

# Recipient and donor genetic variants associated with mortality after allogeneic hematopoietic cell transplantation

Paul J. Martin,<sup>1,2</sup> David M. Levine,<sup>3</sup> Barry E. Storer,<sup>1</sup> Sarah C. Nelson,<sup>3</sup> Xinyuan Dong,<sup>3</sup> and John A. Hansen<sup>1,2</sup>

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

## Key Points

- Donor MICA-129Met/Val alleles alter risks of GVHD and mortality after allogeneic HCT.
- Other previously reported genetic associations with mortality after allogeneic HCT were not validated.

Many studies have suggested that genetic variants in donors and recipients are associated with survival-related outcomes after allogeneic hematopoietic cell transplantation (HCT), but these results have not been confirmed. Therefore, the utility of testing genetic variants in donors and recipients for risk stratification or understanding mechanisms leading to mortality after HCT has not been established. We tested 122 recipient and donor candidate variants for association with nonrelapse mortality (NRM) and relapse mortality (RM) in a cohort of 2560 HCT recipients of European ancestry with related or unrelated donors. Associations discovered in this cohort were tested for replication in a separate cohort of 1710 HCT recipients. We found that the donor rs1051792 A allele in *MICA* was associated with a lower risk of NRM. Donor and recipient rs1051792 genotypes were highly correlated, making it statistically impossible to determine whether the donor or recipient genotype accounted for the association. Risks of grade 3 to 4 graft-versus-host disease (GVHD) and NRM in patients with grades 3 to 4 GVHD were lower with donor MICA-129Met but not with MICA-129Val, implicating MICA-129Met in the donor as an explanation for the decreased risk of NRM after HCT. Our analysis of candidate variants did not show any other association with NRM or RM. A genome-wide association study did not identify any other variants associated with NRM or RM.

## Introduction

Many studies have suggested that genetic variants in donors and recipients are associated with the overall survival (OS), progression-free survival (PFS), nonrelapse mortality (NRM), or relapse mortality (RM) of patients after allogeneic hematopoietic cell transplantation (HCT). The associations identified in these studies, however, were not confirmed in a large independent HCT cohort of unrelated donors and recipients from the Center for International Blood and Marrow Transplant Research (CIBMTR).<sup>1,2</sup> The CIBMTR studies of candidate single-nucleotide polymorphisms (SNPs)<sup>1</sup> and genome-wide exome polymorphisms<sup>2</sup> were limited to patients with unrelated donors, and the genome-wide testing did not include nonexonic polymorphisms. Moreover, our review of the updated and previously published literature identified 73 variants that were not included in the CIBMTR studies. Consequently, the utility of testing genetic variation in donors and recipients for risk stratification or understanding mechanisms leading to mortality after HCT has not been established.

We took a twofold approach to address gaps in the CIBMTR studies. First, from a review of the literature through December 2019, we compiled a list of 130 variants, associated with 79 genes, reported as individual SNPs, SNP haplotypes, SNP combinations, or indels having statistically significant associations with mortality after HCT. Forty-seven of these variants were included in the CIBMTR

Submitted 23 March 2020; accepted 14 June 2020; published online 20 July 2020.  
DOI 10.1182/bloodadvances.2020001927.

Deidentified individual participant data are available indefinitely in the National Center for Biotechnology Information (accession #phs001918).

The full-text version of this article contains a data supplement.  
© 2020 by The American Society of Hematology

analysis,<sup>1</sup> which encompassed reports published through 2016. Sixty-eight other variants were identified in reports published before 2017 but were not included in the CIBMTR analysis, and 15 were identified in reports published after 2016. We tested 122 of these candidate recipient and donor variants for associations with NRM (ie, death before recurrent or progressive malignancy) and RM (ie, death after recurrent or progressive malignancy) in a cohort of 2560 HCT recipients of European ancestry with related or unrelated donors at our center. Associations discovered in this cohort were tested for replication in a separate cohort of 1710 HCT recipients. Second, we conducted a genome-wide association study (GWAS) to identify variants associated with NRM and RM using the same discovery and replication approach.

## Methods

### Study population

All recipient and donor blood samples were collected before HCT according to research protocols approved by the Fred Hutchinson Cancer Research Center (FHCRC) Institutional Review Board. Project-specific institutional review board approval was obtained for the use of clinical data and research biospecimens.

Table 1 summarizes demographic, clinical, and transplantation characteristics of patients in the study cohort. The overall study cohort included 4270 recipients of European ancestry who had a first allogeneic HCT at the FHCRC or Seattle Cancer Care Alliance from 1990 through 2011. Syngeneic or cord blood donors were excluded. A single prior autologous HCT was allowed. The study was limited to recipients of European ancestry because the number of available non-European recipients was too small for a meaningful analysis in this study. European ancestry was defined by using the minimum covariant determinant method as implemented by Conomos et al.<sup>3</sup> HLA-A, -B, -C, -DRB1, or -DQB1 mismatching was present in 7% of the related donor-recipient pairs and in 37% of the unrelated pairs. Indications for HCT included hematological malignancy or myelodysplasia. Conditioning regimens were categorized as myeloablative or nonmyeloablative according to the intensity of chemotherapy and total-body irradiation. Grafts used for HCT included bone marrow cells or growth factor–mobilized blood cells. Supplemental Figure 1 shows the cumulative incidence frequencies of NRM and RM in the study cohort.

### Literature search

We performed a PubMed search to identify studies that reported a recipient or donor SNP, SNP haplotype, or SNP combination associated with OS, NRM, or PFS after allogeneic HCT at a statistical significance level of  $P \leq .05$ . Other genetic variants such as tandem repeats associated with these outcomes were excluded because the genotyping arrays used for our study are not informative for such variants. Supplemental Table 1 shows the candidate variants and reported phenotypes, with supporting references.

### Sample preparation and genotyping

Details regarding sample preparation have been described previously.<sup>4</sup> Donor samples were not available for 214 pairs (2%), and recipient samples were not available for 370 pairs (4%). Genomic DNA was extracted from blood mononuclear cells or Epstein-Barr virus–transformed B-lymphocyte cell lines using a Puregene kit

**Table 1. Characteristics of the study cohort (N = 4270)**

Characteristic	n (%)
Recipients genotyped	3850 (90)
Donors genotyped	3977 (93)
<b>Patient age at transplantation, y</b>	
Median	43
Range	0-78
<b>Diagnosis</b>	
Acute leukemia	1761 (41)
Chronic myeloid leukemia	1032 (24)
Myelodysplastic syndrome or myeloproliferative neoplasm	736 (17)
Chronic lymphocytic leukemia	116 (3)
Malignant lymphoma or multiple myeloma	625 (15)
<b>Disease risk*</b>	
Low	936 (22)
Intermediate	1199 (28)
High	1881 (44)
Not classified	254 (6)
<b>Donor-recipient sex combination</b>	
Male to male	1438 (34)
Male to female	953 (22)
Female to male	1038 (24)
Female to female	839 (20)
<b>Donor type</b>	
Related HLA-A, -B, -C, -DRB1, or -DQB1 matched	1855 (43)
HLA-DP mismatched†	84 (5)
Related HLA-A, -B, -C, -DRB1, or -DQB1 mismatched	130 (3)
HLA-DP mismatched†	67 (52)
Unrelated HLA-A, -B, -C, -DRB1, or -DQB1 matched	1447 (34)
HLA-DP mismatched†	1117 (77)
Unrelated HLA-A, -B, -C, -DRB1, or -DQB1 mismatched	838 (20)
HLA-DP mismatched†	661 (79)
<b>Graft source</b>	
Bone marrow	2325 (54)
Mobilized blood cells	1945 (46)
T cell–depleted graft	31 (1)
<b>Conditioning regimen</b>	
Myeloablative, <900 cGy TBI	1515 (35)
Myeloablative, ≥900 cGy TBI	2021 (47)
Nonmyeloablative	734 (17)
Antithymocyte globulin	55 (1)
<b>Posttransplantation immunosuppression</b>	
Cyclosporine or tacrolimus	4139 (97)
Methotrexate or mycophenolate mofetil	3926 (92)
Cyclophosphamide	40 (1)

TBI, total-body irradiation.

\*Low risk is chronic myeloid leukemia in chronic phase or myelodysplastic syndrome–refractory anemia; intermediate risk, acute leukemia, chronic lymphocytic leukemia, or non-Hodgkin lymphoma in remission; high risk, all others.

†No. and percentage of the row above.

(Qiagen, Valencia, CA). Genotyping assays used 3 different platforms: the Affymetrix 5.0 Human GeneChip (cohort 1; 1048 recipients and 1029 donors), the Illumina 1M Quad (cohort 2a; 1904 recipients and 1828 donors), and the Illumina 2.5M BeadArray (cohort 2c; 688 recipients and 926 donors). Amplification and hybridization for the Affymetrix 5.0 array were performed at the Affymetrix Service Laboratory (Santa Clara, CA), and amplification and hybridization for the Illumina BeadArrays were performed by the FHCRC Genomics Shared Resource Laboratory. The genotypes of the candidate variants were determined separately for each platform using the BRLMM algorithm for the Affymetrix array<sup>5</sup> and the GeneCall algorithm for the Illumina arrays.<sup>6</sup>

## Quality assurance and quality control

Quality assurance and quality control (QC) were primarily performed separately for variants and samples from each cohort following standard methods.<sup>7</sup> Briefly, for each cohort, we evaluated batch quality (groups of DNA samples processed together), individual sample quality, ancestry differences, and missing data. We identified large chromosomal anomalies (duplications, deletions, and acquired uniparental disomy) that may cause errors affecting genotype calls relative to the germ line. For the variants in each cohort, we evaluated the missing call rate, Hardy-Weinberg equilibrium, Mendelian errors, heterozygosity, and allele frequency differences between males and females and discordant genotypes between duplicate sample pairs. We used a common subset of variants genotyped on each platform to estimate the pairwise relatedness of all samples and 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. This effort was further facilitated by using the SNP genotypes to impute the HLA type of each sample for comparison with the clinical typing records.

After QC, cohort 1 had 2813 samples and 390 412 variants, cohort 2a had 4072 samples and 1 003 751 variants, and cohort 2c had 2215 samples and 2 327 361 variants. The resulting post-QC sample set contained a total of 3850 recipients and 3977 donors.

## HLA matching

For each sample, we calculated a uniform set of 4-digit allele assignments for HLA-A, -B, -C, -DRB1, and -DQB1 using HIBAG software<sup>8</sup> to impute the HLA alleles from the SNP genotypes. In general, the imputed alleles had high concordance with the prior genotyping methods used for the clinical typing. A human expert reviewed and resolved all discrepancies. This approach also resolved ambiguous HLA codes, typed previously unidentified alleles, and provided a uniform set of 4-digit allele assignments.

## Imputation

We used SHAPEIT2<sup>9</sup> and IMPUTE2<sup>10</sup> software to separately prephase and impute, respectively, the samples in each cohort to a common set of variants from the 1000 Genomes Project phase 3.<sup>11</sup> We also imputed a small number of sporadic missing genotype calls. We limited the list of imputed variants to those with at least 4 minor allele copies in any of the 5 1000 Genomes Project superpopulations (AFR, AMR, EAS, EUR, SAS). The posterior probability of the most likely genotype was calculated as the probability of observing an unobserved genotype at the imputed

locus, given all observed genotypes in the flanking region. The imputed SNP genotype was retained only if the posterior probability of the most likely genotype exceeded 0.9, and it was otherwise set to missing.

To harmonize the genotyped and imputed variants and samples from all 3 genotyping platforms into a single integrated data set, we ensured, for each variant, that the genomic location (chromosome and position) and plus-strand alleles<sup>12</sup> matched across platforms and aligned to genome build GRCh37/hg19. We used IMPUTE2 metrics (information score and average posterior probability) to retain only well-imputed variants and our own measures (Mendelian errors between sibling pairs, missing call rate, and duplicate genotype discordance across cohorts) to further remove low-quality variants and/or those exhibiting batch effects between cohorts. The result was a data set with 9100 samples and 29 826 485 post-QC variants: 28 889 566 autosomal and 936 919 on the X chromosome. Of these, 15 282 884 passed genotyping and imputation QC in all 3 cohorts, an additional 2 279 374 passed QC in the 2 Omni cohorts but not in the Affymetrix cohort, and another 12 264 227 passed QC in only 1 of the 3 cohorts. Among the 29 826 485 autosomal and X chromosomal post-QC variants, 8 832 649 had minor allele frequencies (MAFs) >0.01. Of these, 8 010 049 were SNPs and 822 600 were indels; 180 017 were in an exon, and 28 535 were annotated as encoding a change in amino acid sequence.

## Statistical analysis

Outcomes tested in this study were NRM and RM. Recurrent or progressive malignancy after HCT was a competing risk for NRM, and NRM was a competing risk for RM. Follow-up of recipients without NRM or RM was censored at last contact and ranged from 1.3 to 25.6 years after HCT (median, 12.9 years). Evaluation of association of genotype with outcomes was based on cause-specific HR analysis using Cox regression. All candidate variants in the recipient and donor genomes were evaluated for allelic and genotypic (recessive and dominant) associations with NRM and RM. For a SNP with a major allele *a* and a minor allele *b*, the recessive model tested the hypothesis that the genotype *bb* would be associated with a higher or lower risk compared with the collective genotypes *ab* and *aa* used as the reference. The dominant model tested the hypothesis that the collective genotypes *bb* and *ab* would be associated with a higher or lower risk compared with the genotype *aa* used as the reference. The allelic model tested the hypothesis that the minor allele *b* would be associated with a higher or lower risk compared with the major allele *a*, and the number of copies of the minor allele was modeled as an additive effect. The GWAS included both the recipient and donor genomes but was limited to the allelic model. For all SNPs, the minor allele was defined as the less frequent allele in the combined recipient and donor, discovery, and replication data set.

All analyses of candidate variants were conducted in 2 phases. The discovery phase included 60% of the subjects in the data set, and the replication phase included the remaining 40% of the subjects. Candidate variants with  $P \leq 5 \times 10^{-3}$  for association with NRM or RM were tested in the replication phase with Bonferroni adjustments for multiple comparisons. In the GWAS discovery phase, variants with MAF >0.01 and  $P \leq 1.0 \times 10^{-6}$  for association with NRM or RM were tested in the replication phase with Bonferroni adjustments for multiple comparisons. Bonferroni adjustment was applied separately for the 4 genome-end point combinations. Post

**Table 2. Association of candidate variants with NRM and RM**

Chr	Gene	SNP	Alleles*	End point	Genome	Model	MAF†	Discovery results		Replication results		
								P	HR (95% CI)	P	HR (95% CI)	Power, %‡
9	<i>CD274</i>	rs10975123	C/T	NRM	Recipient	Recessive	0.18	.00004	2.05 (1.5-2.8)	.46	0.84 (0.5-1.4)	39
9	<i>CD274</i>	rs10975123	C/T	NRM	Recipient	Allelic	0.18	.0008	1.24 (1.1-1.4)	.99	1.00 (0.9-1.2)	100
6	<i>HLA-DOA</i>	rs429916	C/A	NRM	Recipient	Recessive	0.07	.004	2.91 (1.6-5.4)	.19	0.34 (0.05-2.4)	6
6	<i>MICA</i>	rs1051792	G/A	NRM	Donor	Allelic	0.27	.002	0.84 (0.8-0.9)	.02	0.85 (0.7-1.0)	100
5	<i>IL12B</i>	rs3212227	T/G	NRM	Donor	Recessive	0.21	.005	0.56 (0.4-0.9)	.39	1.21 (0.8-1.8)	48
3	<i>CCR5</i>	rs1800023	A/G	RM	Donor	Dominant	0.36	.002	1.29 (1.1-1.5)	.14	1.16 (1.0-1.4)	100

Chr, chromosome; CI, confidence interval; HR, hazard ratio.

\*Plus strand major/minor alleles.

†MAF in the samples used for the test.

‡Post hoc power estimates to detect an HR  $\geq 1.5$  or  $\leq 0.67$  at a 2-sided significance level of .05 were based on the estimated standard error of the log HR in the replication cohort.

hoc power estimates to detect an HR  $\geq 1.5$  or  $\leq 0.67$  at a 2-sided significance level of .05 were based on the estimated standard error of the log HR in the replication cohort.

## Results

### Evaluation of candidate variants for associations with NRM and RM

Supplemental Table 1 summarizes published reports describing 130 variants having statistically significant associations with OS, PFS, or NRM after allogeneic HCT, and supplemental Table 2 summarizes the QC assessment of the candidate variants genotyped or imputed on the 3 platforms used for our study. Six SNPs were not analyzed: 3 did not pass QC (rs333, rs3969913, and rs4715333), 2 were monomorphic (rs4986893 and rs671), and 1 was not genotyped or imputed (rs396991). We were also unable to analyze the H1-homozygous *CCR5* haplotype reported by McDermott et al<sup>13</sup> or the CYP2C19 combination reported by Elmaagacli et al.<sup>14</sup>

Supplemental Table 3 shows results of testing candidate variants for association with NRM. In the discovery cohort, 2 recipient variants met criteria for replication testing (Table 2), but results were not statistically significant in the replication cohort. Post hoc power was 100% in testing rs10975123 in the allelic model. The HR point estimates for rs10975123 in the recessive model and rs429916 were notably on opposite sides of 1.0 in the discovery and replication cohorts. Taken together, these results suggest that the lack of replication did not result from insufficient power.

In the discovery cohort, 2 donor variants met criteria for replication testing (Table 2). Results for rs3212227 were not statistically significant in the replication cohort, and here again, the HR point estimates were on opposite sides of 1.0. In contrast, results for rs1051792 (*MICA*) met the Bonferroni-adjusted 0.025 threshold of statistical significance in the replication cohort. The respective HR point estimates in the discovery and replication cohorts were 0.84 and 0.85, respectively, for association of the donor rs1051792 A allele with NRM (Table 2).

Supplemental Table 4 shows results of testing candidate variants for association with RM. In the discovery cohort, only 1 donor SNP met criteria for replication testing (Table 2), but results were not statistically significant in the replication cohort. The HR point

estimates were similar, and post hoc power was 100%, indicating that the lack of replication did not result from insufficient power.

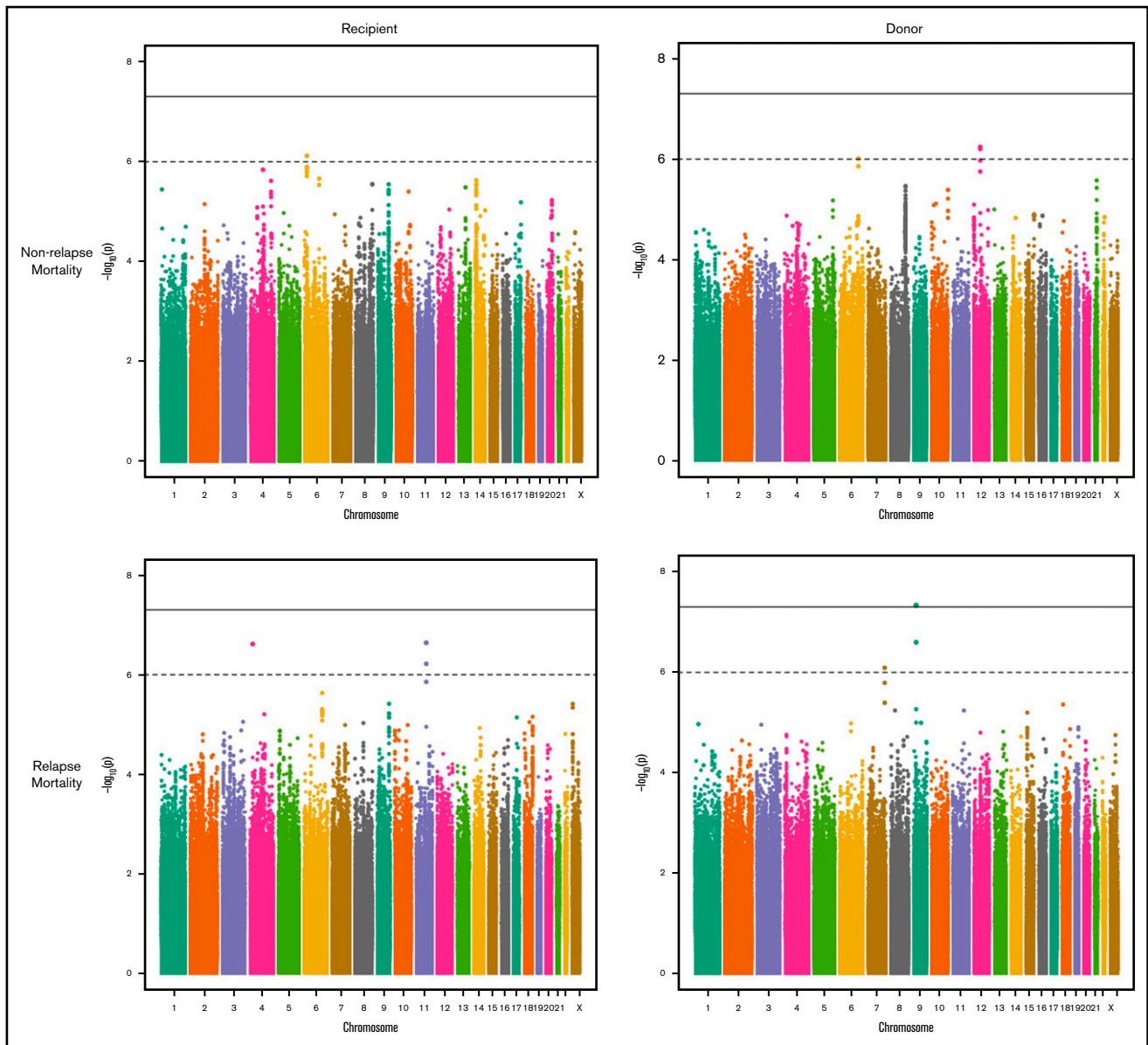
### GWAS of SNP associations with NRM and RM

In the discovery phase, the GWAS of NRM yielded 6 candidate variants that met criteria for replication testing (Figure 1). In people of European ancestry, rs9350084 and rs9350085 are in close linkage disequilibrium (LD) on chromosome 6 ( $r^2 = 1.0$ ), and rs11559982, rs10876550, and rs11451044 are in close LD on chromosome 9 ( $r^2 \geq 0.98$ ; supplemental Tables 5-6). Within each LD group, we selected the variant with the lowest *P* value for replication testing (Table 3). Replication testing thereby encompassed 3 variants, 1 from each of the 2 LD groups (recipient rs9350085 and donor rs11451044, respectively) plus donor rs9492413. Power for testing in the replication phase was 100% for all 3 variants, but none of the associations with NRM was statistically significant ( $P = .65, 1.0, \text{ and } .86$ , respectively).

In the discovery phase, the GWAS of RM yielded 6 candidate variants that met criteria for replication testing, but rs148290359 and rs74782705 are in LD on chromosome 11 ( $r^2 = 0.93$ ), and rs141251357 and rs60125430 are in LD on chromosome 9 ( $r^2 = 0.75$ ; Figure 1). In replication testing, none of the 4 variants that remained after pruning for LD had a statistically significant association with RM, although power was limited for all but 1 of these variants. Supplemental Tables 7 and 8 and supplemental Figure 2 summarize the combined discovery and replication GWAS analysis of variants with MAF  $> 1\%$  and  $P \leq 1 \times 10^{-6}$  for association with NRM or RM with the allelic model for others to test for replication.

### Further evaluation of rs1051792 (*MICA*) association with NRM

The association of the donor rs1051792 A allele with NRM remained unchanged after adjustment for clinical covariates, including donor relation, HLA mismatch, female-to-male sex mismatch, recipient and donor age, source of stem cells, and conditioning intensity (HR, 0.85; 95% CI, 0.74-0.98;  $P = .02$ ; Table 4). Isernhagen et al<sup>15</sup> previously reported that the rs1051792 A allele in the recipient was associated with lower overall mortality after allogeneic HCT, whereas our discovery and replication identified an association of NRM with the rs1051792 A allele in the donor. Because of HLA matching of donors and recipients and



**Figure 1. Manhattan plots show recipient and donor SNP associations with NRM and RM in the discovery cohort.** Each panel shows the  $-\log_{10}(P$  value) for post-QC variants with MAF  $>1\%$  for autosomes and chromosome X. The results in each panel represent  $6.47 \times 10^6$  variants. The solid line shows genome-wide significance ( $5 \times 10^{-8}$ ). The dotted line shows the threshold used to select variants for replication ( $10^{-6}$ ). Vertically aligned associations reflect variants that are strongly correlated by linkage disequilibrium. Recipient genome inflation values<sup>29</sup> were 1.005 for NRM and 1.026 for RM, and donor genome inflation values were 1.030 for NRM and 1.009 for RM.

the close LD between HLA-B and MICA, the donor and recipient genotypes for rs1051792 are highly correlated. For pairs with both genotypes, the concordance was 96%, making it statistically impossible to distinguish donor vs recipient associations. The recipient rs1051792 A allele showed an association with NRM similar to that of the donor allele in the replication cohort, although this did not reach statistical significance in the somewhat different set of patients with genotype data (HR, 0.88; 95% CI, 0.8-0.1.0;  $P = .07$ ; Figure 2A-B).

Previous reports have indicated that MICA mismatching can contribute to NRM.<sup>16,17</sup> Two observations suggest that mismatching cannot explain the association of donor rs1051792 genotypes

with NRM in our study. First, the association of the donor rs1051792 genotypes with NRM was similar in HLA genotypically identical sibling recipients who had no MICA mismatching and in unrelated recipients who could have had MICA mismatching. The respective HR estimates were 0.82 and 0.87. Second, 27 of the 1203 HLA-A, -B, -C, -DRB1, or -DQB1 matched unrelated donor-recipient pairs had graft-versus-host or host-versus-graft rs1051792 allele mismatching, which was not associated with NRM (HR, 0.91; 95% CI, 0.43-1.93;  $P = .81$ ). Similarly, 27 of the 67 unrelated pairs with a single recipient HLA-B mismatch had rs1051792 mismatching, which was also not associated with NRM (HR, 0.51; 95% CI, 0.23-1.13;  $P = .10$ ). Taken together, these results support the hypothesis that the association of rs1051792

**Table 3. Association of most significant GWAS variants per LD block with NRM and RM**

Chr	Gene	SNP	Alleles*	End point	Genome	Model	MAF†	Discovery results		Replication results		
								P	HR (95% CI)	P	HR (95% CI)	Power, %‡
6	MIR548A1HG	rs9350085	C/T	NRM	Recipient	Allelic	0.37	7.7E-07	1.29 (1.2-1.4)	.65	0.97 (0.9-1.1)	100
6	LOC105377999	rs9492413	C/T	NRM	Donor	Allelic	0.19	9.8E-07	1.35 (1.2-1.5)	.86	0.99 (0.9-1.1)	100
12	none	rs11451044	CA/C	NRM	Donor	Allelic	0.45	5.8E-07	1.28 (1.2-1.4)	1.00	1.00 (0.9-1.1)	100
4	GBA3	rs114484584	C/T	RM	Recipient	Allelic	0.02	2.4E-07	2.46 (1.8-3.3)	.23	0.63 (0.3-1.4)	16
11	CCDC81	rs148290359	C/T	RM	Recipient	Allelic	0.02	2.3E-07	0.14 (0.04-0.4)	.46	1.23 (0.7-2.1)	33
7	none	rs117446118	G/A	RM	Donor	Allelic	0.02	8.1E-07	2.91 (2.0-4.2)	.10	0.60 (0.3-1.2)	22
9	ADAMTSL1	rs60125430	C/T	RM	Donor	Allelic	0.09	4.6E-08	1.74 (1.4-2.1)	.37	1.15 (0.9-1.5)	77

\*Plus strand major/minor alleles.

†MAF in the samples used for the test.

‡Post hoc power estimates to detect an HR  $\geq 1.5$  or  $\leq 0.67$  at a 2-sided significance level of .05 were based on the estimated standard error of the log HR in the replication cohort.

with NRM is attributable to an intrinsic protective effect of the minor allele.

Because the association with NRM might be explained by differences in the risk of acute or chronic graft-versus-host disease (GVHD), we evaluated the incidence of grade 2 to 4 acute GVHD, grade 3 to 4 acute GVHD, NRM among patients with grade 3 to 4 GVHD, and chronic GVHD according to the donor and recipient rs1051792 genotypes in the combined discovery and replication cohorts (Table 5). The results showed that the donor rs1051792 A allele was not associated with the incidence of grade 2 to 4 acute GVHD but was marginally associated with a decreased risk of grade

3 to 4 acute GVHD (HR, 0.90; 95% CI, 0.81-1.01;  $P = .05$ ; Figure 2C) and a decreased risk of NRM in patients with grade 3 to 4 GVHD (HR, 0.88; 95% CI 0.76-1.01;  $P = .06$ ; Figure 2E). The donor rs1051792 A allele was not associated with the risk of chronic GVHD or recurrent malignancy. The recipient rs1051792 A allele showed no statistically significant association with any of these end points (Table 5; Figure 2D,F). Taken together, these results suggest that the lower risk of NRM associated with the donor rs1051792 A allele may be explained in part both by a lower risk of grade 3 to 4 GVHD and a lower risk of mortality among patients with grade 3 to 4 acute GVHD.

## Discussion

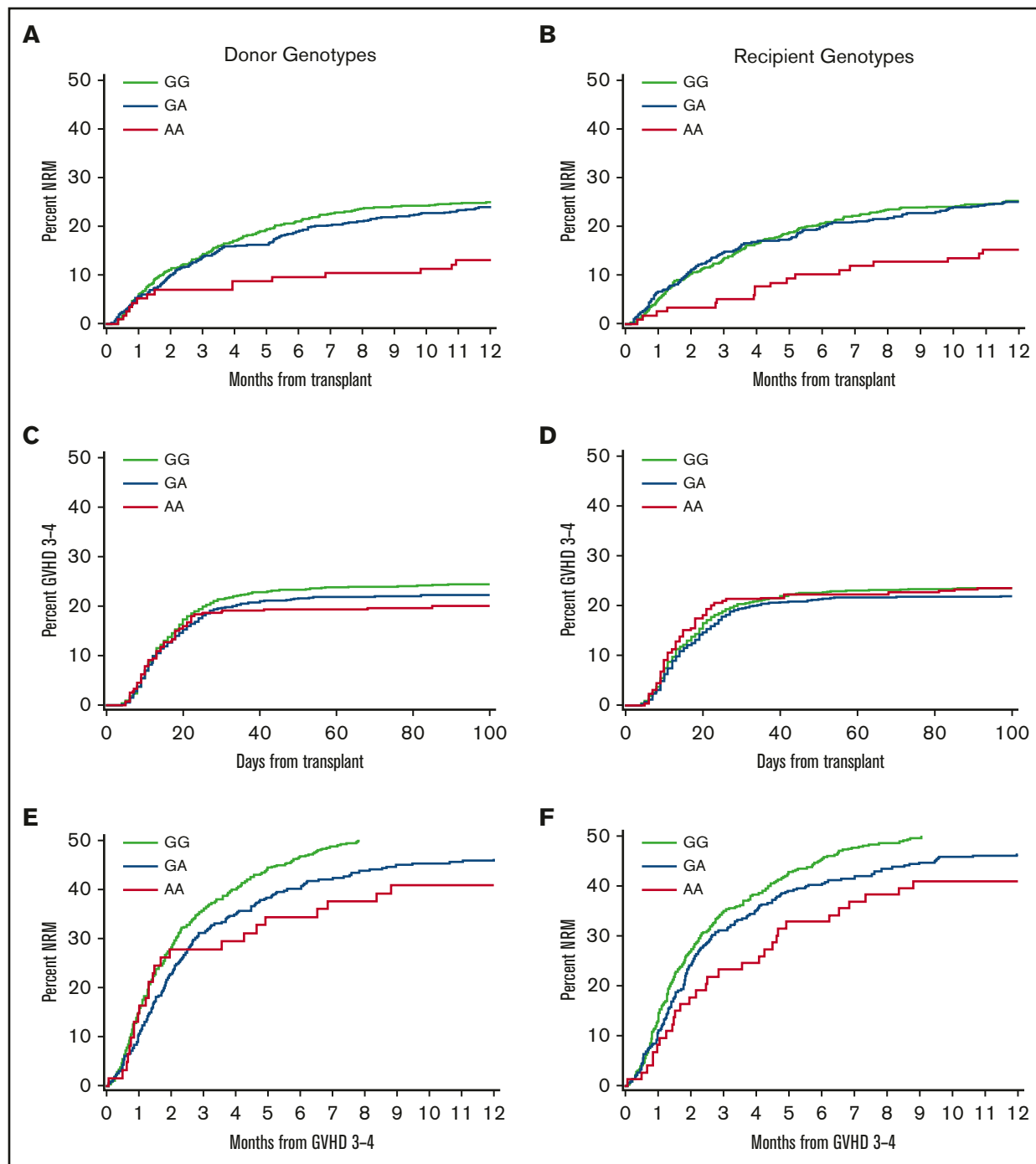
With 1 exception, broadly applicable associations with NRM or RM could not be demonstrated for previously studied donor and recipient genetic variants. The 1 exception was rs1051792, the minor allele A of which encodes methionine and major allele G of which encodes valine at position 129 in the MICA protein. No previous study has demonstrated an association of donor MICA-129Met with NRM after allogeneic HCT.<sup>16,18</sup> Fuerst et al<sup>16</sup> showed an association of donor-recipient MICA-129 mismatching with NRM after unrelated HCT, but this result was not replicated in other studies<sup>17,18</sup> or in our study, although MICA mismatching at large has been associated with increased risks of NRM and acute and chronic GVHD and a decreased risk of recurrent malignancy after HLA-A, -B, -C, -DRB1, or -DQB1 matched unrelated HCT.<sup>17</sup> Isernhagen et al<sup>15</sup> reported that the presence of MICA-129Met in the recipient was associated with improved survival after allogeneic HCT, particularly in patients who had not been treated with antithymocyte globulin as part of the pretransplantation conditioning regimen. The hazard of overall mortality associated with each additional rs1051792 A allele in the recipient was 0.77 (95% CI, 0.6-1.0;  $P = .04$ ), but the association with the hazard of NRM was not statistically significant (HR, 0.51; 95% CI, 0.2-1.1;  $P = .09$ ). In addition, the recipient MICA-129Met variant was associated with an increased frequency of grade 2 to 4 acute GVHD but lower risk of fatality in patients with acute GVHD. MICA-129Met has high avidity for binding to its ligand NKG2D, with more efficient initial NKG2D signaling, but the high-avidity binding reduces NKG2D expression and function in NK cells and CD8<sup>+</sup> T cells over time, whereas the lower-avidity binding of MICA-129Val does not have this effect.<sup>15,19</sup> Isernhagen et al<sup>15</sup> suggested that the lower NRM associated with

**Table 4. Multivariate analysis of risk factors for NRM in the replication cohort**

Risk factor	HR (95% CI)	P
Donor rs1051792 (per A allele)*	0.85 (0.74-0.98)	.02
<b>Donor relation</b>		
Related	1.0	
Unrelated	1.17 (0.94-1.46)	.15
<b>HLA match</b>		
Matched	1.0	
Mismatch	1.50 (1.22-1.86)	.0002
<b>Donor-recipient sex</b>		
Other	1.0	
Female to male	1.41 (1.17-1.70)	.0004
Recipient age (per 10 y)	1.23 (1.14-1.33)	<.0001
Donor age† (per 10 y)	1.09 (1.00-1.19)	.05
<b>Graft source</b>		
Bone marrow	1.0	
Mobilized blood cells	0.79 (0.64-0.99)	.04
<b>Conditioning</b>		
Myeloablative, <900 cGy TBI	1.0	
Myeloablative, $\geq 900$ cGy TBI	1.46 (1.18-1.81)	.0006
Nonmyeloablative	1.08 (0.81-1.45)	.61

\*Supplemental Table 9 summarizes the distribution of genotypes in the replication cohort.

†Missing donor age was accommodated with indicator variable.



**Figure 2. The donor rs1051792 A allele in *MICA* is associated with lower risk of NRM after allogeneic HCT, lower risk of grade 3 to 4 acute GVHD, and lower risk of NRM among patients with grade 3 to 4 acute GVHD.** (A) NRM among patients in the replication cohort, according to the donor rs1051792 genotypes. (B) NRM among patients in the replication cohort, according to the recipient rs1051792 genotypes. (C) Grade 3 to 4 GVHD among patients in the combined discovery and replication cohorts, according to the donor rs1051792 genotypes. (D) Grade 3 to 4 GVHD among patients in the combined discovery and replication cohorts according to the recipient rs1051792 genotypes. (E) NRM from the onset of GVHD among patients with grade 3 to 4 GVHD in the combined discovery and validation cohorts, according to the donor rs1051792 genotypes. (F) NRM from the onset of GVHD among patients with grade 3 to 4 GVHD in the combined discovery and validation cohorts, according to the recipient rs1051792 genotypes. Differences according to donor and recipient genotypes partly reflect different sets of donor-recipient pairs and cannot be attributed entirely to the respective genomes of the donors and recipients.

**Table 5. Association of donor and recipient rs1051792 genotypes with risks of GVHD and recurrent malignancy in the combined discovery and replication cohorts**

End point	Donor		Recipient	
	HR (95% CI)	P	HR (95% CI)	P
Grade 2-4 GVHD	0.96 (0.90-1.01)	.13	1.00 (0.94-1.06)	.98
Grade 3-4 GVHD	0.90 (0.81-1.01)	.05	0.96 (0.87-1.07)	.46
NRM after grade 3-4 GVHD	0.88 (0.76-1.01)	.06	0.90 (0.78-1.03)	.12
Chronic GVHD	1.00 (0.92-1.1)	.99	1.03 (0.96-1.11)	.38
Recurrent malignancy	1.00 (0.92-1.1)	.99	1.01 (0.93-1.10)	.78

All associations were tested in the allelic model.

MICA-129Met in the recipient could be explained by the attenuation of NKG2D expression and function over time after HCT.

Our results showing decreased risks of grade 3 to 4 GVHD and NRM in patients with grade 3 to 4 GVHD associated with donor MICA-129Met but not with MICA-129Val implicate MICA-129Met in the donor as an explanation for the decreased risk of NRM after HCT. Although MICA expression has been widely recognized as a marker of epithelial stress, it is also expressed by activated T cells and monocytes. Cerboni et al<sup>20</sup> showed that antigen-activated human T cells express MICA and other NKG2D ligands, thereby becoming susceptible to fratricidal killing by autologous NK cells. Subsequent studies showed that donor NK cell-mediated killing of alloactivated donor T cells attenuates the severity of acute<sup>21</sup> and chronic<sup>22</sup> GVHD in mice without impairing graft-versus-tumor effects,<sup>21</sup> although in some circumstances, exogenous hyperactivation can cause NK cells to produce proinflammatory cytokines that sustain induction of GVHD by T cells.<sup>23</sup> Based on these results, we hypothesize that alloactivated donor T cells with MICA-129Met have increased susceptibility to killing by donor NK cells or CD8<sup>+</sup> T cells<sup>24</sup> as a result of stronger binding to NKG2D when compared with donor T cells with MICA-129Val. Our results do not refute effects mediated by recipient MICA-129Met. Instead, they highlight a previously unrecognized mechanism that could explain effects mediated by donor MICA-129Met. Although linkage of MICA with the major histocompatibility complex would not allow clinical application of our results through donor selection, it is possible that addition of donor NK cells to the graft could attenuate the severity of acute GVHD.<sup>21-23</sup>

Our study and the CIBMTR studies<sup>1,2</sup> had similarities and many differences but nonetheless reached the same general conclusion regarding the association of genetic variants with mortality-related end points after HCT. Similarities included the selection of candidate variants from a comprehensive screen of the literature reporting statistically significant SNP associations with OS, PFS, NRM, and RM, the focus on recipients of European ancestry, the use of myeloablative and reduced-intensity conditioning regimens, and the use of marrow or growth factor-mobilized blood cells for grafting.

At the same time, our approach differed from the CIBMTR approach in many respects, as summarized in Table 6. The CIBMTR candidate SNP analysis included 47 variants statistically associated with survival outcomes reported in previous publications, whereas our analysis included more than twice that number of variants

**Table 6. Differences between studies**

Characteristic	CIBMTR	Current study
Centers	Multiple	Single
Donor types	Unrelated	Related and unrelated
HLA-A, -B, -C, -DRB1, and -DQB1 mismatching included	No	Yes
No. of recipients, candidate SNP study	Up to 2887	Up to 2560 for discovery
No. of recipients, genome-wide study	1970 for discovery	Up to 2560 for discovery
Recipient diseases	AML, ALL, MDS	Any hematologic malignancy
End points	OS, PFS, NRM, RM	NRM, RM
End point adjudication	Yes	No
No. of individual candidate variants	47	122
Candidate SNP genetic models	Allelic	Allelic, dominant, recessive
Adjustment for clinical covariates	Yes	No
Truncation of follow-up	1 y	None
Tested recipient SNP allele mismatching	Yes	No
GWAS scope	Exomes	Whole genome
Statistical replication or validation	Metaanalysis	3:2 discovery/replication split
Gene-level analysis	Yes	No

ALL indicates acute lymphoid leukemia; AML, acute myeloid leukemia; MDS, myelodysplastic syndrome.

(n = 122). The Karaesmen et al<sup>1</sup> study drew a distinction between replication and validation.<sup>25</sup> Replication testing requires matching the same demographic, disease, and donor type inclusion and exclusion criteria and the end points used for each prior study, whereas validation allows differences to determine whether results apply more broadly. In our study, prior reports were used solely to identify candidate variants, and we made no attempt to match the inclusion and exclusion criteria used in prior studies. The scope of the CIBMTR GWAS was limited to exons, whereas our GWAS encompassed all variants that passed QC and had an MAF > 1%. The only statistically significant result replicated or validated in the CIBMTR cohort was the association of the donor *IL6* rs1800795 C allele with OS (HR, 1.11; 95% CI, 1.0-1.2; P = .018 in the allelic model).<sup>1</sup> In our cohort, however, rs1800795 donor and recipient genotypes showed no statistically significant association with the risk of NRM or RM in recessive, dominant, or allelic models (supplemental Tables 3 and 4).

The negative results of our study and other studies<sup>1,26-28</sup> give rise to the question of whether NRM or RM can be considered as a true phenotype in HCT recipients. With respect to NRM, HCT recipients have variable vulnerabilities related to age and the residual morbidity caused by the underlying disease and prior treatment, and they are subjected to variable stressors such as pretransplantation conditioning regimen, GVHD, and infections. Vulnerability to RM depends on the nature of the underlying disease, the extent to which prior treatment selected for malignant cells that are resistant to the pretransplantation conditioning regimen, and the burden of disease at the time of HCT, whereas higher-intensity conditioning regimens



and graft-versus-leukemia effects decrease the risk of recurrent or progressive malignancy after HCT. As a result, death after HCT has many causes. Accordingly, 1 SNP could affect NRM or RM only if it has wide-ranging effects across many pathophysiological mechanisms or such a large effect on any single pathophysiological mechanism that an association with NRM or RM could be detectable with the size of the available HCT cohorts that have extensive genotype data available. Results showing that previously reported genetic associations do not apply broadly should discourage additional studies treating death as a phenotype. Future studies should focus instead on associations with the individual intermediate phenotypes that precede death.

## Acknowledgments

The authors thank Veronika Groh for assistance with the interpretation of our results.

This work was supported by National Institutes of Health grants from the National Institute of Allergy and Infectious Diseases (AI33484 and AI149213), National Cancer Institute (CA015704 and CA18029), and National Heart, Lung, and Blood Institute (HL087690, HL088201, HL094260, HL105914, and K23HL69860).

## References

1. Karaesmen E, Rizvi AA, Preus LM, et al. Replication and validation of genetic polymorphisms associated with survival after allogeneic blood or marrow transplant. *Blood*. 2017;130(13):1585-1596.
2. Zhu Q, Yan L, Liu Q, et al. Exome chip analyses identify genes affecting mortality after HLA-matched unrelated-donor blood and marrow transplantation. *Blood*. 2018;131(22):2490-2499.
3. Conomos MP, Laurie CA, Stilp AM, et al. Genetic diversity and association studies in US Hispanic/Latino populations: applications in the Hispanic Community Health Study/Study of Latinos. *Am J Hum Genet*. 2016;98(1):165-184.
4. Chien JW, Zhang XC, Fan W, et al. Evaluation of published single nucleotide polymorphisms associated with acute GVHD. *Blood*. 2012;119(22):5311-5319.
5. Affymetrix. BRLMM: an improved genotype calling method for the GeneChip Human Mapping 500K Array Set. [http://media.affymetrix.com/support/technical/whitepapers/brlmm\\_whitepaper.pdf](http://media.affymetrix.com/support/technical/whitepapers/brlmm_whitepaper.pdf). Accessed 14 January 2014.
6. Illumina. Illumina GenCall data analysis software. [https://www.illumina.com/documents/products/technotes/technote\\_gencall\\_data\\_analysis\\_software.pdf](https://www.illumina.com/documents/products/technotes/technote_gencall_data_analysis_software.pdf). Accessed 6 July 2016.
7. 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.
8. Zheng X, Shen J, Cox C, et al. HIBAG-HLA genotype imputation with attribute bagging. *Pharmacogenomics J*. 2014;14(2):192-200.
9. Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3 (Bethesda)*. 2011;1(6):457-470.
10. Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods*. 2013;10(1):5-6.
11. Auton A, Brooks LD, Durbin RM, et al; 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74.
12. Nelson SC, Doheny KF, Laurie CC, Mirel DB. Is "forward" the same as "plus"?...and other adventures in SNP allele nomenclature. *Trends Genet*. 2012;28(8):361-363.
13. McDermott DH, Conway SE, Wang T, et al. Donor and recipient chemokine receptor CCR5 genotype is associated with survival after bone marrow transplantation. *Blood*. 2010;115(11):2311-2318.
14. Elmaagacli AH, Koldehoff M, Steckel NK, Trensche R, Ottinger H, Beelen DW. Cytochrome P450 2C19 loss-of-function polymorphism is associated with an increased treatment-related mortality in patients undergoing allogeneic transplantation. *Bone Marrow Transplant*. 2007;40(7):659-664.
15. Isernhagen A, Malzahn D, Viktorova E, et al. The MICA-129 dimorphism affects NKG2D signaling and outcome of hematopoietic stem cell transplantation. *EMBO Mol Med*. 2015;7(11):1480-1502.
16. Fuerst D, Neuchel C, Niederwieser D, et al. Matching for the MICA-129 polymorphism is beneficial in unrelated hematopoietic stem cell transplantation. *Blood*. 2016;128(26):3169-3176.
17. 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.

## Authorship

Contribution: J.A.H., P.J.M., D.M.L., and B.E.S. were responsible for the study concept and design, interpretation of results, and writing of the manuscript; S.C.N. assisted with imputation and genomic phasing; X.D. assisted in curation of candidate variants and GWAS analyses; D.M.L. and B.E.S. were responsible for data acquisition, imputation, informatics analyses, quality control, and statistical analyses; and P.J.M., D.M.L., B.E.S., S.C.N., and X.D. revised draft manuscripts and approved the final version of the manuscript.

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

John A. Hansen died on 31 July 2019.

ORCID profiles: P.J.M., 0000-0001-9051-1215; D.M.L., 0000-0002-9292-570X; S.C.N., 0000-0002-2109-6465.

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

18. Askar M, Sobecks R, Wang T, et al. MHC class I chain-related gene A (MICA) donor-recipient mismatches and MICA-129 polymorphism in unrelated donor hematopoietic cell transplantations has no impact on outcomes in acute lymphoblastic leukemia, acute myeloid leukemia, or myelodysplastic syndrome: a Center for International Blood and Marrow Transplant research study. *Biol Blood Marrow Transplant*. 2017;23(3):436-444.
19. Steinle A, Li P, Morris DL, et al. Interactions of human NKG2D with its ligands MICA, MICB, and homologs of the mouse RAE-1 protein family. *Immunogenetics*. 2001;53(4):279-287.
20. Cerboni C, Ardolino M, Santoni A, Zingoni A. Detuning CD8+ T lymphocytes by down-regulation of the activating receptor NKG2D: role of NKG2D ligands released by activated T cells. *Blood*. 2009;113(13):2955-2964.
21. Olson JA, Leveson-Gower DB, Gill S, Baker J, Beilhack A, Negrin RS. NK cells mediate reduction of GVHD by inhibiting activated, alloreactive T cells while retaining GVT effects. *Blood*. 2010;115(21):4293-4301.
22. Noval Rivas M, Hazzan M, Weatherly K, Gaudray F, Salmon I, Braun MY. NK cell regulation of CD4 T cell-mediated graft-versus-host disease. *J Immunol*. 2010;184(12):6790-6798.
23. Simonetta F, Alvarez M, Negrin RS. Natural killer cells in graft-versus-host-disease after allogeneic hematopoietic cell transplantation. *Front Immunol*. 2017;8:465.
24. Roberts AI, Lee L, Schwarz E, et al. NKG2D receptors induced by IL-15 costimulate CD28-negative effector CTL in the tissue microenvironment. *J Immunol*. 2001;167(10):5527-5530.
25. Igl BW, Konig IR, Ziegler A. What do we mean by "replication" and "validation" in genome-wide association studies? *Hum Hered*. 2009;67(1):66-68.
26. Nguyen Y, Al-Lehibi A, Gorbe E, et al. Insufficient evidence for association of NOD2/CARD15 or other inflammatory bowel disease-associated markers on GVHD incidence or other adverse outcomes in T-replete, unrelated donor transplantation. *Blood*. 2010;115(17):3625-3631.
27. Pearce KF, Lee SJ, Haagenson M, et al. Analysis of non-HLA genomic risk factors in HLA-matched unrelated donor hematopoietic cell transplantation for chronic myeloid leukemia. *Haematologica*. 2012;97(7):1014-1019.
28. Arora M, Lee SJ, Spellman SR, et al. Validation study failed to confirm an association between genetic variants in the base excision repair pathway and transplant-related mortality and relapse after hematopoietic cell transplantation. *Biol Blood Marrow Transplant*. 2016;22(8):1531-1532.
29. Devlin B, Roeder K. Genomic control for association studies. *Biometrics*. 1999;55(4):997-1004.