

TRANSPLANTATION

Genome-wide minor histocompatibility matching as related to the risk of graft-versus-host disease

Paul J. Martin,^{1,2} David M. Levine,³ Barry E. Storer,¹ Edus H. Warren,^{1,2} Xiuwen Zheng,³ Sarah C. Nelson,³ Anajane G. Smith,¹ Bo K. Mortensen,¹ and John A. Hansen^{1,2}

¹Division of Clinical Research, Fred Hutchinson Cancer Research Center, Seattle, WA; ²Department of Medicine, and ³Department of Biostatistics, University of Washington, Seattle, WA

Key Points

- Variable mismatching for minor histocompatibility antigens correlates with the risk of severe GVHD after HCT with sibling donors.
- HLA-mismatching likely explains most of the increased risk of GVHD after HCT with unrelated donors.

The risk of acute graft-versus-host disease (GVHD) is higher after allogeneic hematopoietic cell transplantation (HCT) from unrelated donors as compared with related donors. This difference has been explained by increased recipient mismatching for major histocompatibility antigens or minor histocompatibility antigens. In the current study, we used genome-wide arrays to enumerate single nucleotide polymorphisms (SNPs) that produce graft-versus-host (GVH) amino acid coding differences between recipients and donors. We then tested the hypothesis that higher degrees of genome-wide recipient GVH mismatching correlate with higher risks of GVHD after allogeneic HCT. In HLA-genotypically matched sibling recipients, the average recipient mismatching of coding SNPs was 9.35%. Each 1% increase in genome-wide recipient mismatching was associated with an estimated 20% increase in the hazard of grades III-IV GVHD (hazard ratio [HR], 1.20; 95% confidence interval [CI], 1.05-1.37; $P = .007$) and an estimated 22% increase in the hazard of stage 2-4 acute gut GVHD (HR, 1.22; 95% CI, 1.02-1.45; $P = .03$). In

HLA-A, B, C, DRB1, DQA1, DQB1, DPA1, DPB1-phenotypically matched unrelated recipients, the average recipient mismatching of coding SNPs was 17.3%. The estimated risks of GVHD-related outcomes in HLA-phenotypically matched unrelated recipients were low, relative to the large difference in genome-wide mismatching between the 2 groups. In contrast, the risks of GVHD-related outcomes were higher in HLA-DP GVH-mismatched unrelated recipients than in HLA-matched sibling recipients. Taken together, these results suggest that the increased GVHD risk after unrelated HCT is predominantly an effect of HLA-mismatching. (*Blood*. 2017;129(6):791-798)

Introduction

It has long been recognized that risk of acute graft-versus-host disease (GVHD) after allogeneic hematopoietic cell transplantation (HCT) is higher in patients with unrelated donors than in those with sibling donors. Previous studies have demonstrated that recipient graft-versus-host (GVH) mismatching for HLA-molecules encoded within the major histocompatibility complex (MHC) contributes to the risk of GVHD after unrelated HCT. HLA-genotypically identical sibling donor and recipient pairs that have inherited the same segments of maternal and paternal chromosomes that encode the MHC have identical genetic sequences across the entire region by virtue of identity by descent (IBD).¹ Unrelated donor and recipient pairs selected for HLA-phenotypic identity are identical only for the specific regions within the MHC that encode the specific antigens included within the scope of HLA-typing. Unrecognized mismatching can occur because a specific locus such as HLA-C or HLA-DP was not included in the matching algorithm or because the methods used for typing could not detect all differences between HLA alleles of the recipient and unrelated donor.

At the same time, it has also been recognized that increased recipient mismatching for minor histocompatibility antigens encoded within or outside of the MHC could also explain the higher risk of acute GVHD after unrelated HCT as compared with related HCT.² Minor histocompatibility antigens comprise peptides complexed with MHC molecules for presentation to T cells.^{3,4} Mismatching for minor histocompatibility antigens between the donor and recipient can be generated by a variety of mechanisms related to DNA sequence and structural variation. Amino acid coding differences resulting from single nucleotide polymorphisms (SNP) represent the most common mechanism.^{3,4}

No previous study has attempted to distinguish the extent to which recipient mismatches for minor histocompatibility antigens contribute to the risk of GVHD and GVHD-related complications after HCT. In the current study, we used genome-wide SNP arrays to enumerate GVH amino acid coding differences between recipients and donors for each donor-recipient pair as a surrogate for possible mismatched minor histocompatibility antigens. We then tested the hypothesis that higher degrees of genome-wide recipient GVH mismatching correlate with higher risks of GVHD after allogeneic HCT.

Submitted 31 August 2016; accepted 15 November 2016. Prepublished online as *Blood* First Edition paper, 21 November 2016; DOI 10.1182/blood-2016-09-737700.

The online version of this article contains a data supplement.

There is an Inside *Blood* Commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2017 by The American Society of Hematology

Methods

Patients

The analysis included 3 groups of recipients: (a) those with full-sibling donors who were genotypically matched at HLA-A, B, C, DRB1, DQA1, DQB1, DPA1, and DPB1 (sibling recipients), (b) those with unrelated donors who were phenotypically matched at HLA-A, B, C, DRB1, DQA1, DQB1, DPA1, and DPB1 (HLA 16/16-matched unrelated recipients), and (c) those with unrelated donors who were HLA-phenotypically matched at HLA-A, B, C, DRB1, DQA1, and DQB1 but GVH-mismatched at HLA-DPA1 or DPB1 (HLA-DP GVH-mismatched unrelated recipients).

Assessment of GVHD

Criteria for grading of acute and chronic GVHD have been described previously.^{5,6} Grade II GVHD was classified as IIa or IIb, where IIa indicates stage 1 gut GVHD with or without rash involving $\leq 50\%$ of the body surface and without liver involvement, and IIb indicates rash involving $>50\%$ of the body surface or stage 1 liver GVHD.⁷ For this study, we used the historical definition of clinical extensive chronic GVHD,⁶ because not all patients had chronic GVHD assessed by National Institutes of Health (NIH) criteria.

Genotyping

All recipient and donor samples were collected before transplantation according to research protocols approved by the Fred Hutchinson Cancer Research Center (FHCRC) Institutional Review Board, in accordance with the Declaration of Helsinki. Genomic DNA was extracted from blood mononuclear cells or Epstein-Barr virus-transformed B-cell lines using a Puregene kit (Qiagen, Valencia, CA). Genotyping assays used 3 different platforms: (I) the Affymetrix 5.0 Human GeneChip (cohort I, 1407 recipients and 1413 donors), (II) the Illumina 1M Quad (cohort IIab, 2072 recipients and 2008 donors), and (III) the Illumina 2.5M BeadArray (cohort IIc, 1007 recipients and 1211 donors). Amplification and hybridization for cohort I were performed at the Affymetrix Service Laboratory (Santa Clara, CA), and amplification and hybridization for cohorts IIab and IIc were performed by the FHCRC Genomics Shared Resource laboratory. The genotypes of the candidate SNPs were determined separately for each platform according to the Bayesian Robust Linear Model with Mahalanobis distance classifier-Perfect match algorithm (http://media.affymetrix.com/support/technical/whitepapers/brlmmp_whitepaper.pdf) from within the Affymetrix Power Tools Software Package version 1.15.2 for cohort I and according to GenCall version 6.3.0 from within Illumina GenomeStudio version 1.9.4 for cohorts IIab and IIc (http://illumina.com/documents/products/technotes/technote_gencall_data_analysis_software.pdf).

Quality assurance and quality control

Data cleaning was done separately for each cohort following standard methods.⁸ Briefly, for each cohort, we evaluated batch quality (groups of DNA samples processed together), individual sample quality, sample identity, ancestry differences, systematically missing data, and SNP quality. We identified large chromosomal anomalies (duplications, deletions, and acquired uniparental disomy) that may cause errors affecting genotype calls relative to the germline.⁹ We used a common subset of SNPs genotyped on each platform to estimate the pairwise relatedness of all samples and to identify samples with questionable identity, expected and unexpected duplicate samples, and incorrectly annotated relationships. Samples with questionable identity were removed, and a fully consistent pedigree was constructed for all related samples. The resulting post-QC sample set consists of 4274 recipient and donor pairs, and an additional 212 recipient and 358 donor singletons. Most donors and recipients ($\sim 90\%$) were of European ancestry as identified by principal component analysis.

Imputation

We selected 693 916 autosomal SNPs genotyped via the 2 Illumina arrays in both cohorts IIab and IIc. Only 84 582 of those SNPs were also genotyped via the

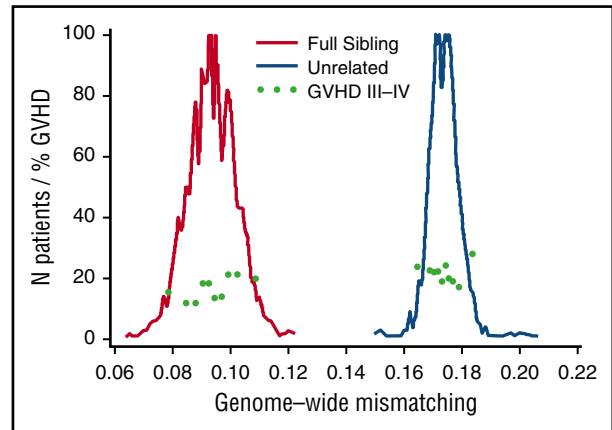


Figure 1. Unrelated donor-recipient pairs have more genome-wide recipient mismatching than sibling pairs. For each coding SNP, recipient mismatching was counted as 0 or 1, where a value of 0 indicates no recipient mismatching and a value of 1 indicates mismatching for 1 or both alleles. The extent of genome-wide mismatching for each recipient was calculated as the proportion of SNPs with mismatching. Histograms show the numbers of patients for each 0.001 increment of genome-wide recipient mismatching for coding SNPs among full-sibling recipients (red, left-side peak) ($N = 1840$) and unrelated recipients (blue, right-side peak) ($N = 1217$). Light green dots show the day-100 cumulative incidence of grades III-IV acute GVHD within deciles across the respective ranges of genome-wide recipient mismatching in sibling and unrelated recipients.

Affymetrix array in cohort I. Therefore, we imputed the SNPs not genotyped in cohort I by using the 1000 Genomes Project phase 3 as a reference panel and the software IMPUTE v2.^{10,11} We also imputed a small number of missing genotype calls for samples in cohorts IIab and IIc. The imputed SNP genotype was retained only if the probability of the most likely genotype exceeded 0.9, and it was otherwise set to missing.

HLA matching and sample selection

The HIBAG software¹² was used to impute the genotypes for HLA-A, B, C, DRB1, DQA1, DQB1, and DPB1 in all samples and HLA-DPA1 in the cohort II samples genotyped on the Illumina platform. HLA-DPA1 genotypes in cohort I were determined by next-generation sequencing (NGS; MiSeq; Illumina System, San Diego, CA) using commercial NGS HLA reagents (Scisco Genetics Inc, Seattle, WA). These HLA alleles had high concordance compared with the prior genotyping methods used for these cohorts. In addition, this approach resolved ambiguous HLA codes, typed previously unidentified alleles, and provided a uniform set of 4-digit allele assignments.

The analysis excluded pairs that had not been graded for GVHD, those whose transplant protocols included treatment with antithymocyte globulin or posttransplant cyclophosphamide, those where either the recipient or donor had a sex chromosome aneuploidy (XXX, XXY, XYY) or Turner's syndrome (XO), related donors who were not full siblings, and any pairs who were mismatched for HLA-A, B, C, DRB1, DQA1, or DQB1 or had ambiguous T-cell epitope assignments (see "HLA-DPB1 T cell epitopes" below). This filtering left 1905 full-sibling pairs and 1217 unrelated pairs for analysis. Of these, 1840 sibling pairs and 194 unrelated pairs were also matched for HLA-DPA1 and HLA-DPB1.

Measurement of genome-wide recipient mismatching across the genome, including the MHC

The GVH vector measures recipient alleles not present in the donor. The GVH vector is defined at a SNP as 0 if the recipient and donor share the same alleles and 1 if the recipient has an allele the donor does not. We defined genome-wide mismatching between recipient and donor as the average of the GVH vector calculated over 19 104 SNPs. These SNPs were selected from the 693 916 autosomal SNPs using IMPUTE2 quality metrics info score >0.8 and certainty score >0.9 (both are 1.0 for genotyped SNPs), a SNP missing call rate $<5\%$, <4

Table 1. Patient characteristics, according to donor type

Characteristic	Sibling recipients* (N = 1840)	Unrelated recipients	
		HLA 16/16–matched† (N = 194)	HLA-DP GVH-mismatched‡ (N = 1023)
Patient age, median years (range)	44 (0–74)	46 (2–75)	44 (0–78)
Chronic myeloid leukemia, N (%)	423 (23)	57 (29)	250 (24)
Female donor for male recipient, N (%)	502 (27)	45 (23)	208 (20)
Mobilized blood cell graft, N (%)	909 (49)	88 (45)	530 (52)
At least 900 cGy total body irradiation, N (%)	691 (38)	95 (49)	477 (47)
Myeloablative conditioning, N (%)	1552 (84)	158 (81)	787 (77)

*HLA-genotypically identical full siblings. Matching includes the entire MHC, including HLA-DPA1 and DPB1.

†Matching includes HLA-A, B, C, DRB1, DQA1, DQB1, DPA1, and DPB1.

‡Matched at HLA-A, B, C, DRB1, DQA1, and DQB1, but GVH-mismatched at HLA-DPA1 or DPB1.

discordances between 214 duplicate sample pairs, no Mendelian errors among parent-offspring pairs, and annotation of the SNP with the term “coding” in the University of California, Santa Cruz known gene table, build hg19. The average pairwise call rate for the 19 104 coding SNPs among donors and recipients in this study was 0.992 (range, 0.809-1.000). Among these 19 104 coding SNPs, we identified a subset of 43 SNPs having alleles known or proposed to encode minor histocompatibility antigens (see supplemental Table 1, available on the *Blood* Web site). The extent of genome-wide mismatching correlated with mismatching for this subset of 43 SNPs (supplemental Figure 1).

HLA-DPB1 T-cell epitopes

We used T-cell epitope (TCE) group definitions according to the method described previously,¹³ but using an updated version of this methodology¹⁴ that classified a wider range of HLA-DPB1 alleles. Even so, 2 HLA-DPB1 alleles, HLA-DPB1*105:01 (previously 06:02) and 107:01 (previously 09:02), were not classified into any TCE group by Crivello et al.¹⁴ However, we noted these 2 alleles have exon 2 sequences identical to those in HLA-DPB1*04:02 and 13:01, respectively,¹⁵ which are classified in the low TCE antigenicity group.¹³ Therefore, we classified HLA-DPB1*105:01 and 107:01 into the low antigenicity group as well.

Statistical analysis

Comparisons between groups of time-to-event endpoints were analyzed by Cox regression. Relapse and death were treated as competing risks for the analysis of GVHD endpoints. Genome-wide mismatching was analyzed as a continuous covariate. For Figure 1, the incidence of acute GVHD was summarized according to deciles of genome-wide mismatching within each donor type.

Results

Table 1 summarizes key GVHD-related risk factors¹⁶ in the related and unrelated HCT cohorts. As expected,² the percentage of recipient coding SNP mismatches was much larger for unrelated donor-recipient pairs than for sibling pairs (Figure 1). Table 2 summarizes the percentages of coding SNPs that differed in recipients as compared with sibling donors, HLA 16/16–matched unrelated donors, and HLA-DP GVH-mismatched unrelated donors (see “Methods” for definitions of the 3 groups). The number of recipient coding SNP mismatches in unrelated recipients was nearly twice the number in related recipients. To the extent that coding SNP mismatches represent possible mismatched minor histocompatibility antigens (supplemental Figure 1), this result suggests that the difference in the number of minor antigen mismatches in unrelated recipients is also nearly twice the number in related recipients.

The distribution of mismatches was notably wider for sibling pairs than for unrelated pairs (Figure 1; Table 2). The wider distribution of recipient mismatching among full-sibling pairs reflects variation in the inheritance of parental chromosomes (Figure 2A) and in the number and size of recombinant sister chromatid segments exchanged between parental chromosomes inherited by siblings during the single generation of meiotic events that connects each child to a parent (Figure 2B). The lower variation in recipient mismatching among unrelated donor transplant pairs (Figure 1; Table 2) occurs as a cumulative result of the multiple generations of inheritance and meiotic recombination events that connect unrelated pairs to their most recent common ancestor.

As a first approach in evaluating the relationship between genome-wide recipient mismatching and the risk of GVHD, we analyzed results separately across the respective ranges of mismatching in related and unrelated recipients. Figure 3A shows the hazard ratios (HRs) of grades II-IV, IIIb-IV, and III-IV acute GVHD, stage 2-4 acute gut GVHD, chronic extensive GVHD, and recurrent or progressive malignancy as related to each 0.01 increase in genome-wide recipient mismatching in sibling recipients. For each 0.01 increase in genome-wide recipient mismatching, results showed an estimated 20% increase in the hazard of grades III-IV GVHD (HR, 1.20; 95% confidence interval [CI], 1.05-1.37; *P* = .007) and an estimated 22% increase in the hazard of stage 2-4 acute gut GVHD (HR, 1.22; 95% CI, 1.02-1.45; *P* = .03). Results for other GVHD-related endpoints showed no statistically significant association with genome-wide recipient mismatching. The left side of Figure 1 shows a weak trend of increasing day-100 cumulative incidence of grades III-IV GVHD across the range of genome-wide recipient mismatching in sibling recipients.

Figure 3B shows results for the same endpoints as related to each 0.01 increase in genome-wide recipient mismatching in unrelated transplant recipients. This analysis was adjusted by classifying each pair as HLA-DPA1 and DPB1–matched, TCE-permissive HLA-DPB1 mismatched, or TCE nonpermissive HLA-DPB1 mismatched. The results showed an estimated 38% increase in the hazard of stages 2-4 acute gut GVHD for each 0.01 increase in genome-wide recipient mismatching (HR, 1.38; 95% CI, 1.03-1.84; *P* = .03). Results for

Table 2. Recipient genome-wide mismatching for coding SNPs, according to recipient type

Recipient type*	Percent genome-wide recipient mismatching, mean (SD)
Siblings	9.35 (0.85)
HLA 16/16–matched unrelated	17.3 (0.56)
HLA-DP GVH-mismatched unrelated	17.4 (0.53)

*See footnote in Table 1 for description of recipient types.

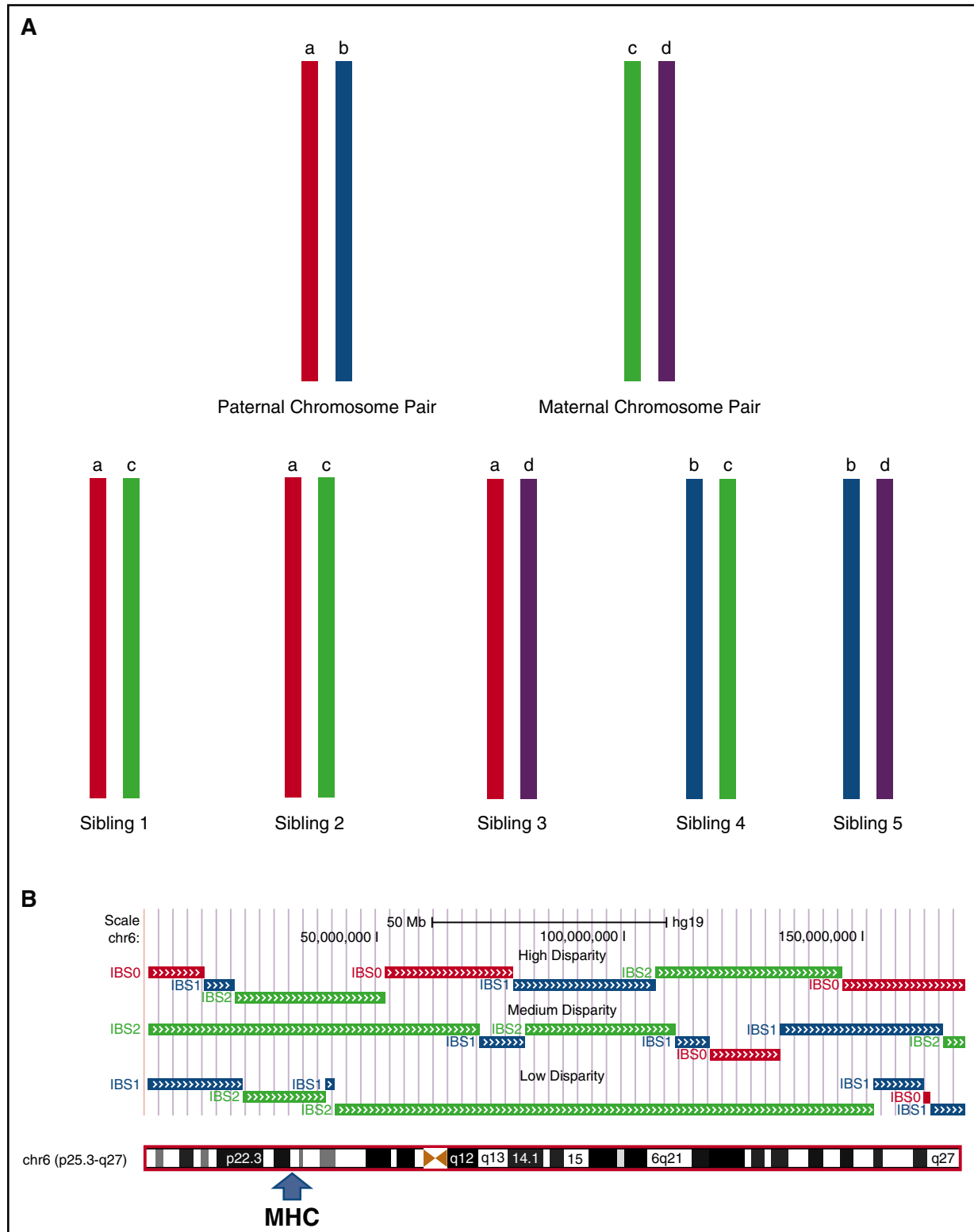
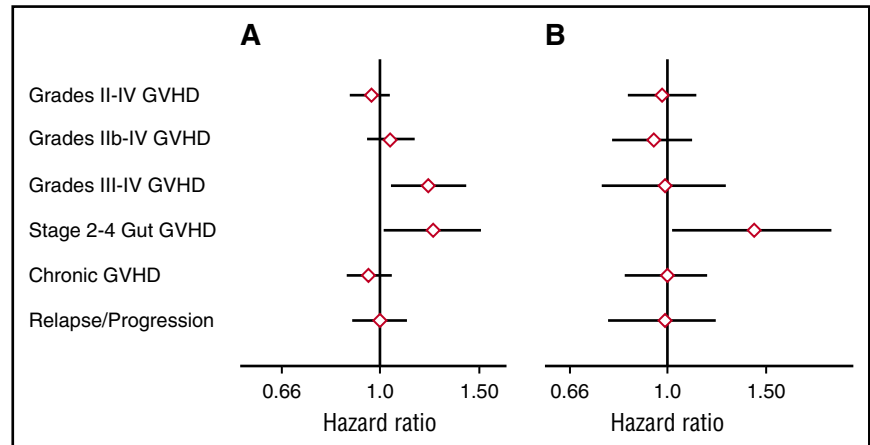


Figure 2. Parental chromosomal segregation and meiotic recombination account for variation in genome-wide recipient mismatching between siblings. (A) Differential inheritance of paternal and maternal chromosomes causes mismatching between sibling pairs. If a pair of paternal chromosomes is designated “a” and “b” and the respective pair of maternal chromosomes is designated “c” and “d,” the chromosomes inherited by the offspring in the absence of meiotic recombination will be “ac,” “bc,” “ad,” or “bd.” Inheritance of the same chromosome pairs by a sibling pair (eg, “ac”) as shown for siblings 1 and 2 is described as IBD with an identity by state (IBS) value of 2, indicating identity for both chromosomes. Inheritance of 1 identical chromosome and 1 nonidentical chromosome by a sibling pair (eg, “ac” and “ad”) is described as partial IBD with an IBS value of 1. Inheritance of 2 nonidentical chromosomes by a sibling pair (eg, “ac” and “bd”) is described as having an IBS value of 0. (B) Meiotic recombination introduces different IBS values across segments of chromosomes inherited by sibling pairs. The figure shows results for 3 sibling pairs with high (top), medium (middle), and low (bottom) genome-wide mismatching across chromosome 6, analyzed as described by Roberson and Pevsner.²⁷ HLA-matching of these pairs is indicated by IBS values of 2 (green) across the MHC region. Genome-wide mismatching increases as the length of segments with IBS values of 2 decreases. Segments with IBS values of 0 (red) and 1 (blue) have different probabilities of recipient mismatching (see supplemental Appendix for details).

Figure 3. Genome-wide recipient mismatching appears to have limited effects on the risks of GVHD-related outcomes. HRs (diamonds) and 95% CIs (lines) show GVHD-related outcomes per 0.01 increment of genome-wide recipient mismatching for coding SNPs among (A) sibling recipients and (B) unrelated recipients, adjusted for HLA-DPB1 T-cell epitope matching. Recipient mismatching was counted as described in Figure 1. Results were similar when recipient mismatching was counted as 0, 1, or 2 for each SNP, corresponding to no mismatching, mismatching for a single allele, or mismatching for both alleles, respectively.



other GVHD-related endpoints showed no statistically significant association with genome-wide recipient mismatching. The statistically significant results in Figure 3 should be interpreted with caution, because no adjustment was made for multiple comparisons.

Given the large difference in genome-wide recipient mismatching between unrelated transplant recipients and sibling transplant recipients, it was of interest to compare the cumulative incidence frequencies of GVHD-related outcomes between the 2 groups, after accounting for the presence of HLA-DP mismatching in unrelated recipients. Figure 4 shows the cumulative incidence frequencies of grades II-IV, IIb-IV, and III-IV acute GVHD, stages 2-4 acute gut GVHD, and clinical extensive chronic GVHD in 3 groups of transplant recipients, those with sibling donors, those with HLA 16/16-matched unrelated donors, and those with HLA-DP GVH-mismatched donors. The cumulative incidence of grades II-IV GVHD was lowest in sibling recipients and was progressively higher in HLA 16/16-matched unrelated recipients and in HLA-DP GVH-mismatched unrelated recipients. Sibling recipients and HLA 16/16-matched unrelated recipients had similar cumulative incidence frequencies of grades IIb-IV and III-IV acute GVHD, stages 2-4 gut GVHD, and chronic GVHD, whereas HLA-DP GVH-mismatched unrelated recipients had higher cumulative incidence frequencies of these events compared with the other 2 groups.

Figure 5A shows the HRs of grades II-IV, IIb-IV, and III-IV acute GVHD, stage 2-4 acute gut GVHD, chronic extensive GVHD, and recurrent or progressive malignancy in the 194 HLA 16/16-matched unrelated recipients as compared with the 1840 sibling recipients. Consistent with Figure 4, despite the far greater degree of genome-wide recipient mismatching between HLA 16/16-matched unrelated recipients and sibling recipients, no statistically significant differences were observed in the risks of the GVHD endpoints, but the risk of recurrent or progressive malignancy was higher in the unrelated recipients than in the related recipients (HR, 1.39; 95% CI, 1.1-1.8; $P = .006$). This analysis was limited by the smaller number of HLA 16/16-matched unrelated recipients, resulting in wide CIs. In contrast, the risks of all GVHD-related outcomes were higher in HLA-DP GVH-mismatched unrelated recipients than in sibling recipients (Figure 5B), whereas the risk of recurrent or progressive malignancy was lower in HLA-DP GVH-mismatched unrelated recipients than in sibling recipients. Results of analyses testing the association of HLA-DP permissive and nonpermissive mismatching with the risks of GVHD and recurrent or progressive malignancy in this cohort were consistent with those reported in previous studies^{13,14} (supplemental Table 2).

Discussion

Taken together, results of this study suggest that although a greater degree of genome-wide recipient mismatching contributes to an increased GVHD risk after HCT from HLA-matched sibling donors, the increased GVHD risk after unrelated HCT is predominantly an effect of HLA-mismatching. Our cohort had only 194 HLA 16/16-matched unrelated recipients, and event frequencies for some outcomes were low. Because of these limitations, the upper 95% confidence limit of the adjusted HRs comparing GVHD-related outcomes in the 194 HLA 16/16-matched unrelated recipients compared with sibling recipients ranged from 1.25 to 1.62 (Figure 5A). With this wide range, we cannot exclude the possibility that genome-wide recipient mismatching also contributes to the increased risk of GVHD after unrelated HCT. As expected, HLA-DP mismatching was associated with a lower risk of recurrent or progressive malignancy. We did not expect to find a higher risk of recurrent or progressive malignancy in HLA 16/16-matched unrelated recipients compared with sibling recipients, and we do not have any simple explanation for this observation.

The analysis of coding SNP variation between recipients and donors has limitations as a surrogate in estimating the extent of mismatching for minor histocompatibility antigens. First, presentation of a minor histocompatibility antigen requires appropriate proteasomal cleavage of the source protein and binding of a polymorphic peptide cleavage product in an HLA molecule expressed by the recipient, in addition to a difference in the amino acid sequence of the peptide product in the recipient as compared with the donor.^{3,4} The extent to which polymorphic peptides can be presented by the HLA molecules expressed in any given individual at an epitope density that is sufficient to elicit an alloimmune response is unknown. In silico efforts to identify minor antigens that cause GVHD could yield potentially interesting results. Several steps in MHC class I peptide processing are not modeled with available algorithms, and identification of minor histocompatibility antigens with the use of prediction tools is a daunting task fraught with high false discovery rates.¹⁷ Moreover, tools for identifying minor histocompatibility antigens presented by MHC class II molecules are less well established than those for MHC class I molecules.¹⁸ Recent work, however, has shown that proteogenomic-based methods can overcome many of these limitations.¹⁷

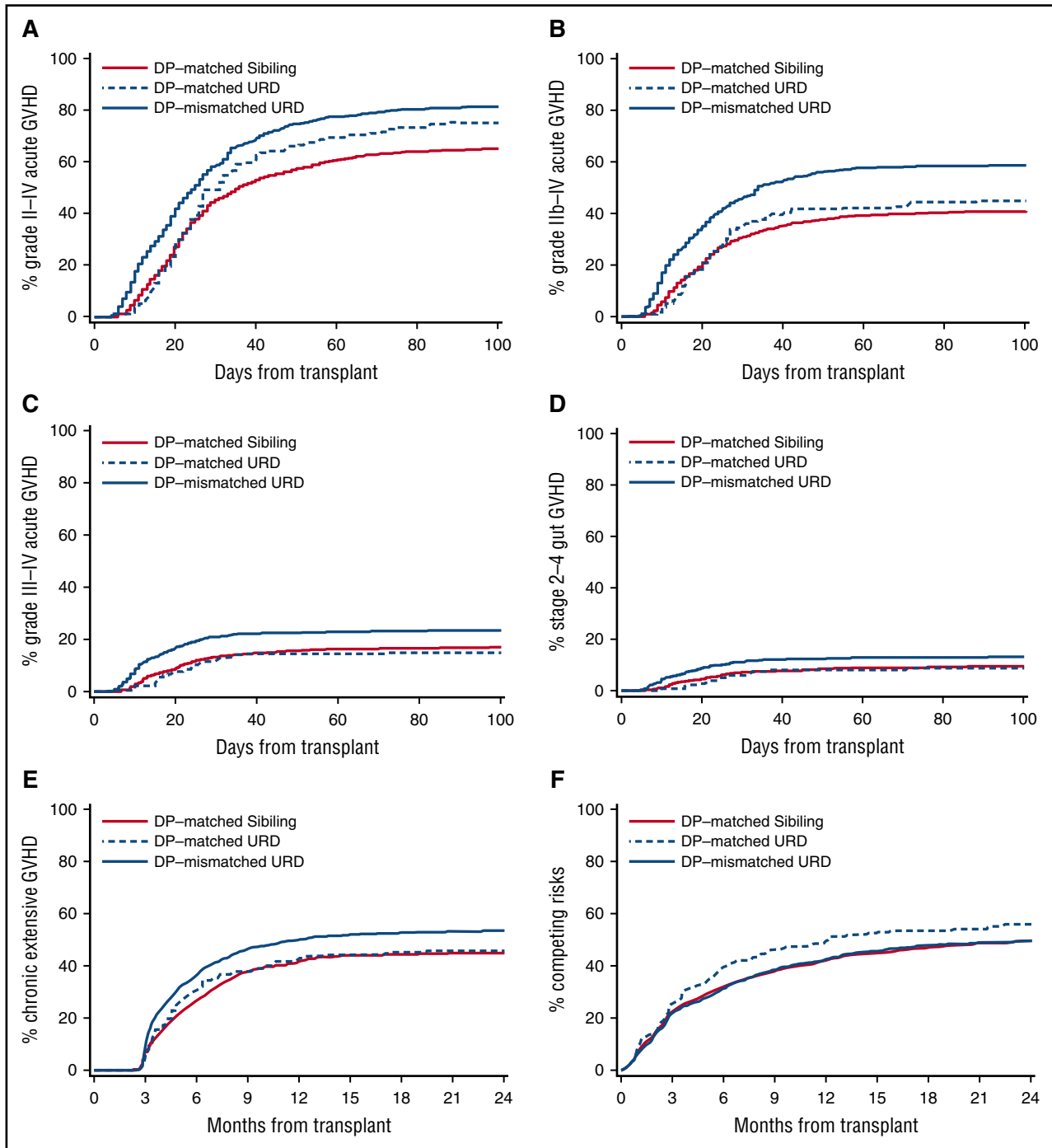
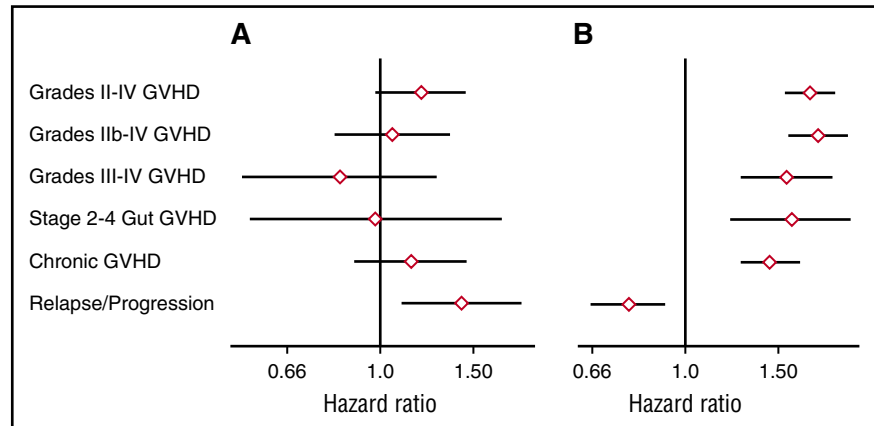


Figure 4. The increased GVHD risk after unrelated HCT is predominantly an effect of HLA-mismatching. Panels compare the cumulative incidence frequencies of GVHD-related outcomes for sibling recipients, HLA 16/16-matched unrelated recipients, and HLA-DP GVH-mismatched unrelated recipients. Results are shown for (A) grades II-IV GVHD, (B) grades IIb-IV GVHD (see “Methods”), (C) grades III-IV GVHD, (D) stages 2-4 gut GVHD, (E) extensive chronic GVHD, and (F) the competing risks of nonrelapse mortality or recurrent or progressive malignancy. The higher incidence of competing risks in the HLA 16/16-matched unrelated recipients is entirely attributable to a higher incidence of recurrent or progressive malignancy (see Figure 5A). DP, HLA-DP; URD, unrelated donor.

The arrays used for our study do not encompass all coding SNPs, and coding SNP variation accounts for only 1 of the mechanisms that can give rise to minor histocompatibility antigens. Other mechanisms include gene insertions and deletions, RNA splice variation, stop codon mutations, differences between Y and X chromosome homologs, and posttranslational modifications of amino acids. Hidden noncanonical open reading frames located outside exomes can also give rise to peptides that form complexes

with MHC molecules for presentation as minor histocompatibility antigens.^{3,4} Within the 19 104 coding SNPs used for our analysis, we identified an average of ~1800 coding SNP mismatches between HLA-identical siblings. On the other hand, it has been estimated that HLA-matched sibling pairs have <100 minor antigen mismatches.¹⁷ Therefore, the extent of mismatching for coding SNPs greatly exceeds the estimated extent of mismatching for minor histocompatibility antigens.^{17,19}

Figure 5. HLA-DP-mismatching has a large effect on the risks of GVHD-related outcomes. HRs (diamonds) and 95% CIs (lines) show GVHD-related outcomes for (A) HLA 16/16-matched unrelated recipients and (B) HLA-DP GVH-mismatched unrelated recipients as compared with sibling recipients. HRs for GVHD-related endpoints were adjusted for the risk factors listed in Table 1, and the HRs for recurrent or progressive malignancy were adjusted for low-risk disease (chronic myeloid leukemia in chronic phase and refractory anemia), high-risk disease (malignancy not in remission), and intermediate-risk disease (all others).



Our results must be interpreted with the understanding that the risk of GVHD depends not only on genetic risk factors but also on certain clinical risk factors and, most importantly, on the intensity and effectiveness of immunosuppressive treatment after HCT. Previous studies have shown that with no posttransplant immunosuppression, all patients develop severe GVHD,²⁰ and with a perfect immunosuppressive regimen (yet to be developed), no patient would develop severe GVHD. Results of the current study are applicable to patients who received T-replete grafts and well-established conventional immunosuppressive regimens after HCT. We excluded patients whose transplant protocols included treatment with antithymocyte globulin or posttransplant high-dose cyclophosphamide. The analysis was adjusted for other factors known to influence the risk of GVHD.¹⁶

Our analysis of genome-wide recipient mismatching might appear to have some similarity to previous studies of HY mismatching (HCT with female donors for male recipients)^{3,4} and UGT2B17 deletion mismatching²¹ in estimating the association of multiple minor histocompatibility antigens with GVHD-related outcomes. The current analysis of genome-wide mismatching did not include polymorphic or nonpolymorphic coding nucleotides in X or Y chromosomes. Mismatching effects attributed to HY and UGT2B17 are independent of outcomes associated with mismatching summarized across the entire genome, and the inability to detect a genome-wide association of minor antigen mismatching with grades II-IV GVHD or chronic GVHD in the current study should not be interpreted as showing that minor antigen mismatching does not influence these outcomes.

Our results suggest that the increased risk of GVHD-related outcomes after unrelated HCT is explained largely by mismatching at HLA-DPA1, HLA-DPB1, or other loci in the MHC.²² Based on the observed difference in the number of coding SNP mismatches, we would estimate that the number of minor antigen differences is approximately twofold higher between HLA-matched unrelated pairs than between HLA-matched related pairs. The much greater immune response against a single HLA-mismatch reflects the number of T cells that recognizes epitopes of the MHC molecule itself and the wide variety of peptides that can be presented by any given MHC allotype.²³ The hypothesis that the increased risk of GVHD-related outcomes is explained by MHC-mismatching could gain support if data were available to show that the risk of GVHD-related outcomes is higher in HLA-DP GVH-mismatched siblings as compared with HLA-matched siblings. Alongside the 1840 donor-recipient sibling pairs examined in the current study, we identified 57 pairs with HLA-DP GVH-mismatching. The number of HLA-DP GVH-mismatched sibling pairs

and the event rates in our cohort do not provide sufficient statistical power for an informative analysis.

The limited effects of genome-wide recipient mismatching across the range observed among siblings and the absence of strikingly increased risks of GVHD-related outcomes despite large differences in genome-wide recipient mismatching between related and unrelated donor-recipient pairs suggest that the *in vivo* alloimmune response against multiple minor antigens is not proportional to the degree of mismatching. This phenomenon might be explained by immunodominance, wherein only a small number of antigens that could elicit an immune response *in vivo* actually do so, whereas the rest do not contribute to the overall response.²⁴ To the extent that the increased GVHD risk after unrelated HCT is predominantly an effect of HLA-mismatching rather than minor antigen mismatching, we would predict that studies with larger numbers of patients would confirm our findings that GVHD risks in optimally HLA-matched unrelated recipients are comparable to those observed with HLA-matched sibling recipients. These results emphasize the value and importance of comprehensive, high-resolution HLA-typing and matching of donors and recipients in unrelated HCT.

Although our study focused on the relative contributions of genome-wide and HLA-DP mismatching to acute and chronic GVHD, donor-recipient disparity also provides the foundation for the therapeutic efficacy of allogeneic HCT for treatment of malignant diseases. Donor T-cell responses against recipient major and minor alloantigens help to eliminate malignant cells surviving in the recipient after the pretransplant conditioning regimen. It is unlikely that matching for minor histocompatibility antigens could be used to decrease the risk of GVHD. Nonetheless, the enormous amount of coding genetic variation that distinguishes any related or unrelated donor-recipient pair offers a rich repertoire of potential minor histocompatibility antigens that could be effectively targeted with vaccine or adoptive transfer approaches.^{25,26}

Observations in this study raise some perplexing questions that remain to be answered. For example, why is the effect of multiple minor disparities limited to the risk of severe acute GVHD, particularly involving the gastrointestinal tract, without affecting the risk of mild (grade II) GVHD, chronic GVHD, or recurrent or progressive malignancy? More provocatively, how is it that the alloimmune response against multiple minor antigens has limits, whereas the response against major histocompatibility antigens does not? As a corollary, why does HLA-mismatching have an effect on the risk of chronic GVHD, whereas the number of minor disparities appears not to have such an effect in siblings? Similar questions were raised by a

previous study indicating that recipient HLA mismatching had a greater effect on the risk of acute GVHD than on chronic GVHD, whereas the use of female donors for male recipients had a greater effect on the risk of chronic GVHD than on acute GVHD.¹⁶ The answers to these questions might lie in differences related to the biology of antigen presentation, T-cell recognition and T-cell activation, the regulation of immune responses against minor and major histocompatibility antigens, and the respective mechanisms leading to acute and chronic GVHD and the immunologic activity of donor cells against malignant cells in the recipient.

Acknowledgments

The authors thank the participants who contributed specimens for this study and especially Jenna Gravley for outstanding assistance in maintaining the biospecimen repository at the FHCRC and verifying the identity of samples used in this study, and the FHCRC Genomics Shared Resource Laboratory (Jeff Delrow, Director) for the genotyping of cohort II.

This work was supported by grants R01-HL105914 and R01-HL87690 from the National Institutes of Health (NIH), National

Heart, Lung, and Blood Institute (J.A.H.), and Cancer Center Support Grant P30-CA015704 from the NIH, National Cancer Institute, Department of Health and Human Services.

Authorship

Contribution: P.J.M., J.A.H., D.M.L., B.E.S., and E.H.W. designed the study, interpreted results, and wrote the manuscript; P.J.M. graded acute GVHD severity; D.M.L., X.Z., and S.C.N. provided genotyping and imputation, data quality control, and coding; A.G.S. interpreted HLA-DP typing; B.K.M. curated minor histocompatibility antigens; and B.E.S. provided statistical analysis. All authors critically revised the manuscript for important intellectual content and approved the manuscript for publication.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: P.J.M., 0000-0001-9051-1215.

Correspondence: Paul J. Martin, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, D2-100, P.O. Box 19024, Seattle, WA 98109-1024; e-mail: pmartin@fredhutch.org.

References

- Weir BS, Anderson AD, Hepler AB. Genetic relatedness analysis: modern data and new challenges. *Nat Rev Genet*. 2006;7(10):771-780.
- Martin PJ. Increased disparity for minor histocompatibility antigens as a potential cause of increased GVHD risk in marrow transplantation from unrelated donors compared with related donors. *Bone Marrow Transplant*. 1991;8(3):217-223.
- Warren EH, Zhang XC, Li S, et al. Effect of MHC and non-MHC donor/recipient genetic disparity on the outcome of allogeneic HCT. *Blood*. 2012;120(14):2796-2806.
- Mullally A, Ritz J. Beyond HLA: the significance of genomic variation for allogeneic hematopoietic stem cell transplantation. *Blood*. 2007;109(4):1355-1362.
- Leisenring WM, Martin PJ, Petersdorf EW, et al. An acute graft-versus-host disease activity index to predict survival after hematopoietic cell transplantation with myeloablative conditioning regimens. *Blood*. 2006;108(2):749-755.
- Lee SJ, Vogelsang G, Flowers ME. Chronic graft-versus-host disease. *Biol Blood Marrow Transplant*. 2003;9(4):215-233.
- Mielcarek M, Storer BE, Boeckh M, et al. Initial therapy of acute graft-versus-host disease with low-dose prednisone does not compromise patient outcomes. *Blood*. 2009;113(13):2888-2894.
- Laurie CC, Doheny KF, Mirel DB, et al; GENEVA Investigators. Quality control and quality assurance in genotypic data for genome-wide association studies. *Genet Epidemiol*. 2010;34(6):591-602.
- Laurie CC, Laurie CA, Rice K, et al. Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat Genet*. 2012;44(6):642-650.
- Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*. 2009;5(6):e1000529.
- Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3 (Bethesda)*. 2011;1(6):457-470.
- Zheng X, Shen J, Cox C, et al. HIBAG—HLA genotype imputation with attribute bagging. *Pharmacogenomics J*. 2014;14(2):192-200.
- Fleischhauer K, Shaw BE, Gooley T, et al; International Histocompatibility Working Group in Hematopoietic Cell Transplantation. Effect of T-cell-epitope matching at HLA-DPB1 in recipients of unrelated-donor haemopoietic-cell transplantation: a retrospective study. [Erratum appears in *Lancet Oncol*. 2012 Apr;13(4):e134-5] *Lancet Oncol*. 2012;13(4):366-374.
- Crivello P, Zito L, Sizzano F, et al. The impact of amino acid variability on alloreactivity defines a functional distance predictive of permissive HLA-DPB1 mismatches in hematopoietic stem cell transplantation. *Biol Blood Marrow Transplant*. 2015;21(2):233-241.
- Reinders J, Rozemuller EH, van Gent R, Arts-Hilkes YH, van den Tweel JG, Tilanus MG. Extended HLA-DPB1 polymorphism: an RNA approach for HLA-DPB1 typing. *Immunogenetics*. 2005;57(10):790-794.
- Flowers ME, Inamoto Y, Carpenter PA, et al. Comparative analysis of risk factors for acute graft-versus-host disease and for chronic graft-versus-host disease according to National Institutes of Health consensus criteria. *Blood*. 2011;117(11):3214-3219.
- Granados DP, Sriranganadane D, Daouda T, et al. Impact of genomic polymorphisms on the repertoire of human MHC class I-associated peptides. *Nat Commun*. 2014;5:3600. doi: 10.1038/ncomms4600.
- Nielsen M, Lund O, Buus S, Lundegaard C. MHC class II epitope predictive algorithms. *Immunology*. 2010;130(3):319-328.
- Ofran Y, Kim HT, Brusci V, et al. Diverse patterns of T-cell response against multiple newly identified human Y chromosome-encoded minor histocompatibility epitopes. *Clin Cancer Res*. 2010;16(5):1642-1651.
- Sullivan KM, Deeg HJ, Sanders J, et al. Hyperacute graft-v-host disease in patients not given immunosuppression after allogeneic marrow transplantation. *Blood*. 1986;67(4):1172-1175.
- McCarroll SA, Bradner JE, Turpeinen H, et al. Donor-recipient mismatch for common gene deletion polymorphisms in graft-versus-host disease. *Nat Genet*. 2009;41(12):1341-1344.
- Carapito R, Jung N, Kwemou M, et al. Matching for the nonconventional MHC-I MICA gene significantly reduces the incidence of acute and chronic GVHD. *Blood*. 2016;128(15):1979-1986.
- Felix NJ, Allen PM. Specificity of T-cell alloreactivity. *Nat Rev Immunol*. 2007;7(12):942-953.
- Korngold R, Wettstein PJ. Immunodominance in the graft-vs-host disease T cell response to minor histocompatibility antigens. *J Immunol*. 1990;145(12):4079-4088.
- Bleakley M, Riddell SR. Exploiting T cells specific for human minor histocompatibility antigens for therapy of leukemia. *Immunol Cell Biol*. 2011;89(3):396-407.
- Granados DP, Rodenbrock A, Laverdure JP, et al. Proteogenomic-based discovery of minor histocompatibility antigens with suitable features for immunotherapy of hematologic cancers. *Leukemia*. 2016;30(6):1344-1354.
- Roberson ED, Pevsner J. Visualization of shared genomic regions and meiotic recombination in high-density SNP data. *PLoS One*. 2009;4(8):e6711. [Electronic Resource]