ASBMT American Society for Blood and Marrow Transplantation

Degree of Predicted Minor Histocompatibility Antigen Mismatch Correlates with Poorer Clinical Outcomes in Nonmyeloablative Allogeneic Hematopoietic Cell Transplantation

Malene Erup Larsen,[1] Brian Kornblit,[2] Mette Voldby Larsen,[1] Tania Nicole Masmas,[2] Morten Nielsen,[1] Martin Thiim,[1] Peter Garred,[3] Anette Stryhn,[4] Ole Lund,[1] Soren Buus,[4] Lars Vindelov[2]

In fully HLA-matched allogeneic hematopoietic cell transplantation (HCT), the main mechanism of the beneficial graft-versus-tumor (GVT) effect and of detrimental graft-versus-host disease (GVHD) is believed to be caused by donor cytotoxic T cells directed against disparate recipient minor histocompatibility antigens (miHAs). The most common origin of disparate miHAs is nonsynonymous single nucleotide polymorphism (nsSNP) differences between donors and patients. To date, only some 30 miHAs have been identified and registered, but considering the many different HLA types in the human population, as well as all the possible nsSNP differences between any 2 individuals, it is likely that many miHAs have yet to be discovered. The objective of the current study was to predict novel HLA-A– and HLA-B–restricted miHAs in a cohort of patients treated with nonmyeloablative conditioning allogeneic HCT (matched related donor, n = 70; matched unrelated donor, n = 56) for a hematologic malignancy. Initially, the cohort was genotyped for 53 nsSNPs in 11 known miHA source proteins. Twenty-three nsSNPs within 6 miHA source proteins showed variation in the graft-versus-host (GVH) direction. No correlation between the number of disparate nsSNPs and clinical outcome was seen. Next, miHAs in the GVH direction were predicted for each patient–donor pair. Using the *NetMHCpan* predictor, we identified peptides encompassing an nsSNP variant uniquely expressed by the patient and with predicted binding to any of the HLA-A or -B molecules expressed by the patient and donor. Patients with more than the median of 3 predicted miHAs had a significantly lower 5-year overall survival (42% vs 70%, $P = .0060$; adjusted hazard ratio [HR], 2.6, $P = .0047$) and significantly higher treatment-related mortality (39% vs 10%, $P = .0094$; adjusted HR, 4.6, $P = .0038$). No association between the number of predicted miHAs and any other clinical outcome parameters was observed. Collectively, our data suggest that the clinical outcome of HCT is affected not by disparate nsSNPs *per se*, but rather by the HLA-restricted presentation and recognition of peptides encompassing these. Our data also suggest that 6 of the 11 proteins included in the current study could contain more miHAs yet to be identified, and that the presence of multiple miHAs confers a higher risk of mortality after nonmyeloablative conditioning HCT. Furthermore, our data suggest a possible role for *in silico* based miHA predictions in donor selection as well as in selecting candidate miHAs for further evaluation in *in vitro* and *in vivo* experiments.

Biol Blood Marrow Transplant 16: 1370-1381 (2010) © 2010 American Society for Blood and Marrow Transplantation

KEY WORDS: Allo-HCT, mHags, Nonsynonymous SNPs, GVHD, Survival

From the [1]Center for Biological Sequence Analysis, DTU Systems Biology, Technical University of Denmark, DK-2800 Lyngby, Denmark; [2]Allogeneic Hematopoietic Cell Transplantation Laboratory, Department of Hematology, Rigshospitalet, Copenhagen, Denmark; [3]Laboratory of Molecular Medicine, Department of Clinical Immunology-7631, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark; and [4]Laboratory of Experimental Immunology, University of Copenhagen, DK-2200 Copenhagen N, Denmark.

Financial disclosure: See Acknowledgments on page 1380.

Correspondence and reprint requests: Malene Erup Larsen, Center for Biological Sequence Analysis, DTU Systems Biology, Building 208, Technical University of Denmark, DK-2800 Lyngby, Denmark (e-mail: mel@cbs.dtu.dk).

Received January 22, 2010; accepted March 23, 2010

1083-8791/$36.00
doi:10.1016/j.bbmt.2010.03.022

INTRODUCTION

In recent years, the role of minor histocompatibility antigens (miHAs) in HLA-matched allogeneic hematopoietic cell transplantation (HCT) has become increasingly evident. miHAs are immunogenic HLA-presented peptides derived from protein products of polymorphic genes that are disparate between patient and donor [1]. Although most of these polymorphic proteins result from nonsynonymous single nucleotide polymorphisms (nsSNPs) in autosomal genes, miHAs also may be caused by gene deletions, genetic variation in noncoding regions affecting gene transcription or the presence of Y chromosome–encoded proteins in sex-mismatched HCT [2-5]. Depending on their tissue distribution, miHAs with broad tissue expression may induce graft-versus-host-disease (GVHD), whereas miHAs, which are expressed only in hematopoietic tissue, may induce a graft-versus-tumor (GVT) effect [1]. Several studies have linked the presence of miHA-specific T cells posttransplantation with graft rejection [6,7], GVHD [8-10], and the GVT effect [11-13]. Because GHVD is a major cause of transplantation-related morbidity and treatment-related mortality (TRM) [14], identification and characterization of miHAs specifically expressed in hematopoietic but not other normal tissues could contribute to the development of selective GVT-oriented immunotherapy by separating the beneficial GVT effect from GVHD. Approximately 30 miHAs have been identified [15] by various methods, including peptide elution from the major histocompatibility complex (MHC) [12,16], expressional cloning [3,17,18], genetic linkage analysis [11,19], and genome-wide association analysis [20-22]. Common to all these methods is that they identify only a few miHAs restricted to no more than few HLA types. Considering the diversity of the HLA system with >3500 known alleles [23-25], as well as the >120,000 known allelic nsSNP variants [26], it seems likely that many miHAs have yet to be identified. If GVT-oriented immunotherapy is to be broadly applicable to a large number of patients, then the number of known miHAs needs to be expanded in a systematic manner and on a larger scale using computerized methods [27]. This has been addressed by several different bioinformatics techniques using algorithms to integrate databases containing information about protein processing, MHC–peptide binding, SNP data, and tissue-specific gene expression [11,28-30]. *NetMHCpan* [31] is an MHC–peptide binding prediction tool capable of predicting the binding of peptides to any MHC molecule with a known protein sequence. The method is based on an Artificial Neural Network trained on experimental MHC–peptide binding data. In 2 recent comparisons, *NetMHCpan* has proven superior to other available predictors in predicting HLA class I binding [32,33]. The purpose of the current project was to investigate the association between the number of predicted miHAs and the outcome after allogeneic HCT with nonmyeloablative (NMA) conditioning. Based on patient and donor HLA-A and -B types and genotype for a number of nsSNPs, miHAs were predicted in proteins already known to contain miHAs. Known miHA source proteins were chosen because the previous discovery of miHAs in these proteins indicates that they are expressed in relevant tissues and have an expression and degradation frequency that allows peptides from the proteins to be presented by HLA molecules.

MATERIALS AND METHODS

Patients

This analysis includes data from 126 consecutive patients who underwent allogeneic HCT with a peripheral blood graft from an HLA-identical related or 10/10 allele-matched unrelated donor after NMA conditioning between April 2000 and July 2007 at the allo-HCT unit, Department of Hematology, Rigshospitalet, Copenhagen. For related donors, donor selection was based on serologic typing for HLA-A, -B, and -C, and on molecular typing for HLA class II. For unrelated donors, donor selection was based on molecular typing for HLA-A, -B, -C, -DRB1, and -DQB1. When available, HLA-identical siblings were preferred to matched unrelated donors, and cytomegalovirus serostatus and sex mismatch were taken into account when possible. Molecular class I typing of related patients and donors was performed retrospectively as part of this study. All patients were treated for a malignant hematologic disease, including acute myelogenous leukemia/myelodysplastic syndrome (n = 58), non-Hodgkin lymphoma (n = 25) (follicular lymphoma, n = 15; diffuse large B cell lymphoma, n = 4; mantle cell lymphoma, n = 3; peripheral T cell lymphoma, n = 3), chronic lymphocytic leukemia (n = 18), multiple myeloma (n = 12), and Hodgkin disease (n = 13). The diseases were classified as low, standard, or high risk according to Kahl et al. [34]. Detailed patient and donor demographic data are summarized in Table 1. Donor treatment, conditioning regimen, and supportive care were as described previously [35]. All patients were conditioned with fludarabine 30 mg/m^2 for 3 days and 2 Gy of total body irradiation (TBI), except for 2 patients who were conditioned with 2 Gy TBI only. Acute and chronic GVHD (aGVHD, cGVHD) was diagnosed according to standard criteria [36]. Informed consent was obtained from all patients, and the local Ethics Committee approved the study design.

Prediction of miHAs

Eleven non–Y chromosomal proteins (Table 2) known to contain miHAs were selected from the dbMinor database [15]. The amino acid sequences of

Table 1. Patient and Donor Characteristics

Variable	Patients (n = 126), n (%)
Patient age, years	
Median=53	
Range 19–69	
Donor age, years	
Median=44	
Range 19–68	
Patient age ≤ 40 years	18 (14%)
Patient age > 40 years	108 (86%)
Donor age ≤ 40 years	50 (40%)
Donor age > 40 years	76 (60%)
Type of donor	
Matched related	70 (56%)
Matched unrelated	56 (44%)
Sex of patient/donor	
Male/female	28 (22%)
Other combinations	98 (78%)
Underlying disease*	
Low risk	25 (20%)
Standard risk	63 (50%)
High risk	38 (30%)
CMV status of patient/donor	
CMV-negative/CMV-negative	25 (20%)
Other combinations	101 (80%)

CMV indicates cytomegalovirus.
*Underlying disease was classified as low, standard, or high risk according to Kahl et al. [34].

the 11 proteins were obtained from RefSeq [37], and the nsSNPs in these were identified using dbSNP [38]. *NetMHCpan* was used to predict the binding to the HLA-A or -B molecules presented by the patients for all peptides with a length of 8-11 amino acids encompassing the nsSNPs. For each HLA allele, binding peptides were defined as those peptides with a predicted binding strength within the top 1% among random natural peptides. A total of 53 nsSNPs were selected for genotyping (Table 3), all with a minor allele frequency of ≥1% in the HapMap CEU population [39] and located within peptides predicted to bind to at least one of the HLA-A or -B molecules represented in the patient cohort. For a peptide to be considered a potential miHA in the graft-versus-host (GVH) direction for a given patient, the peptide should be predicted to bind at least one of the patient's HLA-A or -B molecules according to the foregoing definition, and the patient should carry the allele coding the binding peptide variant, whereas the donor should be homozygous for the alternative allele. This definition also allows for the donor's variant of the peptide to be a predicted binder, because the donor's T cells might recognize the difference between the 2 variant peptides.

Genotyping

Pretransplantation DNA from patients and DNA from donors were genotyped for the 53 nsSNPs using a 12-plex format GenomeLab SNPstream genotyping system (Beckman Coulter, Brea, CA) according to the manufacturer's protocol. The genotype of each of the polymorphisms was validated in 5-10 samples by direct Sanger sequencing (ABI Prism 3100 Genetic Analyzer; Applied Biosystems, Foster City, CA) using PCR primers designed for the SNPstream Genotyping system (autoprimer.com; Beckman Coulter) and purification by ethanol precipitation as described previously [40]. In some cases, failed or missing genotypes could be inferred from linkage disequilibrium (LD) with the successfully genotyped nsSNPs. The criterion for inferring genotypes in this way was complete LD ($R^2 = 1$) using the CEU population in the HapMap database [39]. To validate the genotyping assay in the event of departure from the Hardy-Weinberg equilibrium (HWE), a control population of 96 healthy Danish Caucasian blood donors was genotyped by direct sequencing for the relevant nsSNPs.

Statistical Analysis

LD, expressed as the squared correlation coefficient, R^2, quantified between all pairs of biallelic loci was estimated using SNPAlyze version 4.0 (Dynacom, Yokohama, Japan). The HWE was assessed separately in the patient and donor populations, and analyzed using gene frequencies obtained by simple gene counting and the χ^2 test. Where applicable, Fisher's exact test was used to compare frequencies.

Cox regression was used to estimate the association between the number of nsSNP differences or predicted miHAs and overall survival (OS), progression-free

Table 2. Proteins Selected from dbMinor and Their Reported miHAs

Protein Symbol	Protein Name	Protein Length, aa	Known miHA	Sequence
HMHA1	Histocompatibility (minor) HA-1	1138	HA1, HA-1/B60	VLHDDLLEA, KECVLHDDL
MYO1G	Myosin 1G	1018	HA-2	YIGEVLVSV
AKAP13	A kinase (PRKA) anchor protein 13	2817	HA-3	VTEPGTAQY
KIAA0020	KIAA0020	649	HA-8	RTLDKVLEV
HMHB1	Histocompatibility (minor) HB1	41	HB-1H, HB-1Y	EEKRGSLHVW, EEKRGSLYW
BCL2A1	BCL2-related protein A1	175	ACC-1, ACC-2	DYLQYVLQI, KEFEDDIINW
LRH1	Purinergic receptor P2X5 isoform A	422	LRH-1	TPNQRQNVC
ECGF1	Endothelial cell growth factor	482	LB-ECGF-1H	RPHAIRRPLAL
CTSH	Cathepsin H	335	CTSH/A31, A33	ATLPLLCAR, WATLPLLCAR
TOR3A	Torsin family 3, member A	397	LB-ADIR-1F	SVAPALALFPA
SP110	SP110 nuclear body protein, isoform A	689	SP110(HwA-9)	SLPRGTSTPK

miHAs indicates minor histocompatibility antigen.

Table 3. Overview of the Selected nsSNP with Variation in the GVH Direction, Predicted miHAs, and Prevalence in Patients

Protein	nsSNPs	Predicted miHAs around nsSNP	Patients with nsSNP in GVH Direction	Patients with Predicted miHA(s)	HLA Types Covered	Example Predicted miHA	Known miHAs around nsSNP
SP110	rs9061	6	15	11*	A0301, A1101, A6801, B3801, B4001	KLTSKMNA(K/E)	
	rs28930679	2	28	11*	B4001	(A/V)EEDSEEMPSL	
	rs1135791	12	47	31	A0301, A1101, A3101, A3102, A3103, B0801, B2702, B4001	(M/T)TLGELLK	
	rs3948463	11	12	12	11 HLA-As, B3505, B1302, B5101, B5201	MLWSCTFCR(I/M)	
	rs3948464	17	25	2	A3001, A3101, B2702, B2705	RTKCARKSR(L/S)K	
HMHB1	rs161557	14	25	15	A3001, A0201, A0203, A2402, 8 HLA-Bs	(Y/H)VWKSELVEV	HB-1H, HB-1Y
AKAP13	rs745191	4	24	16	A0101, 11 HLA-Bs	PSDLALL(V/G)	
	rs2061821	8	33	26	A0101, A2902, A3002, A8001, A2501, A2601,11 HLA-Bs	V(M/T)EPGTAQY	HA-3
	rs2061822	8	35	21	A0301, A1101, 10 HLA-Bs	LMNPDATV(W/R)K	
	rs2061824	2	34	6	A3001, B4001	(R/C)EESADAPV	
	rs4075254	7	35	19	A0101, A1101, A0301, 7 HLA-Bs	NTDSSLQS(V/M)	
	rs4075256	6	35	24	B0702, B5501, B5601, B1401, B1402, B3901, B3701, B4001	RPLEDRA(V/A)GL	
	rs4843074	2	33	0	A0203, B3502, B3503	DALNCSQ(P/A)SPL	
	rs4843075	8	36	2	A6802, B4001, B4901, B1302, B3701, B4501, B5001	CEVSG(D/N)VTV	
	rs7162168	8	36	16	A0301, A3101, A3102, A3103, A6601, A6801, 7 HLA-Bs	V(M/T)RAPPSGR	
	rs7177107	4	15	4	A6801, A0301, B4501	KLCDNIVS(K/E)	
	rs34434221	4	5	0	A1101, B0702, B5501, B5601, B3801, B5101	(Q/K)PVDKISV	
	rs35624420	5	2	0	7 HLA-As, 7 HLA-Bs	RAVGLSTS(F/S)	
BCL2A1	rs1138357	16	29	25	10 HLA-As, 9 HLA-Bs	YLQ(Y/C)VLQI	ACC-1
	rs1138358	17	29	24	8 HLA-As, 7 HLA-Bs	VLQ(K/N)VAFSV	
	rs3826007	14	27	16	A3201, A2501, 14 HLA-Bs	KEFEDDII(G/D)II	ACC-2
MYO1G	rs3735485	10	27	12	8 HLA-As, B1518, B0801, B0809	D(M/T)HHRHHL	
	rs7792760	9	26	8	A0301, A3001, A3101, A3102, A3103, A6801, B3901	RLKTL(Q/R)DK	
KIAA0020	rs2173904	7	33	19*	A0301, A1101, A3001, A2301, A2402, 9 HLA-Bs	KSADH(R/P)TLDK	HA-8
	rs2270891	11	6	19*	A0201, A0301, A1101, A3001, A3201, 10 HLA-Bs	LE(V/L)QPEKL	
	rs10968457	3	6	2	A0301, A3001	KQFTGK(S/N)TK	

miHA indicates minor histocompatibility antigen, nsSNP, nonsynonymous single nucleotide polymorphism.

For each nsSNP, the following are listed: number of predicted miHAs, number of patients with the nsSNP difference in the GVH direction, the number of patients with at least one predicted miHA around the nsSNP, the HLA types to which the miHAs around the nsSNP are predicted to bind, an example miHA, and the name of any known miHAs around the nsSNP.

*Number of patients with predicted miHAs containing either of 2 close SNPs.

survival (PFS), relapse incidence (RI), relapse-related mortality (RRM), TRM, grade II-IV aGVHD, grade III-IV aGVHD, or extensive cGVHD. Probability of OS and PFS was estimated by the Kaplan-Meier method, and comparisons were made with the log-rank test, whereas the cumulative incidences of RI, RRM, TRM, and GVHD were compared using Gray's *K* test [41,42]. In the estimates of RI, RRM, TRM, and GVHD, death before relapse, death with or without relapse, death without GVHD, and retransplantation were handled as competing events when appropriate [42]. OS was measured from the time of transplantation until death from any cause. Patients still alive at the time of analysis were censored at the date of last follow-up. PFS was calculated from the date of transplantation to the date of first relapse or death. Patients who were alive and in remission were censored at date of last follow-up. TRM was defined as death in complete remission (CR) or death where it was not possible to assess disease status before death. RRM was defined as death during relapsed or progressive disease. In the multivariate Cox regression models, all of the covariates listed in Table 1, along with the presence of GVHD (time-dependent covariate), were entered one by one into a pairwise model together with the number of nsSNP differences or predicted miHAs. The covariates were kept in the final model if they remained significant ($P < .05$) or altered the association with the number of nsSNP differences or predicted miHAs by >10%. All P values were 2-tailed, and $P < .05$ was considered significant.

RESULTS

Transplantation Outcome

In our cohort of 126 patients, the median follow-up was 837 days (range, 30-3178 days). The 5-year OS and PFS were 58% and 49%, respectively. The probability of grade II-IV aGVHD within the first year was 69%, and the 3-year probability of extensive cGVHD was 44%.

Genotyping of Patients

The patient and donor cohorts were successfully genotyped for 31 of the 53 selected nsSNPs, and a variation in the GVH direction was observed in 23 of these. There was no significant difference in the distribution of genotype frequencies between patients and donors, except for rs2061821 ($P = .036$) and rs1135791 ($P = .046$) (Table 4). Sixteen of 23 nsSNPs adhered to the HWE ($P > .05$). Of the 7 polymorphisms that departed from the HWE, 6 were in strong LD located in *AKAP13* (Table 5), and 1 was located in *SP110* (rs9061). Genotypes of SNPs that failed the HWE assumption were validated by direct sequencing of approximately 10% of the patient and donor cohorts. Furthermore, to ensure unbiased genotyping, the assay for these 7 nsSNPs was further validated in a cohort of 96 healthy controls (data not shown). When genotypes in the 96 control individuals were analyzed together with donor samples, all 7 nsSNPs adhered to HWE (data not shown). Three of the nsSNPs (rs4843075, rs7162168, and rs10968457) that failed genotyping were inferred based on complete LD ($R^2 = 1$) according to the HapMap CEU population to some of the 23 varying nsSNPs, thus resulting in a total of 26 varying nsSNPs. In detail, rs4843075 and rs7162168 in the AKAP13 protein were in LD with a block of 5 nsSNPs (rs4843074, rs2061821, rs2061824, rs4075256, and rs4075254), which were successfully genotyped. In the KIA0020 protein, rs10968457 was in LD with rs2270891.

Effect of Number of nsSNPs in the GVH Direction on Outcome

The median number of nsSNP differences in the GVH direction between patient and donor was 4 (range, 0-17). Patients with ≤4 nsSNP differences in the GVH direction had a nonsignificant higher 5-year OS and PFS than patients with >4 nsSNP differences (Table 6 and Figure 1A). Likewise, patients with ≤4 nsSNP differences had a nonsignificant lower 5-year TRM than patients with >4 nsSNP differences (Figure 1B). No difference in outcome was observed for any of the other clinical parameters ($P > .30$).

Identification of Potential miHAs

A total of 26 nsSNPs within 6 of the 11 proteins showed variation in the GVH direction (Table 3). Whenever a patient–donor pair had an nsSNP difference in the GVH direction, the binding of peptides containing the nsSNP to the patient's HLA-A and -B molecules was assessed using *NetMHCpan* [31]. A binding strength threshold of 1% (binding strength falling within the top 1% compared with a large set of random natural peptides) was used in the analyses. Thresholds of 0.5% and 2% were tested without significantly altering the outcome of the analyses. In some cases, more than one peptide of length 8-11 aa was a predicted binder, and thus each nsSNP in the GVH direction could result in 1, 2, or more potential miHAs for a given patient. Significantly fewer nsSNP differences were present between patients and donors when the donor was matched related compared with matched unrelated (2 vs 7; $P < 10^{-5}$; Mann-Whitney *U* test). Similarly, related patient–donor pairs had fewer predicted miHAs than unrelated pairs (1 vs 5; $P < 10^{-3}$; Mann-Whitney *U* test). Figure 2A shows the distribution of patient–donor pairs according to number of nsSNP differences in the GVH direction,

Table 4. Distribution of Genotypes

				Patients					Donors				
nsSNP	Major allele A	Minor allele a	Fisher's exact test	Failed Genotypes, %	HWE	AA, %	Aa, %	aa, %	Failed Genotypes, %	HWE	AA, %	Aa, %	aa, %
rs9061	G	A	0.45	0	0.80	81.7	16.7	1.6	0.8	**0.01**	84.8	12.0	3.2
rs28930679	C	T	0.27	2.4	0.86	60.2	35.0	4.8	1.6	0.15	55.6	33.9	10.5
rs1135791	T	C	**0.05**	4.0	0.08	23.1	58.7	18.2	4.0	0.33	30.6	44.6	24.8
rs3948463	G	A	1.00	3.2	0.64	83.6	15.6	0.8	3.2	0.59	82.8	16.4	0.8
rs3948464	C	T	0.38	0.8	0.46	71.2	24.8	4.0	0	0.19	78.6	18.2	3.2
rs161557	C	T	0.75	2.4	0.83	57.7	36.6	5.7	3.2	0.42	59.8	32.8	7.4
rs745191	G	T	0.22	1.6	0.12	53.2	43.6	3.2	3.2	0.32	45.1	47.5	7.4
rs2061821	C	T	**0.04**	4.8	**0.01**	28.3	60.8	10.9	4.8	**0.003**	40.8	55.0	4.2
rs2061822	C	T	0.25	0.8	**0.01**	32.8	59.2	8	1.6	**0.02**	41.9	53.2	4.8
rs2061824	T	C	0.08	1.6	**0.002**	27.4	62.9	9.7	4.0	**0.01**	39.7	55.4	4.9
rs4075254	A	G	0.06	3.2	**0.01**	27.9	61.5	10.6	4.8	**0.01**	40.0	55.0	5.0
rs4075256	C	T	0.16	2.4	**0.005**	27.6	61.8	10.6	2.4	**0.02**	38.2	55.3	6.5
rs4843074	G	C	0.11	7.1	**0.001**	25.6	64.1	10.3	5.6	**0.02**	37.8	55.5	6.7
rs7177107	G	A	0.15	5.6	0.23	67.2	26.9	5.9	5.6	0.34	64.7	33.6	1.7
rs34434221	A	C	0.27	10.3	0.81	95.6	4.4	0	4.0	0.89	98.3	1.7	0
rs35624420	C	T	1.00	0	0.93	98.4	1.6	0	0	0.93	97.6	2.4	0
rs1138357	G	A	0.94	3.2	0.52	56.6	35.2	8.2	2.4	0.56	58.5	34.1	7.4
rs1138358	T	G	1.00	0.8	0.46	57.6	34.4	8.0	0.8	0.61	58.4	34.4	7.2
rs3826007	G	A	0.64	7.1	0.38	57.3	34.2	8.5	4.8	0.73	61.7	32.5	5.8
rs3735485	C	T	0.74	3.2	0.99	73.8	23.8	2.4	3.2	0.27	77.0	19.7	3.3
rs7792760	G	A	1.00	0	0.20	74.6	21.4	4.0	0.8	0.44	75.2	21.6	3.2
rs2173904	G	C	0.60	2.4	0.65	33.3	46.4	20.3	7.1	0.72	28.0	47.5	24.5
rs2270891	G	T	1.00	1.6	0.49	92.7	6.5	0.8	2.4	0.49	92.7	6.5	0.8

Observed frequencies of genotypes in patients and donors separately. The minor alleles were defined as the alleles with the lowest frequency, whereas the major alleles were defined as the alleles with the highest frequency. Differences in genotype distribution between patients and donors for each nsSNP were assessed by Fisher's exact test. P values <.05 are in bold type.

and Figure 2B shows the distribution according to predicted miHAs in the GVH direction.

Effect of Number of Predicted miHAs in the GVH Direction on Outcome

There was a median of 3 predicted miHAs per patient–donor pair (range, 0-16). A total of 215 miHAs were predicted for the HLA types and nsSNP differences represented in our cohort (Table 3), and 172 of these matched at least one patient. Patients with >3 predicted mismatched miHAs in the GVH direction had a significantly lower probability of 5-year OS and PFS and a higher probability of 5-year TRM compared with patients with ≤3 predicted miHAs (Table 6 and Figure 3A and B). The presence of >3 predicted miHAs also was a significant risk factor associated with 5-year OS, PFS, and TRM in both the unadjusted and adjusted Cox regression models (Table 6 and Table 7). No association between the number of miHAs and any other clinical outcome parameter was observed. Other cutoffs besides the median of 3 predicted miHAs were tested as well. The difference in OS and PFS was significant for all cutoffs between 0 and 6 predicted miHAs (data not shown). The same was true for TRM, with the exception of a cutoff of 1 predicted miHA ($P = .07$) (data not shown). The probability of 5-year OS showed a successive decrease with 0, 1-2, and >2 predicted miHAs per patient (hazard ratio [HR], 2.4; 95% confidence interval [CI], 1.5-4.0; $P = .0005$), implying a miHA dosage effect (Figure 4). Patients with any predicted miHAs had a significantly lower 5-year OS (46% vs 93%; HR, 8.1; 95% CI, 1.9-34; $P = .60\ 10^{-3}$) than patients with 0 predicted miHAs. Apart from the association between the number of predicted miHAs in a patient–donor pair and outcome, some protein-, nsSNP-, and predicted minor-specific associations with OS, PFS, or TRM were observed. The presence of any miHAs in SP110 and AKAP13 (Table 8), patient homozygosity for the minor allele of 3 nsSNPs in tight LD in AKAP13 (rs2061821, rs2061822, rs4075254) (Table 5 and 8),

Table 5. Pairwise Linkage Disequilibrium, Expressed as R^2 between AKAP13 Polymorphisms out of HWE, in the Patient and Donor Populations

Patient/Donor	Rs2061821	Rs2061822	Rs2061824	Rs4075254	Rs4075256
Rs2061822	0.85 / 0.97				
Rs2061824	1/1	0.86 /1			
Rs4075254	1/1	0.86 /1	1/1		
Rs4075256	1/1	0.86 /1	1/1	1/1	
Rs4843074	1/1	0.85 /1	1/1	1/1	1/1

Table 6. Univariate Analyses of the Effect of Number of Predicted miHAs on Different Clinical Outcome Parameters after 5 Years

Parameter	nsSNP Differences ≤ 4 vs > 4*	HR (95% CI)	*P*	Predicted miHAs ≤3 vs >3†	HR (95%CI)	*P*
OS	66.3 % vs 48.9 %	1.7 (0.9–3.1)	.09	70.1 % vs 42.2 %	2.3 (1.2–4.2)	**.0060**
PFS	54.5 % vs 43.2 %	1.5 (0.9–2.6)	.13	58.3 % vs 36.7 %	2.0 (1.2–3.5)	**.0082**
TRM	12.3 % vs 34.0 %	2.2 (0.9–5.5)	.09	9.9 % vs 39.2 %	3.4 (1.3–8.9)	**.0094**
RRM	24.4 % vs 25.9 %	1.3 (0.6–3.0)	.63	22.2 % vs 30.7 %	1.7 (0.7–3.8)	.39
RI	37.3 % vs 30.6 %	1.2 (0.6–2.3)	.76	34.8 % vs 34.6 %	1.5 (0.8–2.9)	.45
Acute GVHD grade II-IV	72.7 % vs 75.7 %	1.0 (0.7–1.5)	.75	71.8 % vs 76.9 %	1.1 (0.7–1.7)	.95
Acute GVHD grade III-IV	21.2 % vs 20.1 %	1.0 (0.4–2.1)	.82	19.6 % vs 22.0 %	1.2 (0.5–2.5)	.82
Extensive chronic GVHD	62.3 % vs 63.7 %	0.9 (0.5–1.6)	.41	62.2 % vs 63.8 %	1.0 (0.6–1.7)	.36

HR, hazard ratio; CI, confidence interval; *P* values <.05 are in bold type.
*Patient–donor pairs are divided into those with ≤4 or >4 nsSNP differences in the GVH direction.
†Patient–donor pairs are divided into those with ≤3 or >3 predicted miHAs.

and 4 predicted miHAs were individually associated with outcome (Table 8). The multiple comparison penalty paid in these analyses increases the Bonferroni-corrected *P* values to well above the .05 threshold. According to dbMinor [15], proteins AKAP13 and KIAA0020 are classified as broadly expressed, whereas SP110, HMHB1, BCL2A1, and MYO1G are classified as hematopoietically expressed. No tissue-specific effect was observed when dividing patients into those with predicted miHAs only from hematopoietically expressed proteins, only from broadly expressed proteins, or from both kinds of proteins (5-year OS, 47% vs 49% vs 41%; $P = .95$).

DISCUSSION

To the best of our knowledge, the present study is the first to investigate the association between the number of predicted miHAs in known miHA source proteins and clinical outcome after matched allogeneic HCT with NMA conditioning. By identifying nsSNP differences and using an Artificial Neural Network tool (*NetMHCpan*) 172 patient–donor specific miHAs were predicted. Compared with the known HLA-A and -B binding miHAs (n = 19; source dbMinor [15]), this represents an almost 10-fold increase, suggesting that the investigated miHA source proteins contain additional miHAs that have yet to be identified. Among the predicted miHAs, 6 were already in the dbMinor database [15]: HA-3 (VTEPGTAQY), HA-8 (RTLDKVLEV), HB-1H (EEKRGSLHVW), HB-1Y (EEKRGSLYVW), ACC-1 (DYLQYVLQI), and ACC2 (KEFEDDIINW). The dbMinor database currently contains 29 miHAs, of which 10 originate from the Y chromosome and thus were not considered in this study. We predicted only 6 of the remaining 19 previously identified HLA-A and -B binding miHAs, because the corresponding nsSNP failed genotyping (6 miHAs), the rs number was not listed in dbSNP (2 miHAs), or the miHA was not caused by an nsSNP (5 miHAs).

In line with the greater degree of genetic variation between unrelated individuals, significantly fewer nsSNP differences and predicted miHAs in the GVH direction were observed with sibling patient–donor pairs compared with matched unrelated pairs. The patient–donor relationship did not significantly influence the transplantation outcome, however. When

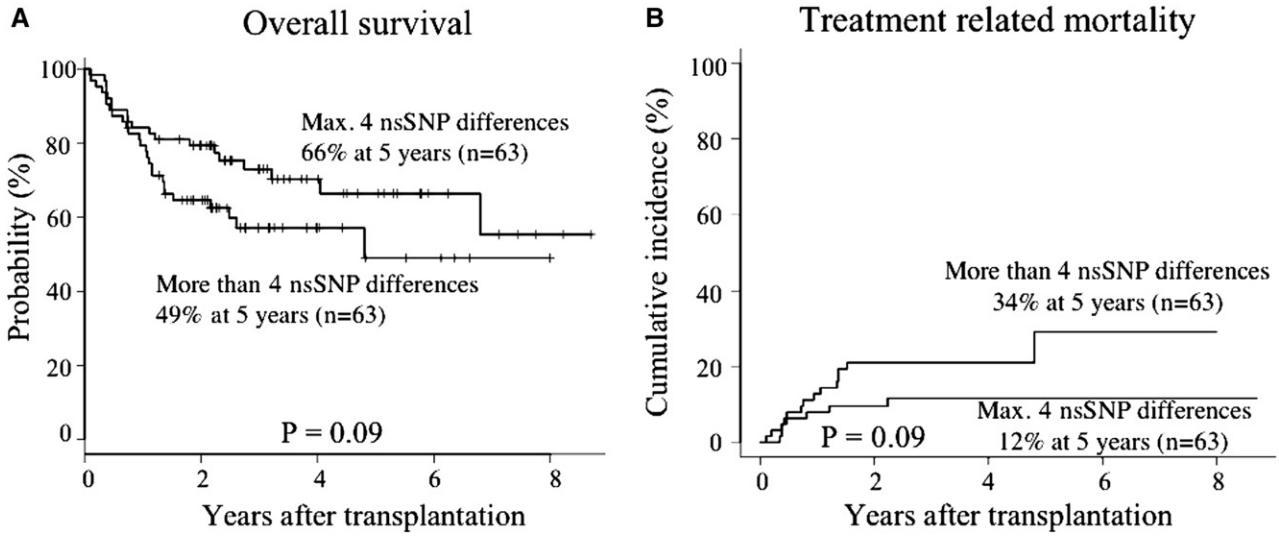

Figure 1. Probability of OS (A) and cumulative incidence of TRM (B) stratified according to the median number of nsSNP differences in the GVH direction.

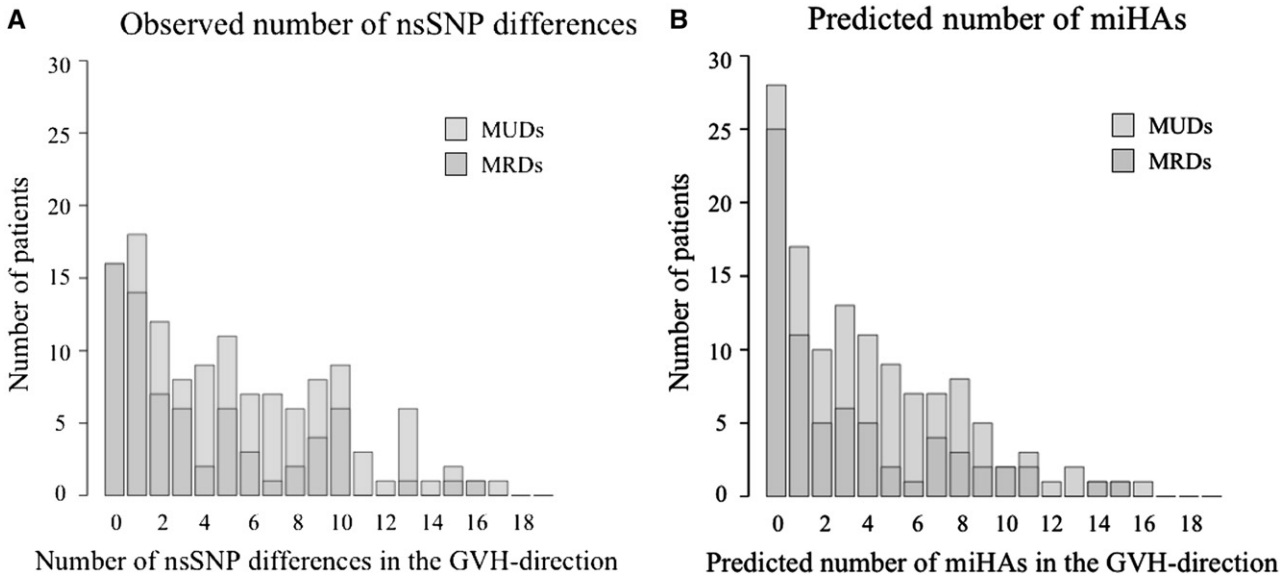

Figure 2. Histograms showing the distribution of patients by number of observed nsSNP differences in the GVH direction (A) and by number of predicted miHAs in the GVH direction (B). MUD, matched unrelated donor; MRD, matched related donor.

restricting the analysis to nsSNP differences in the GVH direction, it was possible to observe only a *trend* toward superior transplantation outcomes in patients with few nsSNP differences. But, when HLA restrictions were also taken into account by using miHA predictions, we were able to show that the presence of the median of ≤3 miHA disparities within a patient–donor pair, was a significant independent factor associated with a higher probability of both OS and PFS and lower risk of TRM. Although the group of patients with 0 predicted miHAs had the highest OS compared with all other patients, the median was chosen to provide an unbiased level for dichotomization in the analyses, because the very few (2) events in the group of patients with 0 predicted miHAs could affect the reliability of the statistical comparison when using 0 predicted miHAs as the level of dichotomization.

These data suggest that the outcome of HCT depends on matching donor and recipient for HLA-restricted miHAs, rather than on the mere matching of nsSNPs. No association between the number of predicted miHAs and aGVHD or cGVHD was observed. Although GVHD is considered one of the main causes of TRM [14], TRM also encompasses patients who succumbed to infection. Because it is unlikely that the number of predicted miHAs is associated with the risk of infection without affecting the incidence of GVHD, the discrepancy between TRM and GVHD most likely results from insufficient study power. Given that no associations with relapse-related

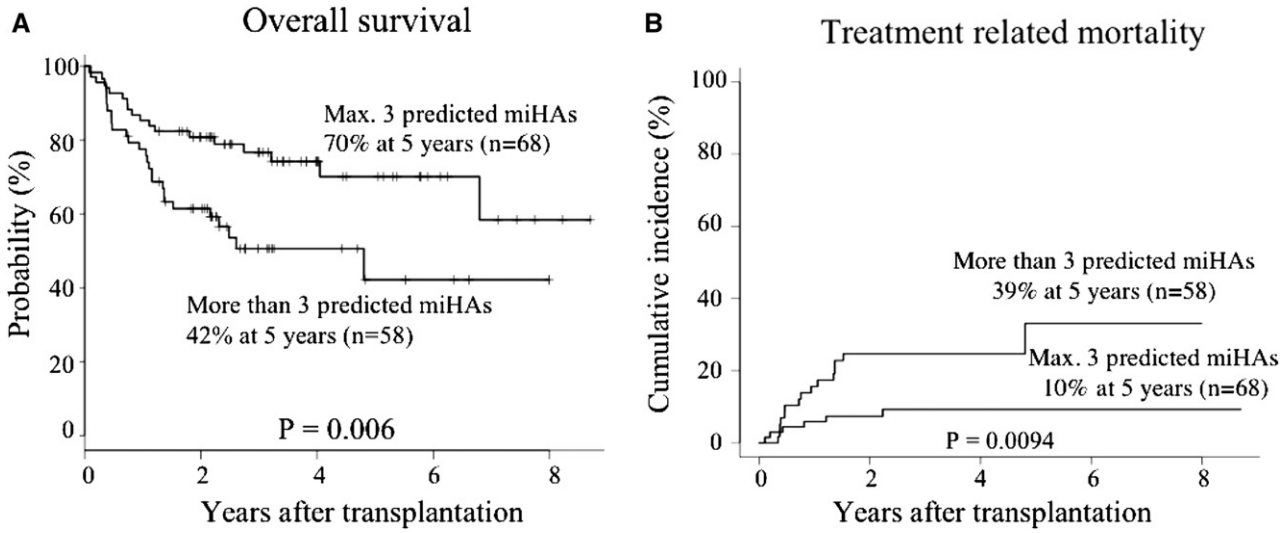

Figure 3. Probability of OS (A) and cumulative incidence of TRM (B) stratified according to the median number of predicted miHAs in the GVH direction within a patient–donor pair.

Table 7. Multivariate Cox Regression Analysis of the Association of Number of Predicted miHAs with 5-Year Transplantation Outcome

Outcome	Covariate		HR	95% CI	*P*
OS	Number of predicted miHAs	≤3	Ref.		
		>3	2.2	1.2–4.0	**.014**
	Acute GVHD grade III-IV	Absence	Ref.		
		Presence	3.2	1.7–5.9	**<.001**
	Extensive chronic GVHD	Absence	Ref.		
		Presence	0.8	0.35–2.0	.67
PFS	Number of predicted miHAs	≤3	Ref.		
		>3	2.0	1.1–3.4	**.014**
	Acute GVHD grade III-IV	Absence	Ref.		
		Presence	2.7	1.5–4.8	**.001**
	Extensive chronic GVHD	Absence	Ref.		
		Presence	1.3	0.60–2.7	.525
TRM	Number of predicted miHAs	≤3	Ref.		
		>3	4.5	1.7–12.3	**.003**
	Patient age	≤40 years	Ref.		
		>40 years	3.8	0.5–29	.198
	Donor age	≤40 years	Ref.		
		>40 years	2.2	0.8–5.9	.126
	Acute GVHD grade III-IV	Absence	Ref.		
		Presence	4.4	1.9–10.6	**.001**
	Extensive chronic GVHD	Absence	Ref.		
		Presence	1.3	0.4–4.9	.669

OS indicates overall survival; PFS, progression-free survival; TRM, treatment-related mortality; GVHD, graft-versus-host disease; miHA, minor histocompatibility antigen.
Covariates were included in the final models only if they changed the estimate of the main variable by at least 10% or were significantly associated with outcome in pairwise analyses. *P* values <.05 are in bold type.

outcome measures were observed, our data suggest that the presence of many miHAs confers an increased risk of death rather than inducing the beneficial GVT effect, implying that mismatching for most miHAs results in decreased survival.

If the extent of interindividual genetic variation and HLA diversity is taken into account, then the current study assesses only a very limited subset of all possible predicted miHAs. Because many of the predicted miHAs likely will not initiate cytotoxic T cell responses because of immunodominance issues [43], it is of interest that the limited subset of miHAs predicted in our study was associated with transplantation outcome. This could be explained by the assessed proteins being the source of most relevant miHAs for transplantation (which we consider unlikely), or by the degree of miHA disparity in these proteins being a surrogate marker for the total genome-wide and HLA-wide miHA disparity within each patient–donor pair. If the degree of miHA disparity in our study only represents a proxy for the real patient–donor discrepancy, the exact level of dichotomization also becomes less important compared with making a distinction between few or many disparities. Apart from the impact of the predicted miHAs on transplantation outcome, the possibility that factors such as the functional aspects of the nsSNPs cannot be excluded, and the general heterogeneity of the patient cohort also could influence the outcome in our cohort.

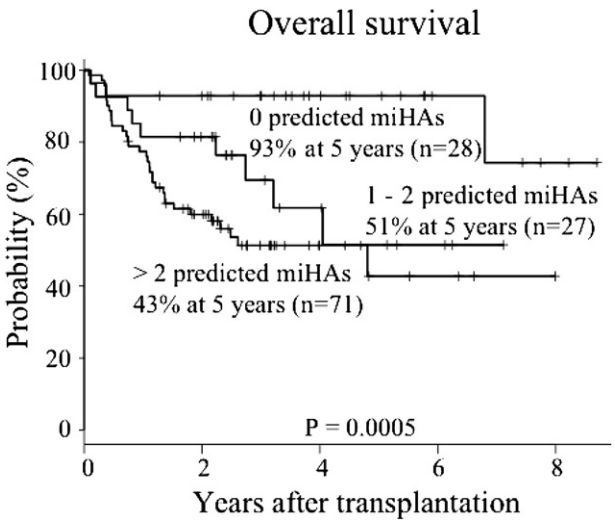

Figure 4. Probability of OS stratified according to number (0, 1-2, or >2) of predicted miHAs in the GVH direction.

In several studies of single or very few patients, the identification and presence of miHA-specific T cells has been associated with remissions of chronic myelogenous leukemia [11,44] or rejection [7]. However, larger studies of single or multiple miHA mismatches restricted to one or few HLA types in patients undergoing allogeneic HCT with sibling donors have uniformly been associated with GVHD without any association with RRM [8-10,45-47]. In line with our study, this suggests that in general, miHA mismatch is not beneficial. In contrast, no association between miHA disparity and outcome was observed in a single study of 730 unrelated HLA-matched allogeneic HCTs [48], possibly illustrating the impact of increased non-HLA genetic variation confounding the observations.

Table 8. Single Predicted miHAs, Predicted miHAs around a Single nsSNP, and Predicted miHAs from a Single Protein with a Significant Impact on OS, PFS, or TRM

Protein	nsSNP Difference	Predicted miHA	Relevant Patients	*P* (OS)	*P* (PFS)	*P* (TRM)
SP110			45	**.025**	.058	.28
BCL2A1			29	.16	.11	.063
	rs1138357		25	.28	.30	.27
		YLQYVLQI*	13	.40	.90	**.025**
		RLAQDYLQYV	13	.40	.90	**.025**
	rs1138358		24	.17	.088	**.013**
		VLQKVAFSV	14	.51	.69	**.040**
AKAP13			49	**.041**	.062	.37
	rs2061821		26	.082	**.047**	.14
		LVMEPGTAQY†	13	**.0062**	**.0022**	**.0040**
	rs2061822		21	.41	.29	**.03**
	rs4075254		19	.20	.20	**.011**

OS indicates overall survival; PFS, progression-free survival; TRM, treatment-related mortality; GVHD, graft-versus-host disease; nsSNP, nonsynonymous single nucleotide polymorphism; miHA, minor histocompatibility antigen.
nsSNPs are listed under the protein in which they occur, and predicted miHAs are listed under the nsSNP that causes the miHA. Bold type denotes $P < .05$ (not corrected for multiple testing).
*Similar to the known miHA *ACC-1*:DYLQYVLQI.
†Similar to the known miHA *HA-3*:VTEPGTAY.

miHAs have been classified into those with a restricted tissue expression encompassing tissues of hematopoietic origin and a broad tissue expression including nonhematopoietic tissues such as skin, gut, and liver [49]. It has been suggested that miHAs with a restricted tissue expression would result in GVT effects without deleterious GVHD, because the GVHD elicited by such miHAs would only result in the removal of normal recipient hematopoiesis. In contrast, miHAs with a broad tissue expression would carry the risk of inducing potentially life-threatening GVHD. Among the 6 miHA source proteins showing nsSNP variation in the GVH direction, only AKAP13 and KIAA0020 (accounting for a total of 87 predicted miHAs in our cohort) are classified as broadly expressed (source: dbMinor [15]). The other 4 proteins—SP110, HMHB1, BCL2A1, and MYO1G, accounting for a total of 128 predicted miHAs—have restricted tissue expression. However, a comparison of HCT outcome in patients with predicted miHAs from broadly expressed proteins, proteins restricted to hematopoietic tissue, or both types of proteins showed no significant differences between the 3 patient groups, challenging either the experimental results on which the classification is based or the theoretical framework for separating GVHD and GVT effects [49]. Alternatively, in addition to creating miHAs, the functional aspects of the nsSNPs also could decisively influence hematopoiesis and thus transplantation outcome. Therefore, it is possible that the current classification of miHAs is simplistic and will require revision as our understanding grows.

Several limitations apply to the current study. Predictions were limited to HLA-A and -B molecules, because *NetMHCpan* is most accurate for these [50]. miHA predictions were planned for peptides surrounding 53 nsSNPs in 11 different non–Y chromosomal proteins. However, technical limitations because of both the genetic sequence surrounding the nsSNPs and the nature of the SNPstream genotyping platform limited the number of successfully genotyped nsSNPs to 31. Because the 53 nsSNPs at best only are surrogate markers for the common genetic variation between individuals, and because the 53 planned and 31 successfully genotyped nsSNPs probably represent similar fractions of the numerous potential nsSNPs in the entire genome, no further effort was made to pursue the genotype of the failed 22 nsSNPs. Although most genotypes adhered to HWE, 7 nsSNPs (6 of which were in strong LD) departed significantly. These observations are likely because of small sample size, because genotypes were confirmed by extensive resequencing, and because adherence to HWE was observed when a control population of 96 healthy blood donors was included. Furthermore, the significantly different distribution of rs2061821 and rs1135791 genotypes between patients and donors also was considered an artifact ascribed to the small study population, rather than a true association with disease susceptibility.

In conclusion, the current study presents a feasible method for large-scale *in silico* prediction of novel HLA-A– and -B–restricted miHAs incorporating any patient–donor HLA types. Although the functional aspects of the predicted miHAs are unknown and the study is purely descriptive, our findings suggest that the level of predicted miHA discrepancy between patient and donor could be associated with transplantation outcome. If these observations were to be validated in independent cohorts, miHA predictions specific for each patient–donor pair could have a place in future risk stratification and possibly in guiding donor selection and therapy. With the current

advancements in microarray-based genotyping, whole genome sequencing, and *in silico* modeling, a genome-wide and HLA-wide approach is within reach. miHA predictions could be expanded from encompassing only a few nsSNPs to all known nonsynonymous genetic variations and both class I and II HLA molecules, providing a unique miHA map of each patient–donor pair. This likely would enhance the prognostic value of the method in selecting the most optimal donor in those cases for which more than one 10/10 allele-matched donor was found using the current donor selection procedures. Apart from the prognostic application, the large-scale miHA predictions also could function as a powerful tool in selecting candidate miHAs involved in the GVT effect and GVHD for further evaluation in *in vitro* and *in vivo* experiments.

ACKNOWLEDGMENTS

They thank Anne Bjørlig for her excellent technical assistance and the staff of the Department of Hematology, Rigshospitalet for their help and support. This work was supported by grants from the Danish Cancer Society, the Danish Medical Research Council, the Novo Nordisk Research Foundation, Rigshospitalet, and the Lundbeck foundation, and a grant from the Danish Research Council for Technology and Production Sciences (project "Disease Gene Finding, Somatic Mutations, and Vaccine Design;"principal funding recipient, Søren Brunak). The genotyping equipment was sponsored by the John and Birthe Meyer Foundation.

AUTHORSHIP STATEMENT

Malene Erup Larsen and Brian Kornblit contributed equally to this work.

Financial disclosure: The authors have no conflicts of interest to disclose.

REFERENCES

1. Hambach L, Spierings E, Goulmy E. Risk assessment in haematopoietic stem cell transplantation: minor histocompatibility antigens. *Best Pract Res*. 2007;20:171-187.
2. Spaapen R, Mutis T. Targeting haematopoietic-specific minor histocompatibility antigens to distinguish graft-versus-tumour effects from graft-versus-host disease. *Best Pract Res*. 2008;21: 543-557.
3. Kawase T, Akatsuka Y, Torikai H, et al. Alternative splicing due to an intronic SNP in HMSD generates a novel minor histocompatibility antigen. *Blood*. 2007;110:1055-1063.
4. Mullally A, Ritz J, Beyond HLA: the significance of genomic variation for allogeneic hematopoietic stem cell transplantation. *Blood*. 2007;109:1355-1362.
5. Murata M, Warren EH, Riddell SR. A human minor histocompatibility antigen resulting from differential expression due to a gene deletion. *J Exp Med*. 2003;197:1279-1289.
6. Goulmy E, Termijtelen A, Bradley BA, et al. Y-antigen killing by T cells of women is restricted by HLA. *Nature*. 1977;266:544-545.
7. Voogt PJ, Fibbe WE, Marijt WA, et al. Rejection of bone-marrow graft by recipient-derived cytotoxic T lymphocytes against minor histocompatibility antigens. *Lancet*. 1990;335:131-134.
8. Goulmy E, Schipper R, Pool J, et al. Mismatches of minor histocompatibility antigens between HLA-identical donors and recipients and the development of graft-versus-host disease after bone marrow transplantation. *N Engl J Med*. 1996;334: 281-285.
9. Perez-Garcia A, De la Camara R, Torres A, et al. Minor histocompatibility antigen HA-8 mismatch and clinical outcome after HLA-identical sibling donor allogeneic stem cell transplantation. *Haematologica*. 2005;90:1723-1724.
10. Akatsuka Y, Warren EH, Gooley TA, et al. Disparity for a newly identified minor histocompatibility antigen, HA-8, correlates with acute graft-versus-host disease after haematopoietic stem cell transplantation from an HLA-identical sibling. *Br J Haematol*. 2003;123:671-675.
11. de Rijke B, van Horssen-Zoetbrood A, Beekman JM, et al. A frameshift polymorphism in P2X5 elicits an allogeneic cytotoxic T lymphocyte response associated with remission of chronic myeloid leukemia. *J Clin Invest*. 2005;115:3506-3516.
12. van Bergen CA, Kester MG, Jedema I, et al. Multiple myeloma–reactive T cells recognize an activation-induced minor histocompatibility antigen encoded by the ATP-dependent interferon-responsive (ADIR) gene. *Blood*. 2007;109:4089-4096.
13. Stumpf AN, van der Meijden ED, van Bergen CA, et al. Identification of 4 new HLA-DR-restricted minor histocompatibility antigens as hematopoietic targets in antitumor immunity. *Blood*. 2009;114:3684-3692.
14. Ferrara JL, Levine JE, Reddy P, et al. Graft-versus-host disease. *Lancet*. 2009;373:1550-1561.
15. Spierings E, Drabbels J, Hendriks M, et al. A uniform genomic minor histocompatibility antigen typing methodology and database designed to facilitate clinical applications. *PloS One*. 2006;1. e42.
16. Spierings E, Brickner AG, Caldwell JA, et al. The minor histocompatibility antigen HA-3 arises from differential proteasome-mediated cleavage of the lymphoid blast crisis (Lbc) oncoprotein. *Blood*. 2003;102:621-629.
17. Warren EH, Vigneron NJ, Gavin MA, et al. An antigen produced by splicing of noncontiguous peptides in the reverse order. *Science*. 2006;313:1444-1447.
18. Dolstra H, Fredrix H, Maas F, et al. A human minor histocompatibility antigen specific for B cell acute lymphoblastic leukemia. *J Exp Med*. 1999;189:301-308.
19. Akatsuka Y, Nishida T, Kondo E, et al. Identification of a polymorphic gene, BCL2A1, encoding two novel hematopoietic lineage-specific minor histocompatibility antigens. *J Exp Med*. 2003;197:1489-1500.
20. Kawase T, Nannya Y, Torikai H, et al. Identification of human minor histocompatibility antigens based on genetic association with highly parallel genotyping of pooled DNA. *Blood*. 2008; 111:3286-3294.
21. Kamei M, Nannya Y, Torikai H, et al. HapMap scanning of novel human minor histocompatibility antigens. *Blood*. 2009; 113:5041-5048.
22. Ogawa S, Matsubara A, Onizuka M, et al. Exploration of the genetic basis of GVHD by genetic association studies. *Biol Blood Marrow Transplant*. 2008;15:39-41.
23. Robinson J, Waller MJ, Parham P, et al. IMGT/HLA Database: a sequence database for the human major histocompatibility complex. *Nucleic Acids Res*. 2001;29:210-213.
24. Robinson J, Waller MJ, Parham P, et al. IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. *Nucleic Acids Res*. 2003;31:311-314.
25. Robinson J, Waller MJ, Fail SC, et al. The IMGT/HLA database. *Nucleic Acids Res*. 2009;37:D1013-D1017.
26. UniProt. Human polymorphisms and disease mutations (humsavar, release 57.12, December 2009), http://www.uniprot.org/docs/humsavar.
27. Kessler JH, Melief CJ. Identification of T-cell epitopes for cancer immunotherapy. *Leukemia*. 2007;21:1859-1874.

28. Schuler MM, Donnes P, Nastke MD, et al. SNEP: SNP-derived epitope prediction program for minor H antigens. *Immunogenetics*. 2005;57:816-820.
29. Halling-Brown M, Quartey-Papafio R, Travers PJ, et al. SiPep: a system for the prediction of tissue-specific minor histocompatibility antigens. *Int J Immunogenet*. 2006;33:289-295.
30. Deluca DS, Eiz-Vesper B, Ladas N, et al. High-throughput minor histocompatibility antigen prediction. *Bioinformatics*. 2009;(25):2411-2417.
31. Nielsen M, Lundegaard C, Blicher T, et al. *NetMHCpan*, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PloS One*. 2007;2:e796.
32. Lin HH, Ray S, Tongchusak S, et al. Evaluation of MHC class I peptide binding prediction servers: applications for vaccine research. *BMC Immunol*. 2008;9:8.
33. Peters B, Bui HH, Frankild S, et al. A community resource benchmarking predictions of peptide binding to MHC-I molecules. *PLoS Comp Biol*. 2006;2:e65.
34. Kahl C, Storer BE, Sandmaier BM, et al. Relapse risk in patients with malignant diseases given allogeneic hematopoietic cell transplantation after nonmyeloablative conditioning. *Blood*. 2007;110:2744-2748.
35. Kornblit B, Masmas T, Madsen HO, et al. Haematopoietic cell transplantation with non-myeloablative conditioning in Denmark: disease-specific outcome, complications and hospitalization requirements of the first 100 transplants. *Bone Marrow Transplant*. 2008;41:851-859.
36. Sullivan KM. Graft-vs-host disease. In: Blume KG, Forman SJ, Appelbaum FR, editors. *Thomas' Hematopoietic Cell Transplantation*, 3rd ed. Malden, MA: Blackwell; 2003. p: 635-664.
37. Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res*. 2007;35:D61-D65.
38. Smigielski EM, Sirotkin K, Ward M, et al. dbSNP: a database of single nucleotide polymorphisms. *Nucleic Acids Res*. 2000;28: 352-355.
39. Consortium. The International HapMap Project. *Nature*. 2003; 426:789-796.
40. Kornblit B, Munthe-Fog L, Petersen SL, et al. The genetic variation of the human HMGB1 gene. *Tissue Antigens*. 2007;70: 151-156.
41. Gray RJ. A class of *k*-sample tests for comparing the cummulative incidence of a competing risk. *Ann Stat*. 1988;16: 1141-1154.
42. Gooley TA, Leisenring W, Crowley J, et al. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Stat Med*. 1999;18:695-706.
43. Roopenian D, Choi EY, Brown A. The immunogenomics of minor histocompatibility antigens. *Immunol Rev*. 2002;190:86-94.
44. Marijt WA, Heemskerk MH, Kloosterboer FM, et al. Hematopoiesis-restricted minor histocompatibility antigens HA-1– or HA-2–specific T cells can induce complete remissions of relapsed leukemia. *Proc Natl Acad Sci USA*. 2003;100:2742-2747.
45. Tseng LH, Lin MT, Hansen JA, et al. Correlation between disparity for the minor histocompatibility antigen HA-1 and the development of acute graft-versus-host disease after allogeneic marrow transplantation. *Blood*. 1999;94:2911-2914.
46. Grumet FC, Hiraki DD, Brown BWM, et al. CD31 mismatching affects marrow transplantation outcome. *Biol Blood Marrow Transplant*. 2001;7:503-512.
47. Cavanagh G, Chapman CE, Carter V, et al. Donor CD31 genotype impacts on transplant complications after human leukocyte antigen-matched sibling allogeneic bone marrow transplantation. *Transplantation*. 2005;79:602-605.
48. Spellman S, Warden MB, Haagenson M, et al. Effects of mismatching for minor histocompatibility antigens on clinical outcomes in HLA-matched, unrelated hematopoietic stem cell transplants. *Biol Blood Marrow Transplant*. 2009;15:856-863.
49. Bleakley M, Riddell SR. Molecules and mechanisms of the graft-versus-leukaemia effect. *Nat Rev*. 2004;4:371-380.
50. Hoof I, Peters B, Sidney J, et al. *NetMHCpan*, a method for MHC class I binding prediction beyond humans. *Immunogenetics*. 2009; 61:1-13.