Impact of HLA class I and class II high resolution matching on outcomes of unrelated donor bone marrow transplantation: HLA-C mismatching is associated with a strong adverse effect on transplant outcome

Neal Flomenberg, Lee Ann Baxter-Lowe, Dennis Confer, Marcelo Fernandez-Vina, Alexandra Filipovich, Mary Horowitz, Carolyn Hurley, Craig Kollman, Claudio Anasetti, Harriet Noreen, Ann Begovich, William Hildebrand, Effie Petersdorf, Barbara Schmeckpeper, Michelle Setterholm, Elizabeth Trachtenberg, Thomas Williams, Edmond Yunis, and Daniel Weisdorf

From the Department of Medicine and Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA; Department of Surgery, University of California, San Francisco, CA; National Marrow Donor Program, Minneapolis, MN; Department of Oncology, Georgetown University, Washington, DC; Children’s Hospital Medical Center, Cincinnati, OH; International Bone Marrow Transplant Registry, Milwaukee, WI; Georgetown University Medical Center, Washington, DC; Jaeb Center for Health Research, Tampa, FL; H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL; Immunology/Histocompatibility Laboratory, Fairview-University Medical Center, Minneapolis, MN; Celera Diagnostics, Alameda, CA; University of Oklahoma Health Sciences Center, Oklahoma City, OK; Fred Hutchinson Cancer Research Center, Seattle, WA; Summercrest Consulting, Columbia, MD; Children’s Hospital Oakland Research Institute, Oakland, CA; Department of Pathology, University of New Mexico, Albuquerque, NM; Department of Cancer Immunology & AIDS, Dana Farber Cancer Institute and Department of Pathology, Harvard Medical School, Boston, MA; University of Minnesota, Minneapolis, MN

Supported by funding from the National Marrow Donor Program, the Health Resources and Services Administration #240-97-0036 and the Office of Naval Research #N00014-93-1-0658, N00014-95-1-0055 and N00014-96-2-0016. The views expressed in this article do not reflect the official policy or position of the Department of the Navy, the Department of Defense or the U.S. government.


Reprints: Neal Flomenberg, M.D., Thomas Jefferson University, 125 S. Ninth St., Suite 801 Sheridan, Philadelphia, PA 19107; e-mail: Neal.Flomenberg@mail.tju.edu.
Abstract
Outcome of unrelated donor marrow transplantation is influenced by donor-recipient matching for Human Leukocyte Antigens (HLA). Prior studies assessing the effects of mismatches at specific HLA loci have yielded conflicting results. The importance of high-resolution matching for all HLA loci has also not been established. We therefore examined the effects of HLA matching (low and/or high resolution) on engraftment, graft-versus-host disease (GVHD), and mortality in 1874 donor-recipient pairs retrospectively typed at high resolution for HLA-A, B, C, DRB1, DQ and DP. Mismatches at HLA-A, B, C and DRB1 each had similar adverse effects on mortality. Only HLA-A mismatches demonstrated significant adverse effects on GVHD. These adverse effects on outcome were more evident in transplants with low-resolution versus only high-resolution mismatches. Mismatches for HLA-DQ or DP did not significantly affect outcome. When high-resolution mismatches at HLA-A, B, C and DRB1 were considered together, adverse effects on survival and GVHD were observed. We therefore conclude that matching for HLA-C should be incorporated into algorithms for unrelated donor selection. High-resolution mismatches at HLA-A, B, C and DRB1 adversely impact outcome, but less so than low-resolution mismatches. When clinical circumstances allow, high-resolution class I typing may help optimize donor selection and improve outcome.
Introduction

Allogeneic hematopoietic stem cell transplantation can potentially cure a variety of lymphohematopoietic and congenital metabolic disorders. While transplants between HLA-identical siblings produce the best outcomes, transplants from unrelated donors using marrow, peripheral blood stem cells, or umbilical cord blood, or using aggressively T-cell depleted mismatched related donors can also yield acceptable outcomes. Of these, the use of bone marrow from well-matched, unrelated donors has been, by far, the most commonly applied approach to date.

The use of unrelated donors introduces a number of questions and problems in donor selection, which do not occur in the context of transplantation from an HLA-identical sibling donor. Ideally, one would hope to identify unrelated donors who are genotypically identical to the patients at all HLA loci, analogous to transplantation between HLA matched siblings. When a “perfectly” matched donor is not available, it remains uncertain whether some mismatches will be more forgiving than others. Controversies remain as to whether mismatches at some loci have more profound clinical consequences than those at other loci. Virtually every HLA locus has been reported to influence outcome of unrelated donor bone marrow transplantation, with conflicting results as to the relative importance of various class I and class II loci. Many of these studies were limited in size, making definitive conclusions difficult.

Additionally, many of the earlier studies relied on serologic typing approaches for HLA class I loci. While serologically based HLA typing is accurate in most circumstances, there are some limitations associated with this technique. In particular, discrimination among certain closely related HLA alleles is beyond the resolution of serologic typing. However, such polymorphisms can be detected by alloreactive T cells, either in the laboratory or clinically in association with GVHD or graft rejection in transplant recipients. It remains uncertain whether mismatches which can only be detected using high-resolution (allele level) nucleic acid techniques are more permissive of clinical success than those mismatches which can be detected using serology or comparable low-resolution (antigen level) DNA based typing approaches.

Resolving these questions has important implications with regard to the number of patients for whom acceptably matched unrelated donors can be identified and the search algorithms used for donor selection. To address these issues, we have performed retrospective high-resolution typing for HLA-A, B, C, DRB1, DQA1, DQB1, DPA1 and DPB1 on over 1800 donor-recipient transplant pairs. The effects of HLA mismatching at high and low resolution on transplant outcome are presented, and the implications of these findings for donor selection are discussed.

Materials and methods

Patient population

Between 1988 and 1996, amongst all bone marrow transplants performed under the auspices of the National Marrow Donor Program (NMDP), retrospective high-resolution typing for HLA-A, B, C, DRB1, DQA1, DQB1, DPA1 and DPB1 was performed for 1874 donor-recipient pairs. This subset of transplants was reflective of the overall case mix with the exceptions of the need for availability of pre-transplant donor and recipient samples for retrospective high-resolution typing and the deliberate overrepresentation of CML cases (45.8% in this analysis vs. 30.1% of controls) to allow a separate outcome study to be performed for that disease (in preparation). Consequently, a reciprocal modest decrease was observed in patients with AML (15.5% vs. 21.0%) and ALL (16.0% vs. 23.6%) amongst the high-resolution typed patients. This likely also accounts for their slightly higher median age (30 vs.
26 years) and percent with Karnofsky scores ≥ 90 (76% vs. 70%). Approximately 9% had nonmalignant disorders, including severe aplastic anemia, Fanconi anemia, metabolic disorders and immune deficiency states. The high-resolution typed group had a higher proportion of cases from the middle years of the study (1992–1994).

**HLA typing**

Retrospective high-resolution molecular typing of class II alleles (DRB1, DQA1, DQB1, DPA1 and DPB1) was performed mainly by sequence specific oligonucleotide probe (SSOP) methods with approximately 64% of the samples tested in duplicate. Of these, 23% were tested by sequence based typing (SBT) methods. All HLA-A, B and C alleles were identified by SBT with duplicate typing performed independently using SSOP methodologies. Duplicate results were compared and discrepancies resolved.

**HLA matching**

Two levels of DNA-based HLA matching were considered in the analyses of clinical outcomes. Donors and recipients were considered high-resolution (allele level) matched for a given locus when their high-resolution typing was identical, indicating that they expressed the identical allele. Low-resolution (serology level or antigen level) DNA matching involved conversion of the DNA-based typing to its lower-level serologic equivalent, usually by collapsing the four-digit typing result back to its first two digits, with the exception of a few HLA-B alleles which were mapped to their corresponding serologic specificities. For HLA-C, low-resolution HLA-C matching was performed by collapsing the four-digit allele back to the first two digits, even though this is less rigorously supported by prior serology than is the case for HLA-A, B and DR.

Though all typing was performed using DNA-based approaches in this study, donors and recipients were considered low-resolution or “serologically equivalent” matched for a given locus when their low-resolution DNA typing assignments for a given locus were identical. Low-resolution matching for a given locus indicates that a donor and recipient express similar gene products (and possibly identical ones) for the locus in question. Conversely, a low-resolution mismatch is one which can be detected using low-resolution typing, while a high-resolution mismatch could only be detected using high-resolution typing techniques. A high-resolution mismatch is used to indicate that the donor and recipient are matched at the serologic or low level of resolution, but differ with regard to the specific allele they express from within that serologic or low-resolution antigenic family.

In the case of HLA-DQ and DP, each of the two protein chains forming the heterodimeric cell surface molecule may contribute to donor-recipient disparities. If there were differences between one HLA-DQA1 type and/or one HLA-DQB1 type, these were assumed to result from a single haplotype and were scored as a single HLA-DQ disparity. Disparity for two HLA-DQA1 and/or two HLA-DQB1 high-resolution types was scored as two HLA-DQ disparities. Comparison of HLA-DP was analogous to HLA-DQ.

**Clinical outcomes**

Evaluation of clinical outcomes was performed using criteria standardized by the NMDP for all its studies. Diagnosis and grading of acute and chronic GVHD utilized standard criteria. Time to engraftment was defined as the first of three consecutive absolute neutrophil counts ≥ 5.0 x 10^8/L. Patients were considered evaluable for engraftment if they survived 21 days after transplant and evaluable for chronic GVHD if they survived at least 80 days.
**Statistical methods**

Comparisons between the subset of cases for which high-resolution typing was performed and the other NMDP-facilitated transplants utilized the Wilcoxon rank sum test for continuous variables (e.g., age) and the likelihood ratio chi-square statistic for categorical variables (e.g., gender).

Cumulative incidences were compared at 100 days for acute GVHD and engraftment, and at two years for chronic GVHD, treating death as a competing risk, and using a Taylor series linear approximation to estimate the variance. Survival rates were estimated up to the date when patient follow-up forms were due at the NMDP, calculated by the method of Kaplan and Meier, and compared using the log-rank statistic.

Logistic regression was used for multivariate analysis of neutrophil engraftment and the proportional hazards model was used for the other outcomes. Each model included disease/stage as a covariate regardless of significance. For HLA-A, B, C and DRB1, donor-recipient matching was considered in three categories: high-resolution match, low-resolution match/high-resolution mismatch, and low-resolution mismatch. Two indicator variables (one for each type of mismatch) were included in the regression model for each of these four loci. A single indicator was included in the model for each of the other HLA loci (DQ and DP) without distinguishing between allele-level and low-resolution mismatches.

Other factors were included in the multivariate models if they demonstrated a statistically significant (Wald chi-square p < 0.05) association with outcome. Factors considered were transplant center; T-cell depletion; cell dose (T-cell replete cases only); recipient and donor age, sex, CMV serology, body mass index and race; interval from diagnosis to transplant; and year of transplant. The interval from diagnosis to transplant was modeled separately for each disease group. Due to non-linear effects, the continuous variables of recipient age and interval from diagnosis to transplant were divided into discrete categories.

Additional regression models were run replacing the HLA indicator variables with two continuous variables counting the total number of low-resolution and high-resolution mismatches, respectively. HLA-DQ and DP were ignored in these models because they had demonstrated no statistically significant effect on transplant outcome. Because of the large number of statistical comparisons performed in this study, only associations with p < 0.01 were considered as statistically significant.

**Results**

**Types of mismatches at HLA-A, B, C and DRB1 present in the study population**

The number and types of mismatches detected by molecular typing in the 1874 cases analyzed are illustrated in Table 1. Table 1A summarizes the types of mismatches observed at each locus. As illustrated, approximately half the mismatches at HLA-A and B were detectable by low-resolution typing, and half required high-resolution typing for detection. Nearly twice as many HLA-C locus disparities were identified compared to HLA-A and B. Over 80% of HLA-C mismatches were detectable with low-resolution typing. Most HLA-DR mismatches could only be detected using high-resolution typing for DRB1, while DQ mismatches were evenly split as to whether they required high-resolution typing for detection or were detectable using low-resolution techniques.
### A. Mismatches by locus

<table>
<thead>
<tr>
<th>Locus</th>
<th>Donor-recipient pairs (n = 1874)</th>
<th>Alleles (n = 3748)</th>
<th>Mismatches detectable at low resolution</th>
<th>Mismatches detectable at high resolution only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total (%) mismatched donor-recipient pairs at this locus</td>
<td>Total (%) mismatched alleles at this locus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLA-A</td>
<td>374 (20%)</td>
<td>386 (10%)</td>
<td>219 (57%)</td>
<td>167 (43%)</td>
</tr>
<tr>
<td>HLA-B</td>
<td>477 (25%)</td>
<td>514 (14%)</td>
<td>209 (41%)</td>
<td>305 (59%)</td>
</tr>
<tr>
<td>HLA-C</td>
<td>749 (40%)</td>
<td>851 (23%)</td>
<td>734 (86%)</td>
<td>117 (14%)</td>
</tr>
<tr>
<td>HLA-DRB1</td>
<td>311 (17%)</td>
<td>342 (9%)</td>
<td>52 (15%)</td>
<td>290 (85%)</td>
</tr>
<tr>
<td>HLA-DQ</td>
<td>415 (22%)</td>
<td>449 (12%)</td>
<td>219 (49%)</td>
<td>230 (51%)</td>
</tr>
<tr>
<td>HLA-DP</td>
<td>1648 (88%)</td>
<td>2255 (60%)</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

### B. Mismatch characteristics

<table>
<thead>
<tr>
<th>Number of donor-recipient pairs</th>
<th>Number of donor-recipient pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matched for all 8 loci</td>
<td>108 (6%)</td>
</tr>
<tr>
<td>High-resolution mismatch at HLA-A,B,C and/or DRB1</td>
<td>631 (34%)</td>
</tr>
<tr>
<td>Low-resolution mismatch at HLA-A,B,C and/or DRB1</td>
<td>452 (24%)</td>
</tr>
<tr>
<td>DQ and/or DP mismatch only</td>
<td>683 (36%)</td>
</tr>
<tr>
<td>Total</td>
<td>1874 (100%)</td>
</tr>
<tr>
<td>Total mismatched alleles at HLA-A,B,C, and/or DRB1</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>791 (42%)</td>
</tr>
<tr>
<td>1</td>
<td>469 (25%)</td>
</tr>
<tr>
<td>2</td>
<td>351 (19%)</td>
</tr>
<tr>
<td>3+</td>
<td>263 (14%)</td>
</tr>
<tr>
<td>Total</td>
<td>1874 (100%)</td>
</tr>
</tbody>
</table>

Table 1A. Shown are the total number of mismatches detected in 3748 alleles typed at each locus in the 1874 donor-recipient pairs. The absolute number and percentage (in parentheses) of mismatches detectable at either low or high resolution only is shown.

Table 1B. Shown are the types of mismatches detected between donor and recipient in the 1874 pairs. “High-resolution mismatch at HLA-A,B,C and/or DRB1” indicates the presence of one or more mismatches at one of these loci detectable with high-resolution typing only. “Low-resolution mismatch at HLA-A,B,C and/or DRB1” indicates the presence of one or more mismatches at one of these loci detectable at low resolution with or without additional high-resolution mismatches. The 8 loci typed include HLA-A, B, C, DRB1, DQA1, DQB1, DPA1, and DPB1.
The match/mismatch characteristics of the 1874 donor-recipient pairs are summarized in Table 1B. Only 6% of the donor-recipient pairs were matched for all 8 of the loci studied. Thirty six percent of the pairs were matched at high resolution for HLA-A, B, C and DRB1 but had a mismatch at DQ and/or DP. Thus between these two groups, 42% of the pairs were matched at high resolution for HLA-A, B, C and DRB1. Of the remaining 58% of the study pairs, 34% were matched for HLA-A, B, C and DRB1 at low resolution but had one or more mismatches at high resolution. Twenty four percent had at least one HLA-A, B, C or DRB1 mismatch which could be detected at low resolution. Within the study population, 25% of pairs exhibited one HLA-A, B, C or DRB1 mismatch, 19% exhibited two mismatches, and 14% exhibited three or more mismatches at these loci. Hurley et al, manuscript in preparation

**Impact of mismatching at individual HLA loci on transplant-related outcomes**

Table 2A shows the results of multivariate analyses of the impact of HLA mismatching on transplant-related outcomes. The cumulative incidence of engraftment in the overall study population was 95 ± 1%. Even among patients with HLA mismatches, engraftment rates remained high. HLA-C mismatching showed the strongest association with graft failure (OR of engraftment 0.54, p = 0.02), though this did not achieve the statistical significance threshold defined for this study. Multiple HLA-C mismatches were not associated with a significantly higher risk of non-engraftment (data not shown).
Table 2. Impact of HLA mismatching at specific loci on transplant-related outcomes.

### A. Associations between HLA mismatch (high- or low-resolution) at specific loci and outcome after unrelated donor BMT

<table>
<thead>
<tr>
<th>Mismatched locus</th>
<th>N</th>
<th>Engraftment</th>
<th>Grade III-IV acute GVHD</th>
<th>Chronic GVHD</th>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>OR</td>
<td>95% CI</td>
<td>P</td>
<td>RR</td>
</tr>
<tr>
<td>HLA-A</td>
<td>374</td>
<td>0.68</td>
<td>(0.39, 1.22)</td>
<td>0.20</td>
<td>1.41</td>
</tr>
<tr>
<td>HLA-B</td>
<td>477</td>
<td>1.07</td>
<td>(0.61, 1.88)</td>
<td>0.82</td>
<td>1.24</td>
</tr>
<tr>
<td>HLA-C</td>
<td>749</td>
<td>0.54</td>
<td>(0.33, 0.89)</td>
<td>0.02</td>
<td>1.19</td>
</tr>
<tr>
<td>HLA-DRB1</td>
<td>311</td>
<td>1.07</td>
<td>(0.57, 2.02)</td>
<td>0.83</td>
<td>1.26</td>
</tr>
<tr>
<td>HLA-DQ</td>
<td>415</td>
<td>0.67</td>
<td>(0.39, 1.14)</td>
<td>0.14</td>
<td>1.03</td>
</tr>
<tr>
<td>HLA-DP</td>
<td>1648</td>
<td>0.69</td>
<td>(0.38, 1.25)</td>
<td>0.22</td>
<td>1.19</td>
</tr>
</tbody>
</table>

### B. Associations between high- versus low-resolution HLA mismatch at specific loci and outcome after unrelated donor BMT

<table>
<thead>
<tr>
<th>Locus</th>
<th>N</th>
<th>Engraftment</th>
<th>Grade III-IV acute GVHD</th>
<th>Chronic GVHD</th>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>OR</td>
<td>95% CI</td>
<td>P</td>
<td>RR</td>
</tr>
<tr>
<td>HLA-A match</td>
<td>1500</td>
<td>1.00</td>
<td>—</td>
<td>—</td>
<td>1.00</td>
</tr>
<tr>
<td>Hi-res mismatch</td>
<td>157</td>
<td>0.95</td>
<td>(0.40, 2.26)</td>
<td>0.91</td>
<td>1.31</td>
</tr>
<tr>
<td>Low-res mismatch</td>
<td>217</td>
<td>0.51</td>
<td>(0.24, 1.08)</td>
<td>0.08</td>
<td>1.51</td>
</tr>
<tr>
<td>HLA-B match</td>
<td>1397</td>
<td>1.00</td>
<td>—</td>
<td>—</td>
<td>1.00</td>
</tr>
<tr>
<td>Hi-res mismatch</td>
<td>270</td>
<td>1.66</td>
<td>(0.75, 3.69)</td>
<td>0.22</td>
<td>1.13</td>
</tr>
<tr>
<td>Low-res mismatch</td>
<td>207</td>
<td>0.72</td>
<td>(0.34, 1.49)</td>
<td>0.37</td>
<td>1.35</td>
</tr>
<tr>
<td>HLA-C match</td>
<td>1125</td>
<td>1.00</td>
<td>—</td>
<td>—</td>
<td>1.00</td>
</tr>
<tr>
<td>Hi-res mismatch</td>
<td>83</td>
<td>0.80</td>
<td>(0.29, 2.24)</td>
<td>0.68</td>
<td>0.83</td>
</tr>
<tr>
<td>Low-res mismatch</td>
<td>666</td>
<td>0.54</td>
<td>(0.31, 0.93)</td>
<td>0.03</td>
<td>1.24</td>
</tr>
<tr>
<td>HLA-DRB1 match</td>
<td>1563</td>
<td>1.00</td>
<td>—</td>
<td>—</td>
<td>1.00</td>
</tr>
<tr>
<td>Hi-res mismatch</td>
<td>260</td>
<td>0.77</td>
<td>(0.39, 1.50)</td>
<td>0.44</td>
<td>1.29</td>
</tr>
<tr>
<td>Low-res mismatch</td>
<td>51</td>
<td>1.61</td>
<td>(0.28, 9.26)</td>
<td>0.59</td>
<td>1.42</td>
</tr>
</tbody>
</table>

Shown are results of multivariate analysis including odds ratio (OR) for engraftment and relative risk (RR) for GVHD and mortality, 95% confidence intervals (CI), and p-values associated with mismatching at each locus.
Table 2A. HLA-A mismatching showed significant associations (shaded) with acute GVHD, chronic GVHD, and mortality. HLA-B, C and DR showed significant associations with mortality. Mismatch at HLA-C showed a trend to association with poorer engraftment, though not reaching the significance threshold defined for this study. Mismatched HLA-B, C, DR all showed trends to more frequent acute GVHD.

Table 2B. Low, but not high-resolution mismatching for HLA-A adversely affected acute GVHD, chronic GVHD, and mortality, while low-resolution mismatch for HLA-B, C and DR showed significant effects on mortality. Similar adverse relative risks and trends (0.01 < p < 0.05) were noted also with high-resolution mismatch at HLA-A for acute GVHD and mortality, HLA-DR for acute GVHD, chronic GVHD and mortality. Low-resolution, but not high-resolution mismatching demonstrated trends toward adverse effects for HLA-B on acute GVHD and for HLA-C on engraftment and acute GVHD.
Mismatching for HLA-A was associated with a significantly increased risk of grades III-IV acute GVHD (RR = 1.41, p = 0.005). Mismatching for HLA-B, C, DR and DP were each associated with relative risks for grades III-IV acute GVHD of around 1.2, but these did not reach independent statistical significance (p = 0.03–0.06). HLA-DQ mismatched recipients demonstrated no increased risk of developing grades III-IV acute GVHD (RR = 1.03, p = 0.76).

HLA-A mismatching was also associated with a significantly higher incidence of chronic GVHD (RR = 1.35, p = 0.006). HLA-DR and DP mismatches were associated with somewhat higher relative risks of developing chronic GVHD, but these associations were not statistically significant. Mismatching for HLA-B, C and DQ was not associated with increased risk of chronic GVHD.

The final and most important outcome variable analyzed was mortality. As illustrated in Table 2A, mismatching for HLA-A, B, C, or DR were each independently associated with significantly higher risks of mortality. In contrast, mismatching for HLA-DQ and DP did not appear to exert any significant effect on survival.

Impact of high- and low-resolution mismatches at individual loci on transplant-related outcomes

Mismatches at HLA-A, B, C and DR were all associated with at least one significant effect on a major clinical outcome. Therefore, we next examined these associations after subdividing the mismatches into those detectable at low resolution versus those requiring high-resolution typing for detection. In this analysis, HLA-DQ and DP mismatches were not included, as these loci showed no significant independent association with any of the transplant outcomes studied.

Clinical outcomes were examined in the subsets of patients with either high- or low-resolution mismatches at each locus (Table 2B) and compared with the results observed when these two subgroups were combined (Table 2A). As shown in Table 2A, mismatching for a particular HLA locus was associated with a statistically significant (p < 0.01) impact on transplant outcome in six cases. Specifically, mismatches at HLA-A, B, C and DR each demonstrated significant associations with mortality, while HLA-A mismatching also demonstrated significant associations with acute and chronic GVHD. As shown in Table 2B, in all six of these instances a similar statistically significant adverse effect was observed in the subset of mismatches detectable with low-resolution typing. Within the limits of the current sample size, there was no instance, either in these six cases or the others analyzed, where high-resolution mismatching showed an independent statistically significant impact on outcome when analyzed for an individual HLA locus. However, if one compares the low- and high-resolution mismatched subsets, the relative risks are similar in four of the six cases noted above (HLA-A and DR on mortality; HLA-A on acute and chronic GVHD).

If one compares the low-resolution mismatched subset (Table 2B) to the composite group of mismatches (i.e., low- and high-resolution mismatches combined [Table 2A]), the relative risks are higher for the low-resolution mismatched subset in five of the six cases noted above and the same in one (impact of HLA-C on mortality). Looking separately at the impact of HLA-A, B, C, or DR mismatching on mortality, the statistical significance of the associations was actually stronger in the low-resolution subsets than in the composite (high and low) groups, despite the fact that the sample sizes were smaller and the degrees of freedom higher.

Thus in each of these six instances where mismatching at one particular HLA locus was associated with an adverse effect on transplant outcome, the risks were most evident in those cases mismatched at low resolution. However, the number of patients within each of these groups is too small to formally prove whether high-resolution and low-resolution mismatches at any given locus truly differ from one another.
with regard to risk, despite the fact that low-resolution mismatches showed statistically significant associations with outcome events while high-resolution mismatches did not.

Impact of class I versus class II high-resolution mismatches on transplant outcome

Within the limits of the current study population, high-resolution HLA mismatches at any single locus did not demonstrate an independent statistically significant adverse effect on unrelated donor BMT outcome. However, when the high-resolution mismatches detectable from DNA-based typing for class I (HLA-A, B and C) were pooled, additional effects on outcome could be demonstrated.

Using HLA-A, B, DR low-resolution matched pairs as our starting group, we assessed the clinical impact of additional mismatches at either HLA class I loci versus mismatches at HLA-DRB1 in these patients. As illustrated in Table 3A, the presence of a single mismatch at class I or a single mismatch for HLA-DRB1 had similar deleterious effects on the incidence of grades III-IV acute GVHD (8–12% more frequent) and mortality (8–12% worse at 5 years). Next, as illustrated in Table 3B, we analyzed cases matched for HLA-A and B at low resolution and DRB1 at high resolution to assess the impact of additional mismatches at HLA-A, B and C versus mismatches for HLA-DQ and DP on transplant outcome. In this group, often clinically described as “6 antigen matched,” additional mismatches for HLA-DQ and DP had no impact on survival while, in contrast, a single additional class I mismatch had a substantial adverse effect (7–8% worse) on survival.

### Table 3. Impact of mismatching on transplant outcome.

**A. Impact of DRB1 versus class I mismatching in HLA-A, B, DR serologically matched pairs**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Grades III-IV acute GVHD</th>
<th>Survival (5 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No high-resolution mismatches</td>
<td>791</td>
<td>30% ± 3%</td>
<td>39% ± 3%</td>
</tr>
<tr>
<td>Single mismatch at class I</td>
<td>317</td>
<td>40% ± 5%</td>
<td>31% ± 5%</td>
</tr>
<tr>
<td>Single high-resolution mismatch at DRB1</td>
<td>77</td>
<td>38% ± 10%</td>
<td>27% ± 10%</td>
</tr>
</tbody>
</table>

**B. Impact of additional mismatching on survival among HLA-A, B low-resolution and DRB1 (“6 antigen”) matched pairs**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Survival (5 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-resolution matched for 8 of 8 loci</td>
<td>108</td>
<td>39% ± 9%</td>
</tr>
<tr>
<td>Mismatch at DQ/DP only</td>
<td>683</td>
<td>39% ± 4%*</td>
</tr>
<tr>
<td>Single class I mismatch</td>
<td>317</td>
<td>31% ± 5%†</td>
</tr>
</tbody>
</table>

Table 3A. Shown are comparisons between mismatch at class I vs. HLA-DRB1 on acute GVHD (p = 0.65) and 5 year survival (p = 0.37). Single mismatches at class I include high-resolution mismatches for HLA-A and B and all mismatches for HLA-C.

Table 3B. 8 locus match includes high-resolution matching for HLA-A, B, C, DRB1, DQA1, DQB1, DPA1, and DPB1. Class I mismatch refers to high-resolution mismatches for HLA-A and B and all mismatches for HLA-C.

* p = 0.93 and † p = 0.08 compared to 8 of 8 matched pairs
In prior reports, HLA-DQ mismatches have been shown to adversely impact transplant outcome.\textsuperscript{25,27} We therefore assessed whether the prior observation might have resulted from the ability of a mismatch for HLA-DQ to aggravate the impact of class I disparities which went undetected in those studies. In patients matched for HLA-DRB1 with a single disparity for HLA-class I, additional disparity for HLA-DQ had no adverse effect on outcome (data not shown).

We then considered the cumulative impact of all allele-level mismatches for HLA-A, B, C and DRB1 on GVHD risk and survival in transplants performed between HLA-A, B low-resolution matched, DRB1 high-resolution matched pairs. As illustrated in Figure 1, there is a statistically significant increase in risk of developing grades III-IV acute GVHD with even a single class I mismatch. Similarly, survival progressively declined as the number of mismatches for HLA-A, B and C increased (Figure 2). In multivariate analysis, the relative risk of grades III-IV acute GVHD increased from 1.0 to 1.53 to 1.78 as the number of HLA-A, B, C and DR mismatches increased from 0 to 1 to 2. Similarly, the risk of mortality increased from 1.0 to 1.32 to 1.53.

![Figure 1: Grades III-IV acute GVHD among HLA-A, B serologic and DRB1 allele-matched pairs by number of class I mismatched loci.](image)

The incidence of grades III-IV acute GVHD was analyzed as a function of the number of class I mismatches detected by high-resolution HLA typing in HLA-A, B low-resolution and DRB1 high-resolution matched donor-recipient pairs. The data presented are adjusted for competing risk factors using the proportional hazards model, rather than univariate analysis. One or more additional mismatches led to more frequent GVHD (p=0.001).
Figure 2. Risk-adjusted survival among HLA-A, B serologic and DRB1 allele-matched pairs by number of mismatched class I loci. Survival after transplant was analyzed as a function of the number of class I mismatches detected by high-resolution HLA typing in HLA-A, B low-resolution and DRB1 high-resolution matched donor-recipient pairs. The data presented are adjusted for competing risk factors using the proportional hazards model, rather than univariate analysis. One or more additional mismatches led to poorer risk-adjusted survival ($p = 0.0003$).

It should be noted that in the above analyses, both high- and low-resolution C locus disparities were included. If one restricts the analysis to pairs who are low-resolution matched for HLA-C as well as the other loci, the trends observed are the same, though they do not reach statistical significance in these smaller groups (data not shown).

Because of the direct implications on donor selection, we constructed additional regression models to compare the impact of low- versus high-resolution mismatches at HLA-A, B, C, or DR on GVHD or survival. Donor-recipient pairs with a single low-resolution mismatch were compared to pairs with a single high-resolution mismatch. Pairs with two high-resolution mismatches were compared to pairs with one high-resolution and one low-resolution mismatch (two total mismatches). Patients with three and four mismatches were compared in similar fashion. In this model, mismatches detectable at low resolution were associated with similar risks of grades III-IV acute GVHD as high-resolution mismatches ($RR = 1.16, p = 0.18$). However, transplants with low-resolution mismatches were associated with significantly worse survival than those with only high-resolution mismatches ($RR = 1.26, p = 0.006$).

**Frequency of high-resolution mismatches in low-resolution matched pairs**

Since matching at HLA-A, B, C and DR were all shown to influence survival after unrelated donor bone marrow transplantation, we assessed the ability of low-resolution HLA-A, B, C and DR typing to predict the outcome of high-resolution typing for these same loci. These results are illustrated in Figure 3. Low-resolution matching for HLA-A, B and DR (6 of 6 match) was associated only with a 56% chance of high-resolution matching for HLA-A, B, C and DR. Among these HLA-A, B, DR low-resolution
matched pairs, 26% had at least one high-resolution mismatch for HLA-A, B, C, or DR, while 12% had two, 4% had three, and 2% had four or more mismatches respectively. Amongst the donor-recipient pairs matched at low resolution for HLA-C as well as for HLA-A, B and DR (8 of 8 match), the frequency of high-resolution matching for all four loci increased to 76%. Nineteen percent of these cases had a single high-resolution mismatch for HLA-A, B, C, or DR and only 5% and 1% of cases exhibited two and three high-resolution mismatches respectively. However, if a single low-resolution mismatch is detected for HLA-A, B, or DR (5 of 6 match), the likelihood of identifying additional occult mismatches after high-resolution typing increases substantially. After high-resolution typing of these 5 of 6 matched pairs, additional mismatches are identified in 76% of pairs. Forty percent had one additional mismatch beyond the original mismatch detected by low-resolution typing (two mismatches in total), 22% had three and 14% had four HLA-A, B, C, or DR mismatches.

Figure 3. Predictive value of HLA A, B, C and DR low-resolution typing for subsequent matching after high-resolution typing. The frequency of unrecognized high-resolution mismatches in pairs selected using low-resolution typing for HLA-A, B, and DR ± HLA-C. “6 of 6” match refers to pairs matched at low resolution for HLA-A, B, and DR (n=1422). “5 of 6” match refers to pairs with a single low-resolution mismatch for HLA-A, B, or DR (n=429). “8 of 8” match refers to pairs matched at low resolution for HLA-A, B, C and DR. (n=1047). The data shows the frequency of pairs with 0, 1, 2, 3 and 4 or more allele-level mismatches in each cohort.

Discussion

We have described a large and diverse analysis of unrelated donor transplant recipients assessing the impact of high-resolution HLA matching for all major class I and class II loci on transplant outcome. The results demonstrate strong negative effects of mismatching for either HLA-A, B, C and DRB1 on survival after unrelated donor BMT. Single mismatches at these loci were associated with significant decrements in survival, and the presence of multiple mismatches was even worse. Low-resolution mismatches appear to have more severe impact on survival than mismatches detectable only with high-
resolution typing techniques, but high-resolution mismatches were also associated with adverse outcomes.

Using rigorous statistical criteria, low- but not high-resolution HLA-DRB1 mismatching was associated with adverse effects on survival. Previous studies reported that HLA-DRB1 mismatching (low- or high-resolution) was associated with worse GVHD and survival.\textsuperscript{24,28} Since that time, HLA-DRB1 matching has been a priority in most transplant programs, driving the broad-based adoption of molecularly-based DRB1 typing. During the last decade, relatively few patients proceeded to transplant with HLA-DR low-resolution mismatched donors. When a high-resolution mismatch for HLA-DRB1 was unavoidable, some transplant centers have attempted to select donors whose DRB1 mismatches had fewer amino acid disparities, conservative amino acid substitutions, or substitutions in portions of the molecules thought to be less functionally significant. Data supporting these assumptions in donor selection have not been reported. However, it is conceivable that such practices have sufficiently skewed the types of DRB1 mismatches occurring in the present study so as to dilute the impact of mismatching at this locus on transplant outcome, partially accounting for the finding that high-resolution HLA-DRB1 mismatching was not clearly associated with any adverse effects on GVHD and survival using the study’s rigorous statistical criteria. Additionally, the small subset of low-resolution DRB1 mismatches had the highest relative risk for mortality. Our data concerning HLA-DRB1 are consistent with earlier observations, despite the lack of an independently significant association, and we encourage continued use of high-resolution HLA-DRB1 matching as a criterion in donor selection.

Analysis of class I allele matching in Japan\textsuperscript{29} identified increased GVHD with HLA-A and HLA-C mismatching, and poorer survival with HLA-A mismatching. HLA-B associations with GVHD and survival were observed in univariate, but not multivariate analyses. HLA class II disparities did not affect outcome. Morishima\textsuperscript{37} subsequently reported that single high-resolution disparities at HLA-A, B, C led to more acute GVHD and graft failure in Japan, while HLA-A and B disparities worsened chronic GVHD and survival. In contrast to Sasazuki’s earlier report,\textsuperscript{29} HLA-DRB1 was also a risk factor for acute GVHD but not other transplant outcomes. These studies did not address the relative importance of low- versus high-resolution mismatches. Smaller studies of Caucasian populations have also raised the question as to whether HLA-C mismatching worsens GVHD and survival.\textsuperscript{22,23}

In contrast, our study demonstrated significant adverse impact for HLA-A, B, or C mismatching on survival. While only HLA-A demonstrated statistically significant effects on the incidence of grades III-IV acute GVHD, HLA-B, C and DR mismatches all showed trends for more frequent acute GVHD (RR = 1.19–1.26; 0.01 < p < 0.05). Moreover, current techniques for scoring GVHD focus only on peak severity and do not capture information about resistance to therapy or the intensity and duration of treatment required. Thus, whether mismatches at HLA loci result in more resistant GVHD and worse survival, despite a similar overall incidence, requires further study. The reasons underlying differences in the consequences of HLA-DRB1 mismatching in Japan and North America are unclear. However, disparities at various HLA loci may differ between the Japanese population and the more heterogeneous NMDP population which includes Caucasians, African Americans, Hispanics, Asians and individuals of mixed ancestry.

Prior North American and Japanese studies have suggested that HLA-C mismatches augment graft failure risks.\textsuperscript{26,37} Petersdorf\textsuperscript{33} reported increased risks of graft failure with low-resolution or multiple high-resolution mismatches for HLA class I loci, though this was primarily in CML patients. We observed only a statistically insignificant trend to greater graft failure risks with HLA-C mismatching, but not with other loci.

Our analysis shows strong adverse effects on survival from mismatches at HLA-A, B, C, or DRB1. The observation concerning HLA-C is particularly important, as this locus is omitted in most matching
algorithms. The present study suggests that HLA-C exerts significant effects on survival, comparable in magnitude to HLA-A, B and DRB1. Strong linkage disequilibrium between HLA-B and C result in their frequent coordinate matching, but it is important that the favorable, independent impact of HLA-C matching on survival be recognized and included in algorithms used for unrelated donor selection.51

The impact of HLA-DQ and DP has not been clearly established from prior reports. Consistent with the report of Sasazuki et al,29 in our analysis mismatches at HLA-DQ and DP had little impact on transplant outcome. This is in contradistinction to other reports which found DQ and DP mismatching to be risk factors for GVHD and other post-transplant outcomes.25,27,31,33 In the present study, HLA-DQ showed RR of 1.03 for grades III-IV acute GVHD and RR of 0.98 for mortality with p-values of 0.76 and 0.80 respectively. To reconcile the present findings concerning HLA-DQ with prior observations,25,27,31,33 we considered the possibility that HLA-DQ disparities may have worsened the clinical consequences of high-resolution class I mismatches30,52 which went undetected as they were not consistently typed for in the earlier studies.25,27,31,33 We therefore assessed outcomes of patients with DQ disparities with or without associated class I disparities. These two groups of patients behaved identically, and thus matching at HLA-DQ was not confirmed as critical to successful outcome of unrelated donor BMT.

Linkage disequilibrium between the DR and DQ loci may have limited the spectrum of DQ mismatches in the study population to more limited mismatches such as DQB1*0301 versus DQB1*0302. It is also conceivable that, analogous to HLA-DRB1, donor selection practices may have similarly skewed the HLA-DQ mismatches in the study to those perceived to be more clinically permissible. However, in contrast to HLA-DRB1, the statistical analysis does not suggest any deleterious effects independently associated with HLA-DQ mismatch.

Prior matching algorithms have often favored class II matching over class I matching when a complete match could not be identified. HLA-DP is particularly problematic because this locus shows little positive genetic linkage disequilibrium with the rest of the major histocompatibility complex. The strong impact of HLA-A, B, C and DRB1 on survival and the lack of any significant impact of HLA-DQ and DP suggest that when mismatching is unavoidable, it may be appropriate to accept a mismatch for HLA-DQ if this will facilitate better matching for HLA-class I loci and DR. The current data do not support expending the resources nor incurring the delays required for HLA-DP typing.

The present study is the first to analyze the relative clinical importance of high- versus low-resolution mismatching for HLA-class I loci for mortality and GVHD. In all six instances where mismatching for HLA-A, B, C, or DR demonstrated a significant effect on clinical outcome (Table 2A, shaded), a statistically significant effect was also noted for the corresponding subset of low-resolution mismatched cases (Table 2B, shaded). Moreover, in the stratified regression analyses performed to compare the relative effects of high- versus low-resolution mismatches, a low-resolution mismatch was associated with significantly worse post-transplant survival. In contrast, we observed similar adverse impact of low- or high-resolution mismatching on acute GVHD. Petersdorf and associates52 have reported that HLA mismatches detectable at low resolution carry a higher risk of graft rejection than mismatches only detectable using high-resolution typing. This is consistent with the concept that low-resolution mismatched donor-recipient pairs are likely to differ at a larger number of immunogenic epitopes than would usually occur in the setting of a high-resolution mismatch. While the above data argue that certain high-resolution mismatches are perhaps more permissive than their low-resolution counterparts, high-resolution mismatches do, when pooled together, still exhibit adverse effects on transplant outcome. In HLA-A, B, DR low-resolution matched pairs, the adverse effects revealed by additional typing information for HLA-A, B and C versus HLA-DRB1 are similar in the two subgroups. Thus the current data would indicate that there is demonstrable clinical benefit to using high-resolution class I typing to guide unrelated donor selection.
It is noteworthy that low-resolution HLA-A, B and DR matching, a commonly used starting point for donor selection, provides little more than a 50% likelihood that a donor will be matched for HLA-A, B, C and DRB1 after high-resolution typing is performed. The addition of HLA-C typing to the algorithm increases the predictive value, in part by eliminating the HLA-C mismatched cases and in part by helping to match for haplotypes rather than alleles.

The current study underscores the importance of recognizing the added information gained from high-resolution typing for HLA-A, B, C and DRB1. Adding more typing information to the process of donor selection should not be used to exclude more patients from the potentially curative benefits of this therapy. Rather it should be used as a means of stratifying patients’ risk and allowing risk-adapted treatment strategies to be based on more complete and precise matching information. Ultimately, new approaches to immune suppression which selectively or preferentially eliminate or permanently inactivate alloreactive T cells while sparing other T cell populations will be required to obviate the adverse effects of HLA mismatching on post-transplant survival.

In summary, the present analysis demonstrates significant adverse clinical effects for HLA-A, B, C and DR mismatch on survival after unrelated donor BMT. Most important, HLA-C should clearly be included in search algorithms. High-resolution class I typing for HLA-A, B and C provides important additional prognostic information regarding transplant outcome. The ability to resolve the DNA sequences of the HLA alleles in even larger numbers of donors and recipients will allow us to address questions such as whether multiple amino acid substitutions will have more significance than fewer substitutions, whether conservative amino acid substitutions will be more permissive clinically, and whether substitutions in some portions of HLA molecules will be better tolerated than others. Answering these questions will allow us to more precisely understand the structural basis of alloreactivity and to hopefully define a subset of HLA disparities with little clinical consequence, which can be clinically tolerated when a perfectly matched donor is unavailable.
References


29. Sasazuki T, Juji T, Morishima Y, Kinukawa N, Kashiwabara H, Inoko H et al. Effect of matching of class I HLA alleles on clinical outcome after transplantation of hematopoietic stem cells from an


53. Santamaria P, Reinsmoen NL, Lindstrom AL, Boyce-Jacino MT, Barbosa JJ, Faras AJ et al. Frequent HLA class I and DP sequence mismatches in serologically (HLA-A, HLA-B, HLA-DR) and molecularly (HLA-DRB1, HLA-DQA1, HLA-DQB1) HLA-identical unrelated bone marrow transplant pairs. [erratum appears in Blood 1994 Jun 15;83(12):3834]. Blood. 1994;83(1):280-287.


Impact of HLA class I and class II high resolution matching on outcomes of unrelated donor bone marrow transplantation: HLA-C mismatching is associated with a strong adverse effect on transplant outcome