Biol Blood Marrow Transplant 25 (2019) 921–931

Biology of Blood and Marrow Transplantation

journal homepage: www.bbmt.org

Direct HLA Genetic Comparisons Identify Highly Matched Unrelated Donor-Recipient Pairs with Improved Transplantation Outcome

Ibrahim Vazirabad[1,†], Saurabh Chhabra[2,†], James Nytes[1], Vatsal Mehra[1], Ravi K. Narra[2], Aniko Szabo[3], James H. Jerkins[2], Binod Dhakal[2], Parameswaran Hari[2], Matthew W. Anderson[1,*]

[1] *Diagnostic Laboratories and Blood Research Institute, BloodCenter of Wisconsin, Milwaukee, Wisconsin*
[2] *Division of Hematology/Oncology, Department of Medicine, Medical College of Wisconsin, Milwaukee, Wisconsin*
[3] *Division of Biostatistics, Institute for Health and Equity, Medical College of Wisconsin, Milwaukee, Wisconsin*

Article history:
Received 16 November 2018
Accepted 3 December 2018

Key Words:
Next-generation sequencing
HLA
Unrelated donor
Allogeneic transplantation

A B S T R A C T

HLA matching by allele-level genotyping is largely based on genetic similarity between a few exons that encode the antigen recognition domain (ARD) of the HLA protein. Next-generation sequencing (NGS) can identify HLA genetic polymorphisms in non-ARD-encoding exons, introns, and untranslated regions, but the impact of these polymorphisms on hematopoietic cell transplantation (HCT) outcome is unclear. We performed NGS-based sequencing of 11 HLA loci on a well-characterized retrospective cohort of 166 unrelated donor-recipient HCT pairs. Genetic differences between HCT pairs were identified and visualized using a novel bioinformatics approach that directly compares phased full-length HLA sequences. Our approach was able to correctly classify HCT pairs without known HLA allele-level mismatches and also to identify a subset of HLA allele-matched HCT pairs with very few to no genetic differences in the sequenced HLA regions. This highly HLA genetically matched unrelated HCT group shows improved overall survival and reduced acute graft-versus-host disease compared with HCT pairs with HLA allele-level mismatches. These results suggest that direct genetic matching of HLA loci may offer an additional means of HCT donor selection beyond traditional HLA allele comparisons and suggests that genetic similarity as defined by HLA sequencing may have a novel role in unrelated HCT donor selection. Finally, our approach can enable larger cohort studies with adequate power to detect differences in other HCT outcomes based on genetic similarity within the HLA loci.

INTRODUCTION

Consistent with their role in immune recognition, HLA genes are the most polymorphic loci in the human genome. Within the HLA genes, the highest degree of genetic polymorphism is found within a few exons that encode the peptide and T cell receptor- binding regions of the HLA proteins, (also termed antigen recognition domains (ARDs) [1]. Exons 2 and 3 encode the peptide- binding site for HLA class I proteins (HLA-A, -B, and -C), whereas exon 2 encodes the peptide- binding site for HLA class II proteins (HLA-DR, -DQ, and -DP). Before the advent of NGS technology, only the exons encoding the ARD of an HLA gene were routinely sequenced to determine a patient's 2-field (high-resolution) HLA genotype. Therefore, clinical studies investigating the role of HLA matching in hematopoietic cell transplantation (HCT) outcome were based essentially on genetic matching of ARD-encoding exonic sequences between donors and recipients (as defined by 2-field HLA allele matching). However, polymorphisms in other regions of the HLA genes (ie, non-ARD exons, introns, and regulatory regions) may create antigenic determinants, alter the immunogenicity of the HLA protein through effects on expression, or impact binding of the HLA protein to other immunoreceptors, such as killer cell immunoglobulin-like receptor. With the advent of full-gene NGS approaches for HLA sequencing, non-ARD polymorphisms are now more easily identified, yet we do not fully understand the clinical impact of non-ARD HLA polymorphisms on transplantation outcome.

One challenge to elucidating the effect of non-ARD polymorphisms on hematopoietic cell transplantation (HCT) outcome is that current HLA allele nomenclature has inherent limitations for comparing donors and recipients at a genetic level. For example, sequential numbering in HLA allele nomenclature is driven by the date of discovery of the allele and is independent of the nature of the genetic difference between 2 alleles. Furthermore, it is difficult to use HLA allele nomenclature to assess the location and functional impact of genetic

Financial disclosure: See Acknowledgments on page 931.

* Correspondence and reprint requests: Matthew W. Anderson, MD, PhD, Diagnostic Laboratories, BloodCenter of Wisconsin, 638 N 18th Street, Milwaukee, WI 53233.
E-mail address: matthew.anderson@bcw.edu (M.W. Anderson).
† Contributed equally as primary co-authors of this manuscript.

https://doi.org/10.1016/j.bbmt.2018.12.006
1083-8791/

differences between donors and recipients. Finally, the IMGT/HLA database contains incompletely referenced alleles, such that not all alleles are equivalent with respect to the amount of non-ARD genetic information available for comparison.

To facilitate the investigation of the effect of HLA genetic differences on HCT outcome, we developed a bioinformatics approach to directly compare and donor and recipient full-length phased HLA consensus sequences generated by NGS. The strategy uses IMGT/HLA sequence data and alleles to define genetic features for alignment but allows for the identification and visualization of nucleotide differences in an unbiased manner. We used this approach to define HLA genetic differences between donor and recipient using a retrospective cohort of 166 unrelated HCT donor-recipient pairs. We then evaluated the HCT outcome measures for this cohort as defined by the degree of HLA genetic similarity.

METHODS

Samples

A total of 168 paired donor and recipient samples (336 samples total) were obtained from the Center for International Blood and Marrow Transplant Research (CIBMTR) repository. The recipients in this study underwent HCT with unrelated donors for hematologic malignancies at the Froedtert and the Medical College of Wisconsin from 1998 to 2015. Predicate 2-field HLA allele genotyping information was available for all donors and recipients for the *HLA-A*, *-B*, *-C*, and *-DRB1* loci at a minimum. The Medical College of Wisconsin's Institutional Review Board approval was obtained for this study.

DNA Extraction and Sequencing

Genomic DNA (gDNA) was extracted from frozen whole blood samples using QIAamp DNA Mini and Blood Mini manual DNA purification kits (Qiagen, Venlo, The Netherlands). The quantity of gDNA was measured by the Promega QuantiFluor dsDNA System (Promega, Madison, WI). The *HLA-A*, *-B*, *-C*, *DRB1/3/4/5*, *-DQA1*, *-DQB1*, *-DPA1*, and *-DPB1* loci were sequenced using the TruSight HLA v2 assay (Illumina, San Diego, CA). In brief, 50 ng of gDNA was used as input for each locus-specific amplification, and subsequent library preparation was performed according to the TruSight HLA v2 protocol [2]. Sequencing was performed on an Illumina MiSeq or MiSeqDx sequencer using either nano or micro flow cells with 2×150 paired-end sequencing. TruSight HLA Assign 2.1 software was used to generate HLA genotypes and phased consensus sequences (FASTA file format) from each locus. The IMGT/HLA version 3.29.0 database served as a source of validated reference sequence and genetic feature (ie, exon, intron, and untranslated region) annotation [3]. For the purposes of this study, we used an extended definition of the ARD-encoding exons to include exons 2, 3, and 4 for HLA class I alleles and exons 2 and 3 for HLA class II alleles. Two paired donor/recipient samples failed sequencing due to poor DNA quality and were excluded from the study, leaving 166 paired samples in the cohort. An additional 8 individual HLA loci failed amplification (1 *HLA-A*, 2 *HLA-C*, 1 *HLA-DPA1*, 2 *HLA-DPB1*, 1 *HLA-DQA1*, and 2 *HLA-DQB1*), and were removed from the analysis. An additional 4 individual *HLA-DQB1* results were removed because of a technical error. Removal of individual failed HLA loci did not affect categorization of the HCT pairs for the analysis of outcome measures. Predicate (Sanger) and NGS HLA genotyping results for each locus for each study subject are reported in Supplementary Table 3.

Patients

To evaluate the clinical outcomes in the unrelated HCT cohort, donor-recipient pairs (n = 166) were categorized into 3 groups based on the presence of high-expression loci (HEL) mismatch, as defined by a nonsynonymous genetic difference within the ARD-encoding exons of the *HLA-A*, *-B*, *-C*, or *-DRB1* loci (equivalent to an allele-level HLA mismatch; n = 27), the absence of any HEL mismatch, and the presence of at least 1 low-expression loci (LEL) mismatch, as defined by a nonsynonymous genetic difference in the ARD-encoding exons of the *HLA-DRB3/4/5*, *-DQA1*, *-DQB1*, *-DPA1*, or *-DPB1* loci (equivalent to an allele-level HLA mismatch; n = 125) and the absence of any sequence differences within the exonic regions (no exonic mismatch [NEM]; n = 14). The patients in the HEL group underwent allogeneic HCT (alloHCT) between January 2005 and November 2014, those in the LEL group underwent alloHCT between January 2005 and April 2015, and NEM patients underwent alloHCT between February 2007 and December 2014.

Software

The DaRMA (donor and recipient matching algorithm) was written in Python 3 (https://www.python.org/downloads/release/python-360) and has 3 external dependencies: pandas (https://pandas.pydata.org), Biopython (http://biopython.org) and MAAFT (https://mafft.cbrc.jp/alignment/software) [4]. MAAFT was used to align the paired consensus sequences for each sequenced HLA gene. The aggregate aligned data produced by DaRMA is imported into VIRAS (visualization interface to represent alignment statistics), a tool that enables visualization of each donor-recipient alignment with corresponding statistics. VIRAS is a single-page web application built using the Angular JS framework (https://angularjs.org). Data visualizations were created using D3 Javascript library (https://d3js.org) and tested for compatibility with multiple web browsers, including Mozilla Firefox, Google Chrome, and Safari.

Statistical Analysis

Patient-, disease-, and transplantation-related variables were compared among the 3 groups using chi square statistics for categorical variables and the Kruskal-Wallis test for continuous variables. Kaplan-Meier estimates were used to calculate the probability of overall survival (OS) and progression- free survival (PFS). OS was defined as the time from HCT to death from any cause or the last follow-up. Surviving patients were censored at last follow-up. PFS was defined as the time from HCT to either relapse or death from any cause, and survivors were censored at the last follow-up if not in relapse. Nonrelapse mortality (NRM) was defined as death from any cause in continuous remission and was summarized by cumulative incidence estimates, with relapse as the competing risk. Relapse was summarized by cumulative incidence estimates, with NRM as the competing risk. For relapse and NRM, patients in continuous complete remission were censored at last follow-up. Acute graft-versus-host disease (GVHD) was defined by standard criteria [5]. Probabilities of acute GVHD were calculated using the cumulative incidence method to account for the competing risk of death without the event. Multivariable analysis (Table 3) was conducted to evaluate hazard ratios (HRs) for OS, PFS (by Cox proportional hazards regression models), relapse, NRM, and acute GVHD (Fine-Gray proportional hazards regression models) among the HEL, LEL, and NEM groups, considering the covariates of recipient age, graft source, donor-recipient cytomegalovirus serostatus, Karnofsky Performance Status score, and $CD34^+$ and $CD3^+$ cell doses. SAS version 9.4 (SAS Institute, Cary, NC) was used for statistical analysis.

RESULTS

DaRMA Alignment of Donor and Recipient Consensus HLA Sequences

To determine the degree of HLA genetic similarity between HCT donors and recipients, we used the phased consensus HLA sequences and HLA genotypes produced by the Assign HLA genotyping software (Illumina). In clinical practice, donors are selected relative to the degree of HLA match with the intended recipient. Therefore, the DaRMA algorithm (Figure 1) was designed to describe and annotate all genetic differences with respect to the recipient. To align donor and recipient phased consensus sequences, DaRMA uses HLA genotyping information to determine which consensus sequences (from each allele at each HLA locus) should be aligned together based on similarity of the HLA allele. For example, in the case of a donor-recipient pair that has a single allele-level mismatch at *HLA-A*, sequences from the matched *HLA-A* alleles will be aligned together, and the mismatched *HLA-A* alleles will be aligned together. Pairwise alignments of the donor and recipient sequences are then performed using MAAFT, and nucleotide-level genetic differences between the donor and recipient sequences are identified. To annotate these differences to specific genetic features (ie, exons, introns, and untranslated regions), exonic sequences associated with the given recipient HLA type are extracted from IMGT/HLA for each recipient allele and used as "anchors" to define both coding and noncoding regions. For alleles in which all exons are fully referenced in IMGT/HLA, this process is straightforward (Figure 1A). For alleles that are not fully referenced in IMGT/HLA (ie, alleles with only exon 2 and 3 nucleotide sequences), generic exonic sequences are generated from all other related alleles in IMGT/HLA to allow for genetic feature assignment across the length of the sequenced region. Once the anchoring process is complete, donor and recipient differences are annotated by genetic features, and amino acid changes are calculated if the

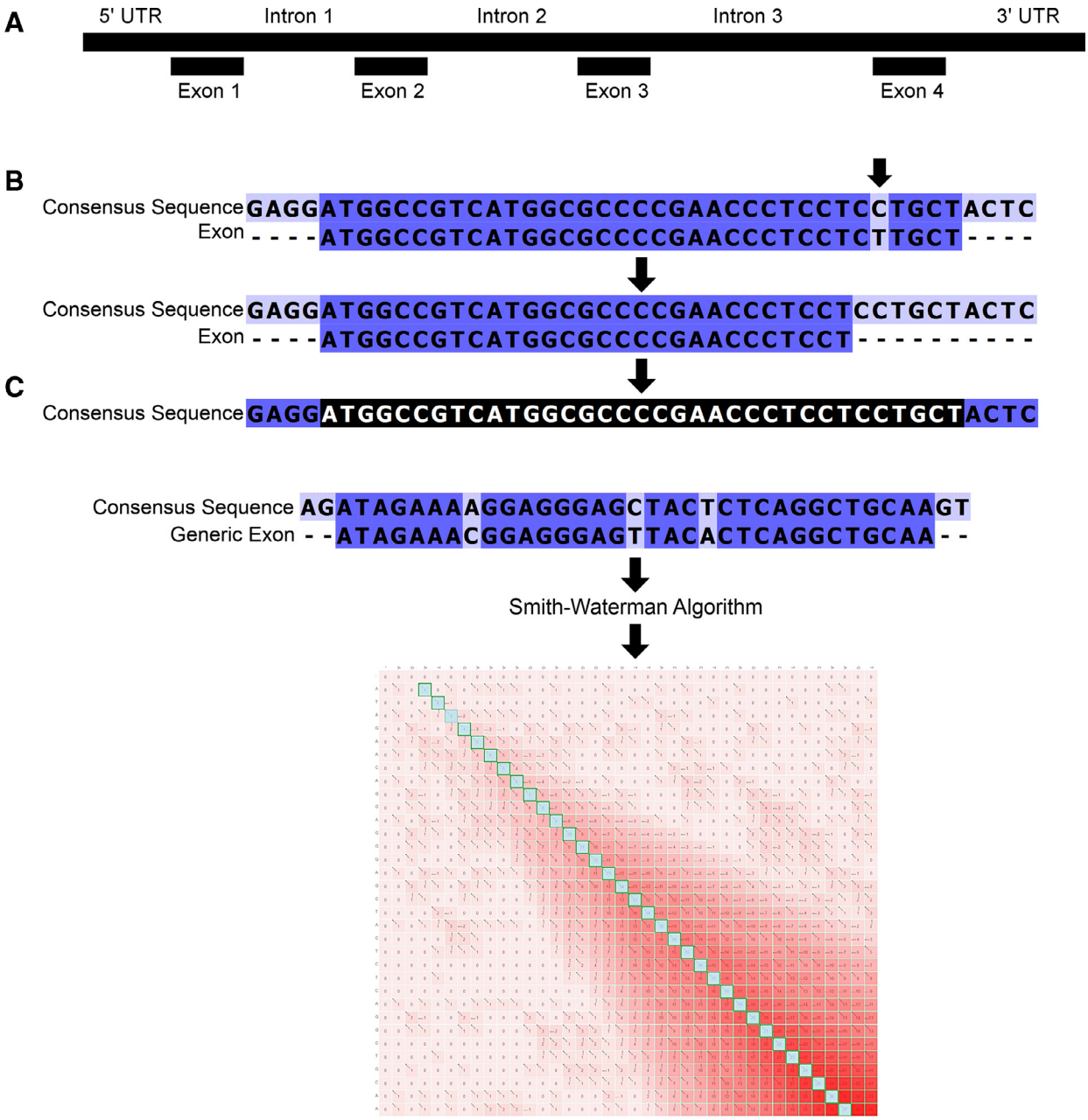

Figure 1. Graphical representation of determination of genetic features with DaRMA. (A) If a consensus sequence has a complete exonic reference in IMGT/HLA, the identification of genetic features is straightforward. (B) If a consensus sequence differs from the IMGT/HLA reference by only a few nucleotides (ie, sequencing errors or novel variants), then DaRMA continues the alignment process in an iterative fashion, excluding nucleotides from the 5′ or 3′ end of the exon until a perfect match is found. (C) For consensus sequences that have incomplete reference exonic sequence in IMGT/HLA, generic exons are used that may have multiple nucleotide differences with respect to the consensus exon. Using an alignment approach (Smith-Waterman algorithm) that compares sequences of all possible lengths, an optimal alignment between the generic exon and the consensus sequence is determined. The depth of shading in the alignment table is directly correlated to the magnitude of identity between the 2 sequences, and the light-shaded path represents the optimal alignment between the 2 sequences.

difference resides within a coding region. Sequential nucleotide differences (insertions or deletions) are also annotated.

The anchoring process is highly dependent on the accuracy of the consensus sequence and the amount of reference sequence available for a given allele in IMGT/HLA. Poor sequence quality can create errors in the consensus sequence, including false insertions/deletions, which then complicate the anchoring process. In addition, long stretches of sequence homogeneity (*e.g.*, homozygosity for a particular HLA locus) can result in incorrect phase assignments within the consensus sequence and false-positive genetic differences. If an exon from IMGT/HLA does not perfectly match the consensus sequence (*e.g.*, error in the consensus sequence, novel sequence variant, localized phase break), DaRMA scales back the IMGT/HLA reference exon from either the 5′ or 3′ direction until a sufficiently long (>20 bp) perfect match is found, then the exon is extended over the mismatched sequence to establish the intron/exon boundary (Figure 1B).

If multiple nucleotide differences are encountered between a consensus sequence exon and the reference exon from IMGT/HLA, alignment failure may occur due to an absence of a >20 bp sequence match. In these cases, the Smith-Waterman algorithm

was used for alignment. The Smith-Waterman approach performs local sequence alignment by creating an $m \times n$ computation matrix, where m and n are the lengths of the 2 sequences being aligned. A predefined substitution matrix for matches, mismatches, and gaps is applied to each element in the matrix. After the computation matrix is scored, the optimal local alignment is traced back, from the bottom up. An example of a finished computation matrix from our study cohort is shown in Figure 1C. A total of 15,552 consensus sequence exons were analyzed using DaRMA. Of that total, 15,063 (96.85%) of the reference exons from IMGT/HLA aligned perfectly to the consensus sequence. An additional 427 exons were aligned successfully by scaling back the alignment from either the 5′ or the 3′ end of the reference exon, and 62 exons were saved from alignment failure by use of the Smith-Waterman algorithm.

Use of VIRAS to Visualize and Group HCT Pairs by Genetic Differences

A web-based visualization tool was created to quickly and easily summarize and visualize genetic differences between HCT pairs (Figure 2, and Supplementary Video 1). Aggregated pairwise comparison data (JSON file format) generated by DaRMA serve as the input files for VIRAS. First, a summary of mismatches for all donor-recipient pairs is presented in an interactive tabular form. This allows the user to perform sorting and filtering of pairs based on the number of genetic differences, location of a difference within a genetic feature, and impact of that difference on the HLA protein (synonymous versus nonsynonymous). To obtain a summary of the genetic differences present at each HLA locus, VIRAS also has a locus-specific view that sorts genetic differences based on different parameters, including total number, type, and location. The tabular data also can be converted into the form of interactive bar-and-donut charts for an overview of the different types of genetic differences within each HLA locus. Finally, the sequence view displays the aligned donor and recipient consensus sequences, enabling the user to visualize genetic differences between the recipient and donor at various levels of resolution. Specific donor and recipient sequence differences also can be directly visualized using a clickable list that allows the user to view a particular mismatch and then analyze surrounding regions of interest.

We first used the alignment, filtering, searching, and sorting processes in DaRMA and VIRAS to determine the overall degree of genetic difference between donor and recipient HLA allele pairs in our HCT cohort. As shown in Figure 3A, there were relatively few (<4%) HEL *HLA-A*, *-B*, *-C*, and *-DRB1* allele pairs with a nonsynonymous core exon genetic difference (equivalent to a 2-field HLA allele mismatch), consistent with our clinical practice of selecting unrelated donors matched at these loci wherever possible. However, up to 20% of LEL *HLA-DQB1*, *-DQA1*, *-DRB3/4/5*, *-DPA1*, and *-DPB1* allele pairs showed a nonsynonymous core exon difference (equivalent to a 2-field HLA allele mismatch). When we restricted the analysis to those allele pairs

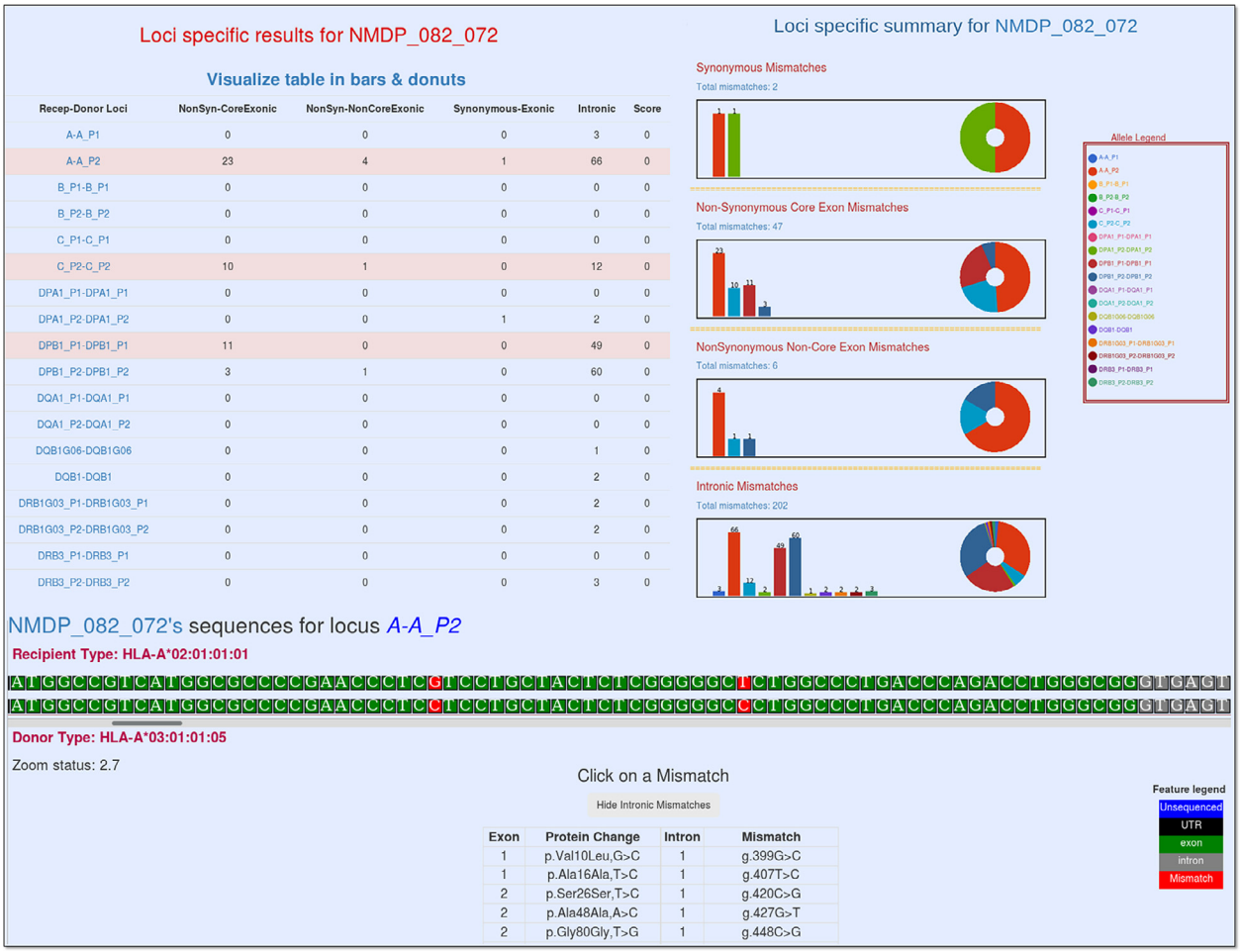

Figure 2. VIRAS provides a graphical representation of matching statistics for each locus along with visualization tools, including sequence-level alignments. In this figure, "core" refers to the ARD-encoding exons.

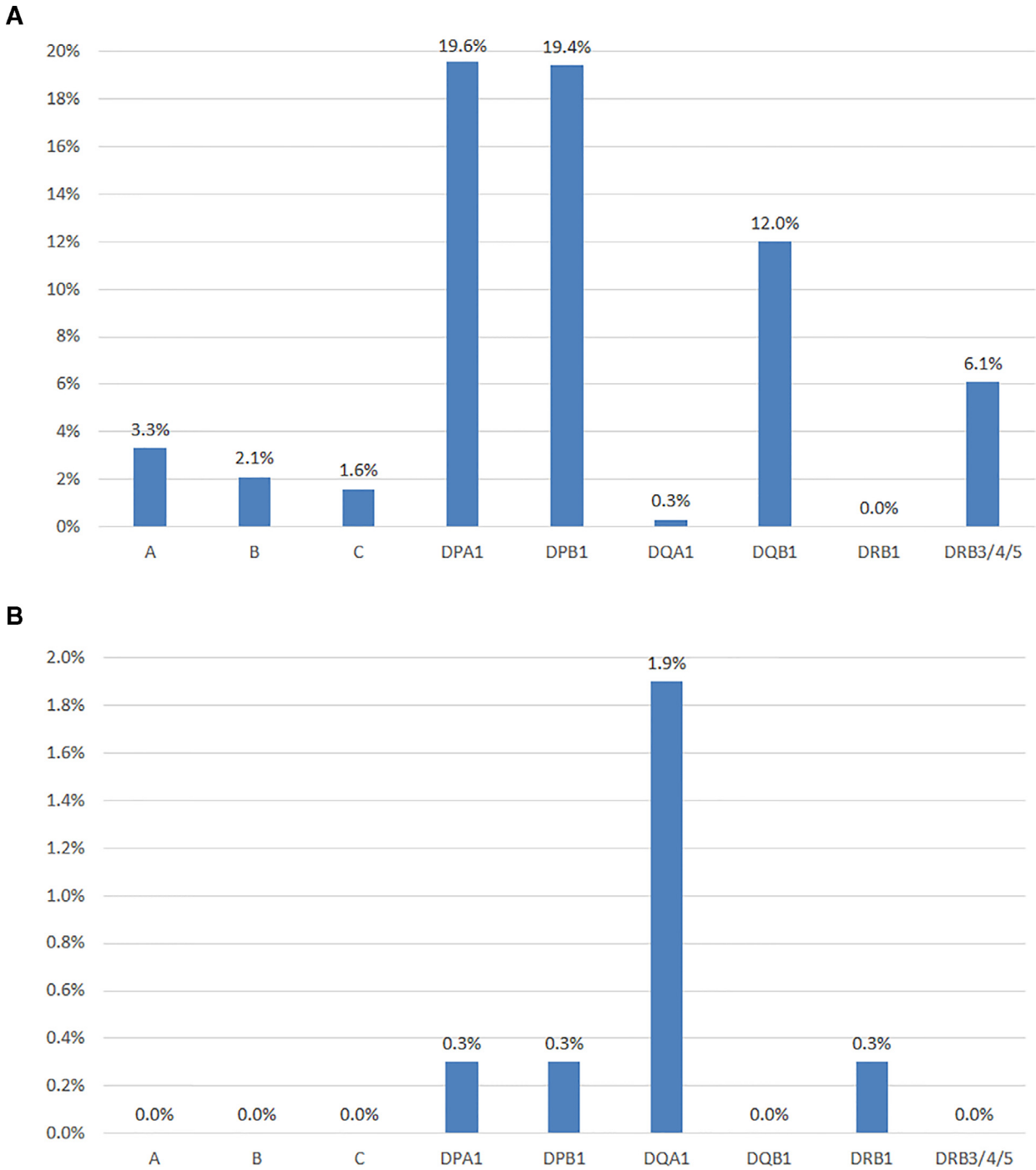

Figure 3. Overall characterization of HLA genetic differences in the study cohort. (A) Enrichment for the LEL is observed when filtering the cohort for HLA allele pairs with at least 1 nonsynonymous genetic difference in both ARD-encoding and non-ARD-encoding exons. (B) In the presence of no nonsynonymous differences in the ARD, very few HLA allele pairs exhibit nonsynonymous changes outside of the ARD-encoding exons.

that had no nonsynonymous genetic differences within the ARD, we found very few nonsynonymous genetic differences across the remaining exonic sequence (<2%; Figure 3B). The percentage of allele pairs with at least 1 single nucleotide difference outside of the ARD-encoding exons varied by locus, ranging from 45% of *HLA-DPA1* allele pairs to <4% of *HLA-A*, *-B*, and *-C* locus pairs (data not shown).

We next used DaRMA and VIRAS to categorize each of the 166 HCT pairs in the study cohort. We identified a total of 139 HCT pairs that exhibited no *HLA-A*, *-B*, *-C*, or *-DRB1* nonsynonymous genetic differences within the ARD-encoding exons (generally equivalent to a 2-field HLA allele match). Of those 139 HCT pairs, 14 were highly genetically matched, exhibiting no nonsynonymous or synonymous genetic differences in any sequenced exon (ie, NEM). The relatively small number of these highly genetically matched NEM HCT pairs was consistent with the amount of LEL mismatching in the study cohort. Five of the 14 NEM pairs showed no genetic difference for any locus, 5 of 14 pairs differed only by a single nucleotide change in intron 1 of *HLA-DPA1* (*HLA-DPA1**01:03:01:02/04; rs9277342), and 4 of 14 pairs exhibited at

least 1 intronic difference in another locus, some of which were not referenced in IMGT/HLA. Of note, the NEM pairs would not have been identified simply by visual comparisons of HLA genotyping results owing to residual genotype ambiguity at some of the HLA loci (Supplementary Table 2).

Transplantation Outcomes as Defined by Degree of HLA Genetic Similarity

We examined the impact of high genetic similarity within the HLA loci by describing and comparing the post-HCT outcomes among the NEM (n = 14), HEL-mismatched (n = 27), and LEL-mismatched (n = 125) patients. The baseline patient, disease, and transplantation characteristics of the 3 groups were similar with a few exceptions (Table 1). The HEL patients were significantly younger (median age, 43 years, compared with 55 and 57 years in the LEL and NEM groups, respectively; $P = .04$). A significantly higher proportion of the HEL group received myeloablative conditioning and antithymocyte globulin (ATG) compared with the other 2 groups. Importantly, the conditioning and GVHD prophylaxis regimens for the 3 groups were determined before alloHCT, based on the results of traditional HLA allele typing exhibiting the allelic match/mismatch. The median duration of follow-up of survivors was 5.0, 4.3, and 3.8 years in the HEL, LEL and NEM groups, respectively. The results of univariate analysis demonstrating the impact of HLA genetic similarity on post-HCT outcomes are presented in Table 2.

On multivariable analysis, OS and PFS were significantly better in the LEL and NEM groups compared with the HEL group (Table 3 and Figure 4A and B). The median OS was 1.3 years in the HEL group, 6.7 years in the LEL group, and was not reached in the NEM group ($P = .001$) (Table 2). The probability of OS at 2 years was 41% in the HEL group, 69% in the LEL group, and 85% in the NEM group. The median PFS was .8 year in the HEL group, 6.7 years in the LEL group, and 4.4 years in the NEM group; $P = .005$) (Table 2). The 2-year PFS was 37%, 61%, and 70% in the HEL, LEL, and NEM groups, respectively. NRM risk was not significantly different among the groups; the 2-year cumulative incidence of NRM was 33% in the HEL group, 18% in the LEL group, and 8% in the NEM group ($P = .13$) (Table 2 and Figure 4C). The cumulative incidence of relapse at 2 years post-HCT was also not significantly different among the HEL (30%), LEL (20%), and NEM (22%) groups (Table 2). $CD3^+$ T cell dose was a significant covariate for NRM on multivariate analysis (HR, .72 for a 10-fold increase in $CD3^+$ cell dose; $P = .02$). The cumulative incidences of grade II-IV and grade III-IV acute GVHD at day +180 post-HCT were 44%, 37%, and 28% ($P = .56$) and 15%, 15%, and 0% ($P = .30$) for the HEL, LEL, and NEM groups, respectively (Table 2 and Figure 4D). The risk of grade III-IV acute GVHD was significantly lower in patients who received a graft from a highly HLA genetically similar donor (NEM) ($P < .0001$; Table 3), albeit as a result of a complete lack of events in the NEM group.

DISCUSSION

The success of HCT depends critically on a high degree of HLA similarity between donor and recipient, largely defined by an analysis of ARD-encoding exonic sequences [6]. With the advent of NGS-based approaches to HLA genotyping, new HLA alleles are being identified at a prodigious rate [7], and more sequence information is being generated from regions outside the ARD-encoding exons. However, the clinical impact of HLA genetic differences outside the ARD remains unclear. By directly comparing donor and recipient consensus HLA sequences to identify and visualize regions of sequence heterogeneity, our approach offers an additional means of comparison beyond traditional HLA allele assignment. DaRMA/VIRAS can identify, localize, and visualize genetic differences between donor and recipient HLA alleles that are not easily described using HLA allele nomenclature. In addition, the automated pipeline enables rapid identification of donor–recipient pairs that demonstrate very few to no sequence differences in the HLA loci. An alternative approach for HLA allele comparison is gene feature enumeration (GFE) [8]. Although GFE can identify the location of genetic differences between HLA alleles, the numbering system is relative to the longest known nucleotide sequence for a given gene feature and does not relate to the impact of a genetic difference (eg, synonymous). In contrast, our approach can facilitate large HCT cohort studies by describing genetic differences specific for a particular pair of donor–recipient HLA alleles, providing a simple and rapid means of classifying HCT pairs into different groups based on HLA genetic similarity.

In this cohort of 166 HCT pairs, we found very few nonsynonymous genetic differences in non-ARD exons when no nonsynonymous differences were found within the ARD-encoding exons. This result is consistent with previously published data suggesting a lesser degree of polymorphism in exonic regions outside the ARD [9]. However, we also found that a significant percentage of HLA allele pairs had at least 1 genetic difference, consistent with the high degree of genetic polymorphism within the HLA loci. Although these genetic differences are not predicted to directly impact the HLA protein sequence, polymorphic determinants outside the ARD may contribute to indirect allorecognition [10] or HLA gene expression [11]. Therefore, it may be important to further investigate donor and recipient differences in these regions. Given that our cohort was composed of unrelated individuals, it is interesting that a significant proportion (14 of 166; 8%) were highly HLA genetically matched, with 5 unrelated donor–recipient pairs showing no genetic differences across all the HLA loci sequenced (encompassing ~64 kb of DNA sequence). Interestingly, another 5 donor–recipient pairs exhibited only 1 nucleotide difference in intron 1 of *HLA-DPA1*. This single nucleotide polymorphism (SNP) is relatively common in the global population, with an allele frequency of .214 (http://gnomad.broadinstitute.org/variant/6-33039804-C-T). Why this SNP was the sole identified genetic difference in these highly genetically similar HCT pairs is unclear, given the strong linkage disequilibrium in the HLA region.

In this study, we have demonstrated that the recipient-unrelated donor pairs with NEM (HR, .74; $P = .02$) or LEL HLA mismatch (HR, .57; $P = .003$) had superior OS compared to those with a HEL HLA mismatch. Both LEL group (HR, .62; $P = .0006$) and the NEM group (HR, .63; $P = .03$) also had significantly better PFS compared with the HEL group, consistent with previous studies [12]. The significantly lower OS and PFS was driven by the increased risk of relapse and NRM with HEL HLA mismatching, which was reflected by the causes of death in the HEL group (Supplementary Table 1). No significant difference in any of the survival outcomes was observed between the NEM and LEL groups. Examining the acute GVHD risk, the NEM group had a significantly low incidence of grade III-IV acute GVHD (0% at day +180). The incidence of grade II-IV acute GVHD in the NEM group was similar to that in the LEL group but was driven entirely by grade II events in the NEM group. In fact, the cumulative incidence of grade II acute GVHD was similar among the 3 groups (30% in the HEL group, 22% in the LEL group, and 28% in the NEM group. In line with the greater degree of genetic variation between unrelated donor-recipient pairs and significantly more disparate minor histocompatibility antigens in the graft-versus-host direction [13], the relatively high incidence of grade II acute GVHD in the NEM group is not surprising. Our analysis also found no statistically significant differences in the risk of NRM among the groups, although it is

likely that a larger sample size would have demonstrated a difference. It is important to recognize that a reasonable number of donor–recipient pairs (10%; 14 of 139) without an HEL HLA mismatch (*i.e.*, the LEL and NEM groups) had NEM.

In general, matched related donors are generally preferred over well-matched unrelated donors owing to a lower incidence of GVHD [14,15]. This observation may be explained in part by the fact that matched related donors are more likely to share HLA haplotypes or large blocks of genetic sequences located on the same

Table 1
Baseline Patient, Disease, and Transplantation Characteristics

Variable	HEL (n = 27)	LEL (n = 125)	NEM (n = 14)	*P* Value
Age, yr, median (range)	43.5 (21.4-67.9)	54.6 (20.1-74.4)	57.4 (26.0-67.0)	
Sex, n (%)				.34*
Female	16 (59.3)	55 (44.0)	7 (50.0)	
Male	11 (40.7)	70 (56.0)	7 (50.0)	
Disease, n (%)				.14*
ALL	2 (7)	9 (7)	1 (7)	
AML	17 (63)	57 (46)	8 (57)	
CLL	0 (0)	8 (6)	0 (0)	
CML	2 (7)	2 (2)	0 (0)	
CMM	0 (0)	4 (3)	0 (0)	
LGL	1 (4)	0 (0)	0 (0)	
MDS	3 (11)	23 (18)	1 (7)	
MM	0 (0)	8 (6)	0 (0)	
MPN	0 (0)	2 (2)	1 (7)	
NHL	1 (4)	7 (6)	3 (21)	
SAA	0 (0)	4 (3)	0 (0)	
cHL	1 (4)	1 (1)	0 (0)	
Chemosensitivity, n (%)				1.00*,†
Chemoresistant	6 (22)	28 (23)	3 (21)	
Chemosensitive	21 (78)	93 (77)	11 (79)	
Disease status at alloHCT, n (%)				.63*,†
Less than PR	8 (30)	30 (25)	5 (36)	
PR or greater	19 (70)	91 (75)	9 (64)	
KPS score, median (range)	90 (70-100)	90 (50-100)	90 (70-100)	.44‡
Pretransplantation conditioning intensity, n (%)				**.046***,†
MAC	23 (85)	75 (60)	9 (64)	
RIC	4 (15)	50 (40)	5 (36)	
ATG in conditioning, n (%)				<**.001***,†
No	9 (33)	121 (97)	14 (100)	
Yes	18 (67)	4 (3)	0 (0)	
TBI in conditioning regimen, n (%)				.59*,†
No	23 (85)	95 (76)	11 (79)	
Yes	4 (15)	30 (24)	3 (21)	
Conditioning regimen, n (%)				<**.001**†
Bu/Cy ± ATG	10 (37)	33 (26)	3 (21)	
Cy/TBI ± ATG	2 (7)	15 (12)	2 (14)	
Flu/Bu2	1 (4)	22 (18)	3 (21)	
Flu/Bu4 ± ATG	9 (33)	25 (20)	3 (21)	
Flu/TBI (2 Gy)	0 (0)	9 (7)	1 (7)	
Others	5 (18)	21 (17)	2 (14)	
GVHD prophylaxis, n (%)				.67*,†
FK + MTX	25 (93)	97 (78)	12 (86)	
FK +MTX + other	0 (0)	17 (13)	1 (7)	
FK + MMF ± other	1 (4)	11 (9)	1 (7)	
Ex vivo T cell depletion	1 (4)	0 (0)	0 (0)	
Graft source, n (%)				.61†
Bone marrow	10 (37)	44 (35)	3 (21)	
Peripheral blood stem cells	17 (63)	81 (65)	11 (79)	
Donor sex, n (%)				.07*,†
Female	11 (41)	33 (26)	1 (7)	
Male	16 (60)	92 (74)	13 (93)	
Donor-recipient sex match, n (%)				.61†
Match	16 (59)	67 (54)	6 (43)	
Mismatch	11 (41)	58 (46)	8 (57)	
Recipient CMV serology, n (%)				.59†
+	9 (33)	51 (41)	4 (29)	
-	18 (67)	74 (59)	10 (71)	
Donor CMV serology, n (%)				.84*,†
+	7 (26)	40 (32)	4 (29)	
-	20 (74)	84 (68)	10 (71)	
ABO match, n (%)				.51†
Match	11 (41)	44 (35)	7 (50)	
Mismatch (major/minor)	16 (59)	81 (65)	7 (50)	
$CD34^+$ cell dose, $\times 10^6$/kg, median (range)	4.9 (.6-19.6)	5.8 (.7-31.7)	7.6 (1.7-28.3)	.29‡

(*continued*)

Table 1 (*Continued*)

Variable	HEL (n = 27)	LEL (n = 125)	NEM (n = 14)	*P* Value
$CD3^+$ cell dose, $\times 10^6$/kg, median (range)	197.0 (0-940.0)	234.0 (7.1-804.0)	348.0 (14.4-526.0)	.30[‡]
Year of transplantation, n (%)				.03[*,†]
2005-2008	8 (29.6)	19 (15.2)	1 (7.1)	
2009-2011	12 (44.4)	38 (30.4)	3 (21.4)	
2012-2015	7 (25.9)	68 (54.4)	10 (71.4)	
Follow-up of survivors, yr, median	5.0	4.3	3.8	

Significant **P** values are in bold type.
AML indicates acute myelogenous leukemia; MDS, myelodysplastic syndrome; MPN, myeloproliferative neoplasm; ALL, acute lymphoblastic leukemia; LGLL, large granular lymphocytic leukemia; CML, chronic myeloid leukemia; CMML, chronic myelomonocytic leukemia; MM, multiple myeloma; NHL, non-Hodgkin lymphoma; cHL, classic Hodgkin lymphoma; SAA, severe aplastic anemia; NA, not applicable; PR, partial response; KPS, Karnofsky Performance Status; MAC, myeloablative conditioning; RIC, reduced-intensity conditioning; TBI, total body irradiation; Bu, busulfan; Cy, cyclophosphamide; Flu, fludarabine; FK, tacrolimus; MTX, methotrexate; MMF, mycophenolate mofetil; CMV, cytomegalovirus.
[*] Chi-square test.
[†] Fisher exact test.
[‡] Kruskal-Wallis test.

Table 2
Comparison of Post-Transplantation Outcomes Based on Recipient-Donor HLA Matching

Outcome	HEL mismatch (n = 27)	LEL mismatch (n = 124)	NEM (n = 15)	*P* Value
OS, yr, median (range)	1.3 (.8-3.3)	6.7 (3.5-∞)	NR (2.5-∞)	**.001**
1 yr OS, %	55.6 (35.2-71.8)	77.6 (69.2-84)	85.1 (52.3-96.1)	
2 yr OS, %	40.7 (22.5-58.2)	69.5 (60.6-76.8)	85.1 (52.3-96.1)	
PFS, yr, median (range)	.8 (.5-2.9)	6.7 (2.3-∞)	4.4 (.8-∞)	**.005**
1 yr PFS, %	40.7 (22.2-58.2)	70.4 (61.6-77.6)	70.1 (38.5-87.6)	
2 yr PFS, %	37 (19.6-54.6)	61.5 (52.4-69.4)	70.1 (38.5-87.6)	
Relapse, %				.41
1 yr	29.6 (13.7-47.5)	16.8 (10.8-23.9)	22.1 (4.5-47.0)	
2 yr	29.6 (13.7-47.5)	20.0 (13.5-27.5)	22.1 (4.5-47.0)	
NRM, %				.13
1 yr	29.6 (13.6-47.6)	12.8 (7.6-19.3)	7.8 (0-30.8)	
2 yr	33.3 (16.3-51.4)	18.4 (12.2-25.7)	7.8 (0-30.8)	
Grade II-IV acute GVHD, %				.56
Day +100	40.7 (22.1-58.6)	31.2 (23.2-39.4)	28.5 (8.2-53.3)	
Day +180	44.4 (25.1-62.2)	37.6 (29.1-46.0)	28.5 (8.2-53.3)	
Grade III-IV acute GVHD, %				.30
Day +100	14.8 (4.5-30.8)	10.4 (5.8-16.5)	0	
Day +180	14.8 (4.5-30.8)	15.2 (9.5-22.1)	0	

Significant *P* values are in bold type.
NR indicates not reached.

chromosome. Petersdorf *et al.* [16] found that patients receiving HLA haplotype–matched unrelated donor grafts had a reduced incidence of GVHD, suggesting that HLA haplotype–matched donors share additional undefined transplant antigens within the HLA region. However, owing to linkage disequilibrium, haplotype–matched unrelated donors are also likely to share extended HLA polymorphisms. Therefore, the increased frequency of GVHD in non-haplotype–matched donors also might reflect immune responses against non-ARD HLA polymorphic determinants or other transplant antigens within the MHC region. Although whether these highly matched pairs identified in our study are haplotype–matched or –mismatched is unclear, we expect that the presence of NEM may indicate broader genetic relatedness across the HLA and MHC regions and thus may be analogous to the haplotype–matched donor–recipient pairs in the study reported by Petersdorf *et al.* [16].

This approach has implications for clinical practice regarding donor selection for allogeneic HCT given that the donor–recipient pairs with NEM had a lower risk of severe (grade III-IV) acute GVHD and a survival advantage compared with HEL-mismatched pairs. The HEL group was expected to have a greater risk of grade II-IV and grade III-IV acute GVHD based on the pretransplantation knowledge of the HLA allele mismatch, and thus two-thirds of the patients in this group were treated with ATG in the conditioning regimen. Of note, a significantly greater proportion of patients (85%; n = 23) in the HEL group received myeloablative conditioning (two-thirds of which [n = 18] included ATG), which possibly contributed to the higher risk of acute GVHD. There were no statistically significant differences in OS, PFS, or relapse risk between the NEM and LEL groups. Future studies with larger NEM cohorts are warranted to conclusively answer the question of whether patients with unrelated NEM donors have better OS compared with those with LEL HLA–mismatched donors owing to a reduced risk of acute GVHD and NRM. In addition, if a greater risk of relapse is demonstrated with NEM versus LEL HLA mismatch in a larger study, it will be critical to discern the mechanism behind this increased risk; without additional data, we can speculate that lower risk of GVHD and potentially decreased graft-versus-malignancy effect contribute to the higher relapse rate.

We acknowledge some limitations of this study, including its retrospective nature, single-center setting, and limited power to detect differences in outcomes. In addition to the smaller sample size of the NEM group, the extreme imbalance of ATG treatment among the 3 groups (67% in the HEL group, 3% in the LEL group, and 0 in the NEM group) is a major limitation. As a result, regression-based adjustment was infeasible. As mentioned previously, the conditioning regimen (including ATG) was chosen before transplantation based on the HLA mismatch by traditional HLA typing. It is possible that with a larger unrelated HCT cohort, we would find additional NEM HCT pairs, which would increase

Table 3
Multivariable Analysis of the Study Population to Evaluate the Impact of HLA Disparity by NGS-Based Genotyping and the Impact of Covariates on Post-Transplantation Outcomes

Variables	HR	95% CI	*P* Value
Overall survival			**.005**
LEL	1.00		
HEL	2.33	1.33-4.09	**.003**
NEM	1.00		
HEL	3.82	1.26-11.54	**.02**
LEL	1.64	.58-4.64	.35
Graft source			.33
Bone marrow	1.00		
Peripheral blood	.68	.32-1.47	
Conditioning intensity			.97
RIC	1.00		
MAC	1.01	.58-1.76	
CMV serostatus			.45
Others	1.00		
D-/R-	.83	.51-1.34	
Disease group			.42
Others	1.00		
AML/ALL	1.23	.74-2.03	
KPS score			.28
<90	1.00		
90-100	.76	.46-1.25	
Age, 1 yr	1.02	.99-1.04	.06
Year of transplantation			.64
2012-2015	1.00		
2009-2011	1.005	.55-1.84	.99
2005-2008	1.37	.65-2.87	.40
CD3, 10-fold increase	.77	.56-1.07	.12
CD34, 10-fold increase	1.21	.46-3.21	.70
PFS			**.002**
LEL	1.00		
HEL	2.65	1.50-4.54	**.0006**
NEM	1.00		
HEL	2.73	.95-5.30	**.03**
LEL	1.03	.46-2.30	.94
Graft source			.55
Bone marrow	1.00		
Peripheral blood	.81	.41-1.61	
Conditioning intensity			.22
RIC	1.00		
MAC	.73	.44-1.21	
CMV status			.55
Others	1.00		
D-/R-	.87	.55-1.37	
Disease group			.90
Others	1.00		
AML/ALL	.97	.61-1.53	
KPS score			.18
<90	1.00		
90-100	.73	.46-1.16	
Year of transplantation			.50
2012-2015	1.00		
2009-2011	.71	.40-1.25	.24
2005-2008	1.15	.59-2.26	.68
Age, 1 yr	1.00	.98-1.02	.86
CD3, 10-fold increase	.84	.63-1.13	.25
CD34, 10-fold increase	1.31	.54-3.16	.54
Relapse			.09
LEL	1.00		
HEL	2.35	1.09-5.06	**.03**
NEM	1.00		
HEL	2.04	.60-6.93	.25
LEL	.87	.30-2.50	.79
Graft source			.16
Bone marrow	1.00		
Peripheral blood	1.87	.78-4.49	
Conditioning intensity			.08
RIC	1.00		
MAC	.53	.26-1.07	
CMV serostatus			.65
Others	1.00		
D-/R-	1.16	.61-2.14	
Disease group			.86
Others	1.00		
AML/ALL	.95	.52-1.73	
KPS score			.45
<90	1.00		
90-100	.77	.39-1.52	
Year of transplantation			.30
2012-2015	1.00		
2009-2011	.52	.22-1.26	.15
2005-2008	.58	.23-1.48	.25
Age, 1 yr	.99	.96-1.01	.40
CD3, 10-fold increase	1.23	.82-1.85	.32
CD34, 10-fold increase	1.68	.56-5.00	.35
NRM			.34
LEL	1.00		
HEL	1.74	.79-3.84	.17
NEM	1.00		
HEL	1.92	.59-6.28	.28
LEL	1.10	.36-3.34	.86
Graft source			.05
Bone marrow	1.00		
Peripheral blood	.35	.12-1.03	
Conditioning intensity			.29
RIC	1.00		
MAC	1.46	.72-2.97	
CMV serostatus			.31
Others	1.00		
D-/R-	.72	.38-1.37	
Disease group			.82
Others	1.00		
AML/ALL	1.08	.54-2.15	
KPS score			.67
<90	1.00		
90-100	.87	.46-1.66	
Year of transplantation			.64
2012-2015	1.00		
2009-2011	1.27	.61-2.63	.52
2005-2008	1.63	.56-4.73	.36
Age, 1 yr	1.02	.98-1.06	.31
CD3, 10-fold increase	.72	.54-.96	**.02**
CD34, 10-fold increase	.65	.15-2.82	.56
Grade II-IV acute GVHD			.40
LEL	1.00		
HEL	1.46	.74-2.84	.27
NEM	1.00		
HEL	2.03	.62-6.59	.24
LEL	1.39	.47-4.10	.54
Graft source			.31
Bone marrow	1.00		
Peripheral blood	1.90	.55-6.64	
Conditioning intensity			.58
RIC	1.00		
MAC	.84	.46-1.56	
CMV serostatus			.99
Others	1.00		
D-/R-	1.00	.57-1.77	
Disease group			.19
Others	1.00		
AML/ALL	.68	.39-1.20	
KPS score			.38
<90	1.00		
90-100	1.30	.72-2.33	
Year of transplantation			.33
2012-2015	1.00		
2009-2011	1.25	.66-2.35	.49
2005-2008	1.80	.82-3.95	.14
Age, 1 yr	1.01	.98-1.03	.44
CD3, 10-fold increase	1.90	.72-5.05	.20
CD34, 10-fold increase	.84	.28-2.50	.75
Grade III-IV acute GVHD			<**.0001**
LEL	1.00		
HEL	1.04	.33-3.25	.95

(*continued*)

Table 3 (*Continued*)

Variables	HR	95% CI	*P* Value
NEM	1.00		
HEL	3.54×10^6	1.00×10^6-12.50×10^6	**<.0001**
LEL	3.42×10^6	1.60×10^6-7.29×10^6	**<.0001**
Graft source			.89
Bone marrow	1.00		
Peripheral blood	1.12	.24-5.27	
Conditioning intensity			.85
RIC	1.00		
MAC	1.09	.42-2.86	
CMV serostatus			.63
Others	1.00		
D-/R-	1.25	.50-3.15	
Disease group			.45
Others	1.00		
AML/ALL	.71	.29-1.73	
KPS score			.61
<90	1.00		
90-100	.80	.34-1.88	
Year of transplantation			.52
2012-2015	1.00		
2009-2011	1.56	.55-4.44	.40
2005-2008	1.95	.53-7.17	.31
Age, 1 yr	1.02	.99-1.06	.23
CD3, 10-fold increase	1.23	.66-2.29	.52
CD34, 10-fold increase	.71	.10-5.25	.74

Significant *P* values are in bold type.
D indicates donor; R, recipient.

the statistical power to detect HCT outcome differences among groups classified by traditional HLA allele matching and direct genetic matching. In support of this approach, a recent report found that ultra-high–resolution HLA matching (which considers non-ARD encoding exons, untranslated regions, and intronic sequence) was associated with improved OS in a cohort of 892 unrelated HCT pairs [17]. Our bioinformatics approach could help facilitate larger HCT cohort studies of the effect of direct HLA genetic matching through the rapid identification of NEM, provided that the HCT pairs have high–quality phased HLA consensus sequences generated by NGS available for analysis.

In summary, this study demonstrates an approach to using the full-length phased sequences produced by NGS to evaluate the degree of genetic similarity within the HLA region between unrelated HCT pairs. We found that a significant proportion (8%) of HCT pairs were highly genetically matched and that a subset of HCT pairs were genetically identical at the sequenced HLA loci despite being unrelated to each other. Furthermore, we showed that this highly HLA genetically matched HCT group had a lower risk of GVHD and equivalent OS compared with the LEL-mismatched group. If these findings are confirmed in a study with a larger sample size and adequate power, evaluating HLA genetic similarity may have clinical ramifications for donor selection. For instance, if multiple unrelated donor options are available, then a low-risk (by Disease Risk Index [18]) patient with multiple medical comorbidities may benefit from a potentially low NRM associated with an NEM donor, if such a donor could be identified among multiple unrelated donors with equivalent 2-field HLA allele-level

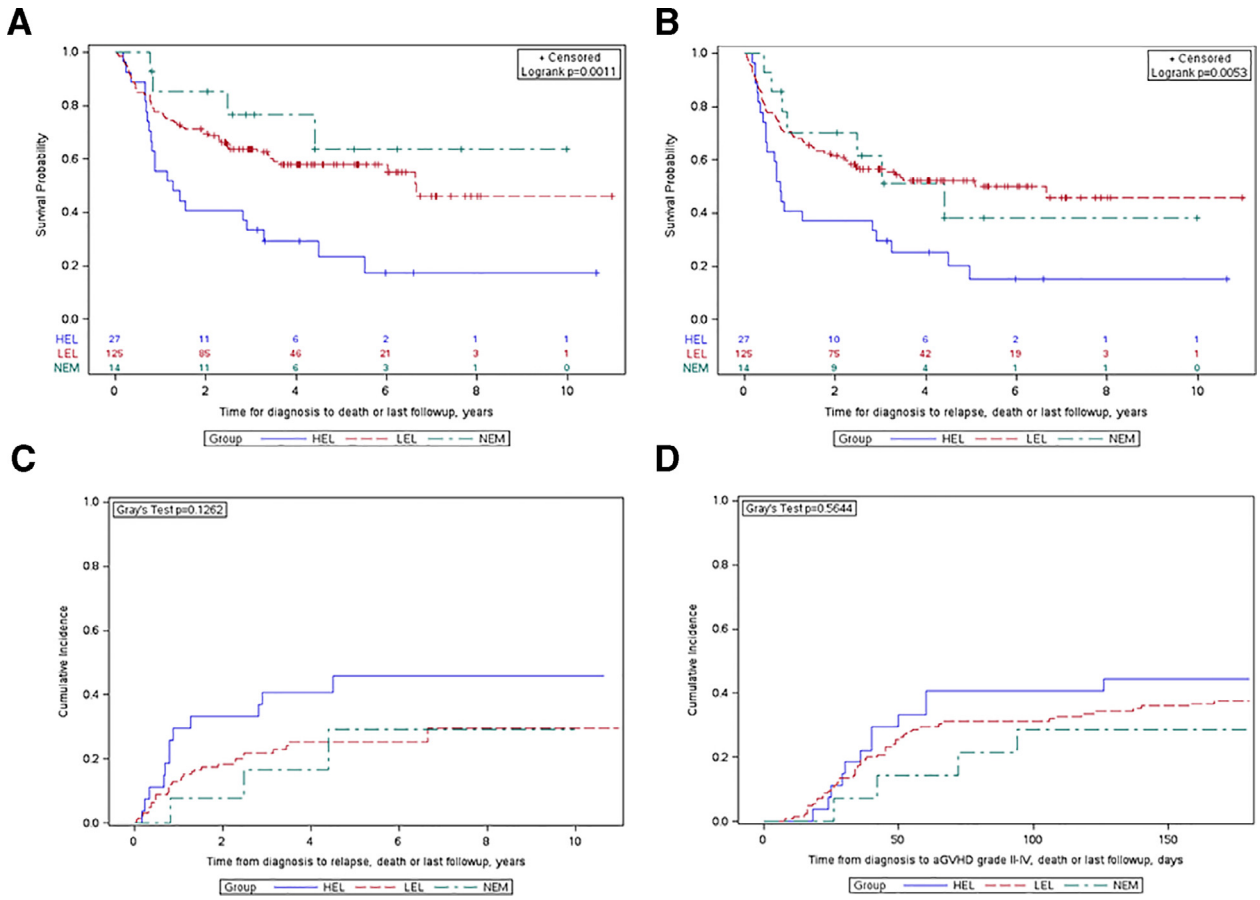

Figure 4. Transplantation outcomes of the HEL-mismatched, LEL-mismatched, and NEM subgroups. (A) Kaplan-Meier plot of OS. (B) Kaplan-Meier plot of PFS. (C) Cumulative incidence plot of NRM. (D) Cumulative incidence plot of grade II-IV acute GVHD.

matching. In contrast, a younger and fit (as reflected by a low HCT Comorbidity Index [19]) patient with high-risk disease (by the Disease Risk Index) could afford to have a multiple LEL-mismatched donor with a higher probability of GVHD and perhaps by the same token, a potentially greater graft-versus-malignancy effect, thereby decreasing the probability of disease relapse and enabling improved survival. In essence, an adequately powered study of unrelated donor–recipient HCT pairs with sufficient follow–up to detect significant differences in the risk of GVHD and survival outcomes between NEM and LEL groups will be critical for the use of direct HLA genetic comparisons for donor selection. Our approach also could help elucidate the effect of genetic similarity as defined by genetic differences between unrelated HCT pairs as opposed to genetic similarity by descent in related sibling HCT pairs. Similar outcome measures between recipients of unrelated donors with NEM and those with HLA-identical sibling donors would further support the use of direct genetic comparison for donor selection in alloHCT.

ACKNOWLEDGMENTS

The authors thank Stephen Spellman, MBS (Principal Investigator, CIBMTR Research Repository) and staff for their assistance in procuring the samples used in the study, and Chao Zhang and Steven Konings (Medical College of Wisconsin) for assistance with database extraction.

This study was presented as a poster at the 2018 BMT Tandem Meetings, Salt Lake City, Utah, February 2018.

Financial disclosure: M.W.A., J.N., and I.V. are employees of Versiti Wisconsin, Inc. V.M. is a former employee of Versiti Wisconsin, Inc. S.C., R.K.N., A.S., J.H.J., B.D., and P.H. are employees of Medical College of Wisconsin, Inc.

Conflict of interest statement: M.W.A. serves on the scientific advisory boards of Omixon and CareDx and has received research funding from Illumina. B.D. serves as a consultant for Celgene and on the advisory boards of Amgen and Takeda Pharmaceutical, and has received research funding from Amgen and Takeda Pharmaceutical. The other authors have no conflicts of interest to report.

Authorship statement: I.V. and S.C. contributed equally to this work as primary co-authors.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found online at doi:10.1016/j.bbmt.2018.12.006.

REFERENCES

1. Marsh SG, Albert ED, Bodmer WF, et al. Nomenclature for factors of the HLA system, 2010. *Tissue Antigens*. 2010;75:291–455.
2. Profaizer T, Lázár-Molnár E, Pole A, Delgado JC, Kumánovics A. HLA genotyping using the Illumina HLA TruSight next-generation sequencing kits: a comparison. *Int J Immunogenet*. 2017;44:164–168.
3. Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SG. The IPD and IMGT/HLA database: allele variant databases. *Nucleic Acids Res*. 2015;43:D423–D431.
4. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30:772–780.
5. Przepiorka D, Weisdorf D, Martin P, et al. 1994 Consensus Conference on Acute GVHD Grading. *Bone Marrow Transplant*. 1995;15:825–828.
6. Lee SJ, Klein J, Haagenson M, et al. High-resolution donor-recipient HLA matching contributes to the success of unrelated donor marrow transplantation. *Blood*. 2007;110:4576–4583.
7. Schöfl G, Lang K, Quenzel P, et al. 2.7 million samples genotyped for HLA by next-generation sequencing: lessons learned. *BMC Genomics*. 2017;18:161.
8. Mack SJ. A gene feature enumeration approach for describing HLA allele polymorphism. *Hum Immunol*. 2015;76:975–981.
9. Hou L, Vierra-Green C, Lazaro A, et al. Limited HLA sequence variation outside of antigen recognition domain exons of 360 10 of 10 matched unrelated hematopoietic stem cell transplant donor-recipient pairs. *HLA*. 2017;89:39–46.
10. Hanvesakul R, Maillere B, Briggs D, Baker R, Larché M, Ball S. Indirect recognition of T-cell epitopes derived from the alpha 3 and transmembrane domain of HLA-A2. *Am J Transplant*. 2007;7:1148–1157.
11. Petersdorf EW, Malkki M, O'hUigin C, et al. High HLA-DP expression and graft-versus-host disease. *N Engl J Med*. 2015;373:599–609.
12. Fernández-Viña MA, Klein JP, Haagenson M, et al. Multiple mismatches at the low-expression HLA loci DP, DQ, and DRB3/4/5 associate with adverse outcomes in hematopoietic stem cell transplantation. *Blood*. 2013;121:4603–4610.
13. Larsen ME, Kornblit B, Larsen MV, et al. Degree of predicted minor histocompatibility antigen mismatch correlates with poorer clinical outcomes in nonmyeloablative allogeneic hematopoietic cell transplantation. *Biol Blood Marrow Transplant*. 2010;16:1370–1381.
14. Robin M, Porcher R, Adès L, et al. Matched unrelated or matched sibling donors result in comparable outcomes after non-myeloablative HSCT in patients with AML or MDS. *Bone Marrow Transplant*. 2013;48:1296–1301.
15. Weisdorf DJ, Nelson G, Lee SJ, et al. Sibling versus unrelated donor allogeneic hematopoietic cell transplantation for chronic myelogenous leukemia: refined HLA matching reveals more graft-versus-host disease but not less relapse. *Biol Blood Marrow Transplant*. 2009;15:1475–1478.
16. Petersdorf EW, Malkki M, Gooley TA, Martin PJ, Guo Z. MHC haplotype matching for unrelated hematopoietic cell transplantation. *PLoS Med*. 2007;4:e8.
17. Mayor NP, Hayhurst JD, Turner TR, et al. Better HLA matching as revealed only by next-generation sequencing technology results in superior overall survival post-allogeneic haematopoietic cell transplantation with unrelated donors. *Biol Blood Marrow Transplant*. 2018;24:S63–S64.
18. Armand P, Kim HT, Logan BR, et al. Validation and refinement of the Disease Risk Index for allogeneic stem cell transplantation. *Blood*. 2014;123:3664–3671.
19. Sorror ML, Maris MB, Storb R, et al. Hematopoietic cell transplantation (HCT)-specific comorbidity index: a new tool for risk assessment before allogeneic HCT. *Blood*. 2005;106:2912–2919.