for overall, graft survival < 1 year after transplant, and graft survival conditioned on having survived at least one year after

transplant. Each model is adjusted for patient risk factors, donor risk factors, and HLA antigen mismatches.

Table S6. Adjusted hazards ratio of graft failure using a stepwise model of individual HLA-A, -B, -DRB1 amino acid mismatched site variables derived from 10 separate HLA imputations

Deceased Donors				Living Donors			
Amino Acid Position	**Mean % of MM**	**Contact site**	**# of imputations where stepwise models resulted in this factor**	**Amino Acid Position**	**Mean % of MM**	**Contact Site**	**# of imputations where stepwise models resulted in this factor**
HLA-A				**HLA-A**			
56	0.10		5	77	0.41	Pocket F	4
116	0.42	Pocket E, F	10	150	0.15	TCR	5
144	0.29		7	158	0.15	TCR	6
150	0.15	TCR	7				
HLA-B				**HLA-B**			
67	0.72	Pocket B	9	63	0.33	Pocket A	7
147	0.08	Pocket E, F	7	147	0.08	Pocket E, F	8
158*	0.06	TCR	10				
HLA-DRB1				**HLA-DRB1**			
10	0.32		10	13	0.74	Pocket 6	10
26	0.31	P	10	32	0.26		6
28	0.29	Pocket 7	10	71	0.61	Pocket 4	5
74	0.48	Pocket 4	10	74	0.48	Pocket 4	5

The total number of amino acid MM variables evaluated in this analysis includes 30 variable sites for HLA-A, 33 for HLA-B, and 23 for HLA-DRB1. The list of these variable sites is as follow:

A locus: A_9 A_12 A_43 A_56 A_62 A_63 A_65 A_66 A_70 A_73 A_74 A_77 A_79 A_95 A_97 A_102 A_109 A_114 A_116 A_142 A_144 A_149 A_150 A_151 A_152 A_156 A_158 A_161 A_163 A_166

B locus: B_4 B_9 B_11 B_12 B_30 B_41 B_45 B_62 B_63 B_67 B_69 B_76 B_77 B_80 B_94 B_95 B_99 B_103 B_109 B_113 B_114 B_116 B_131 B_143 B_145 B_147 B_152 B_156 B_158 B_163 B_167 B_171 B_178

DR locus: DR_9 DR_10 DRB1_13 DRB1_14 DRB1_16 DR_26 DRB1_28 DR_31 DRB1_32 DRB1_33 DRB1_37 DR_38 DR_40 DR_47 DR_57 DR_58 DR_67 DR_70 DRB1_71 DRB1_74 DRB1_77 DR_85 DR_86.

Correlations of HLA amino acid mismatches for HLA-DRB1 was evaluated using a Pearson correlation between the probabilities of no mismatch at one position compared with the probability of no mismatch at another position using a sample size of 23,095 donor-recipient pairs from the cohort examined in this study. The observed highly correlated amino acid MM sites for each HLA- DRB1 locus is due to the well-known complex linkage disequilibrium patterns of the polymorphic amino acids in the binding site of HLA class I and class II molecules[11].

The R2 values of HLA-A amino acid MM sites that are highly correlated with graft failure are shown in this Table include:

Position 56 correlates with **17** (0.71)

Position 77 (pocket F) correlates with **76** (TCR) [0.76], **166** (TCR) [0.56], and **167** (pocket A) [0.56]

Position 116 (pocket E, F) correlates with **66** (pocket A, B) [0.77], **95** (pocket F) [0.60], **114** [0.72]

Position 144 correlates with **151**(TCR) [0.73]

Position 150 (TCR) correlates with **67** (pocket B) [1.00], **158** (TCR) [1.00], **163** (TCR) [0.53], **166** (TCR) [0.50], and **167** (pocket A) [0.50]

Position 158 (TCR) correlates with **67** (pocket B) [1.00], **158** (TCR) [1.00], **163** (TCR) [0.53], **166** (TCR) [0.50], and **167** (pocket A) [0.50]

R2 values of highly correlated HLA-B amino acid MM sites shown in this table include:

Position 67 correlates with **63** (pocket A) [0.44]; **70** (pocket B, C) [0.54], and **97** (pocket C, E) [0.50]

Position 147 (pocket F) correlates with **143** (pocket F) [1.00]

R2 values of highly correlated HLA-DRB1 amino acid MM sites shown in this table include:

Position 10 correlates with **9** (pocket 6, 9) [0.52], **11** (pocket 6) [0.57], **13** (pocket 6) [0.51], and **32** [0.50]

Position 13 (pocket 6) correlates with **9** (pocket 6, 9) [0.52], 1**0** [0.51], and **11** (pocket 6) [0.85]

Position 28 (pocket 7) correlates with **30** (peptide-contact site) [0.75]

Position 60 (TCR contact) correlates with **14** (peptide contact) [0.66), **25** [0.66], **57** (pocket 9) [0.86], **74** (pocket 4) [0.51], and **78** (peptide-binding) [0.74]

Position 74 (pocket 4) correlates with **14** (peptide contact) [0.51], **25 [**0.51], **57** (pocket 9) [0.56], **60** TCR contact() [0.51], **73** (TCR contact) [0.63], and **78** (peptide contact) [0.53]

Table S7. Adjusted hazards ratio of graft failure using a stepwise model of expected HLA-A, -B, -DRB1 amino acid mismatched site variables[‖]

	Mean % of MM	Contact Site	HR	P		Mean % of MM	Contact Site	HR	P
Deceased Donor					**Living Donor**				
HLA-A					**HLA-A**				
17	0.06		1.04	0.0144	62	0.72	TCR	1.034	0.0491
116	0.42	Pocket E, F	1.02	0.0304	158	0.15	TCR	1.087	<.0001
144	0.29		1.022	0.0142					
150	0.15	TCR	1.033	0.0015					
					HLA-B				
					63	0.33	Pocket A	1.038	0.0149
					143	0.08	Pocket F	1.077	0.0065
					177	0.25		0.966	0.0407
HLA-DRB1					**HLA-DRB1**				
10	0.32		1.044	<.0001	13	0.74	Pocket 6	1.058	0.0043
26	0.31	Peptide	1.018	0.0705	14	0.13	Peptide	1.052	0.0316
28	0.29	Pocket 7	1.032	0.0031	32	0.26		1.043	0.0146
60	0.23	TCR	0.977	0.0487	71	0.61	Pocket 4	1.065	0.0006
74	0.48	Pocket 4	1.046	<.0001	73	0.22	TCR	0.936	0.0021
					74	0.48	Pocket 4	1.062	0.0027

[‖] Due to the well-known complex linkage disequilibrium patterns of polymorphic amino acids in the peptide binding site of HLA class I and class II molecules, the inclusion of one over the other of two highly correlated MM sites can be influenced by which variables is retained in the model during prior steps in the selection process.

Correlations of HLA amino acid mismatches for HLA-DRB1 was evaluated using a Pearson correlation between the probabilities of no mismatch at one position compared with the probability of no mismatch at another position using a

sample size of 23,095 donor-recipient pairs from the cohort examined in this study. The observed highly correlated amino acid MM sites for each HLA- DRB1 locus is due to the well-known complex linkage disequilibrium patterns of the polymorphic amino acids in the binding site of HLA class I and class II molecules[11].

R^2 values of each two highly correlated HLA-A amino acid MM sites shown in this table are listed below:

Position 17 correlates with **56** (0.71)

Position 62 (TCR contact site) correlates with**. 9** (pocket B, C) [0.50], **66** (pocket A, B) [0.52], **95** (pocket F) [0.65], and **97** (pocket C, E) [0.64]

Position 116 (pocket E, F) correlates with **66** (pocket A, B) [0.77], **95** (pocket F) [0.60], **114** [0.72]

Position 144 correlates with **151**(TCR) [0.73]

Position 150 (TCR) correlates with **67** (pocket B) [1.00], **158** (TCR) [1.00], **163** (TCR) [0.53], **166** (TCR) [0.50], and **167** (pocket A) [0.50]

Position 158 (TCR) correlates with **67** (pocket B) [1.00], **158** (TCR) [1.00], **163** (TCR) [0.53], **166** (TCR) [0.50], and **167** (pocket A) [0.50]

R^2 values of each two highly correlated HLA-B amino acid MM sites shown in this table include:

Position 63 (pocket A) correlates with **67** (pocket B)[0.44]

Position 143 (pocket F) correlates with **147** (pocket F) [1.00]

Position 177 correlates with **131** [0.93], and **180** [1.00]

R^2 values of each two highly correlated HLA-DRB1 amino acid MM sites shown in this table include:

Position 10 correlates with **9** (pocket 6, 9) [0.52], **11** (pocket 6) [0.57], **13** (pocket 6) [0.51], and **32** [0.50]

Position 13 (pocket 6) correlates with **9** (pocket 6, 9) [0.52], 1**0** [0.51], and **11** (pocket 6) [0.85]

Position 14 (peptide contact) correlates with **25** [1.00], **57** (pocket 9) [0.57], **60** (TCR contact) [0.66], **73** (TCR contact) [0.56], **74** (pocket 4) [0.51], and **78** (peptide-contact) [0.89]