REGULAR ARTICLE

blood advances

A novel predictive approach for GVHD after allogeneic SCT based on clinical variables and cytokine gene polymorphisms

Carolina Martínez-Laperche,[1,2,*] Elena Buces,[1,2,*] M. Carmen Aguilera-Morillo,[3,*] Antoni Picornell,[2,4] Milagros González-Rivera,[4,5] Rosa Lillo,[3] Nazly Santos,[6] Beatriz Martín-Antonio,[7] Vicent Guillem,[8] José B. Nieto,[9] Marcos González,[10] Rafael de la Cámara,[11] Salut Brunet,[12] Antonio Jiménez-Velasco,[13] Ildefonso Espigado,[14] Carlos Vallejo,[15] Antonia Sampol,[16] José María Bellón,[2] David Serrano,[1,2] Mi Kwon,[1,2] Jorge Gayoso,[1,2] Pascual Balsalobre,[1,2] Álvaro Urbano-Izpizua,[7] Carlos Solano,[8] David Gallardo,[6] José Luis Díez-Martín,[1,2,17] Juan Romo,[3] and Ismael Buño,[1,2,18] on behalf of the GVHD/Immunotherapy Committee of the Spanish Group for Hematopoietic Transplantation

[1]Department of Hematology, Hospital General Universitario (H.G.U.) Gregorio Marañón, Madrid, Spain; [2]Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain; [3]Department of Statistics, Universidad Carlos III de Madrid, Madrid, Spain; [4]Department of Oncology and [5]DNA Sequencing and Genotyping Core Facility, H.G.U. Gregorio Marañón, Madrid, Spain; [6]Department of Hematology, Instituto Catalán de Oncología Hospital Josep Trueta, Girona, Spain; [7]Department of Hematology, Hospital Clínic, Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain; [8]Department of Hematology, Hospital Clínico de Valencia, Valencia, Spain; [9]Department of Hematology, Hospital Universitario Morales Meseguer, Murcia, Spain; [10]Department of Hematology, Hospital de Salamanca, Salamanca, Spain; [11]Department of Hematology, Hospital Universitario de La Princesa, Madrid, Spain; [12]Department of Haematology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain; [13]Department of Hematology, Hospital Regional de Málaga, Málaga, Spain; [14]Department of Hematology, Hospital Universitario Virgen del Rocío, Seville, Spain; [15]Department of Hematology, Hospital Central de Asturias, Oviedo, Spain; [16]Department of Hematology, Hospital Universitario Son Espases, Palma de Mallorca, Spain; [17]Department of Medicine, Facultad de Medicina, Universidad Complutense de Madrid, Madrid, Spain; and [18]Genomics Unit, H.G.U. Gregorio Marañón/Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain.

Key Points

- A risk model using donor and recipient cytokine gene polymorphisms and clinical variables significantly improves GVHD risk stratification.
- The model is useful in identifying patients with low-risk of developing severe GVHD, but results must be confirmed in prospective studies.

Despite considerable advances in our understanding of the pathophysiology of graft-versus-host disease (GVHD), its prediction remains unresolved and depends mainly on clinical data. The aim of this study is to build a predictive model based on clinical variables and cytokine gene polymorphism for predicting acute GVHD (aGVHD) and chronic GVHD (cGVHD) from the analysis of a large cohort of HLA-identical sibling donor allogeneic stem cell transplant (allo-SCT) patients. A total of 25 SNPs in 12 cytokine genes were evaluated in 509 patients. Data were analyzed using a linear regression model and the least absolute shrinkage and selection operator (LASSO). The statistical model was constructed by randomly selecting 85% of cases (training set), and the predictive ability was confirmed based on the remaining 15% of cases (test set). Models including clinical and genetic variables (CG-M) predicted severe aGVHD significantly better than models including only clinical variables (C-M) or only genetic variables (G-M). For grades 3-4 aGVHD, the correct classification rates (CCR1) were: 100% for CG-M, 88% for G-M, and 50% for C-M. On the other hand, CG-M and G-M predicted extensive cGVHD better than C-M (CCR1: 80% vs. 66.7%, respectively). A risk score was calculated based on LASSO multivariate analyses. It was able to correctly stratify patients who developed grades 3-4 aGVHD ($P < .001$) and extensive cGVHD ($P < .001$). The novel predictive models proposed here improve the prediction of severe GVHD after allo-SCT. This approach could facilitate personalized risk-adapted clinical management of patients undergoing allo-SCT.

Introduction

Allogeneic hematopoietic stem-cell transplantation (allo-SCT) is a curative therapeutic approach for patients with hematologic malignancies. Patients undergoing allo-SCT receive a donor graft containing hematopoietic stem cells, as well as various other cell types, including alloreactive T cells. T cells promote

Submitted 10 August 2017; accepted 11 May 2018. DOI 10.1182/bloodadvances.2017011502.

*C.M.-L., E.B., and M.C.A.-M. contributed equally to this work.

The full-text version of this article contains a data supplement.

hematopoietic engraftment, T-cell immunity reconstitution, and mediate graft-versus-leukemia effect, which may prevent tumor relapse. However, donor T cells may also cause graft-versus-host disease (GVHD), which is the main complication after allo-SCT and the most important cause of nonrelapse morbidity and nonrelapse mortality (NRM).[1]

There are 2 forms of GVHD, acute GVHD (aGVHD) and chronic GVHD (cGVHD). aGVHD is a complex process that takes place in 3 phases.[2] In the first phase, conditioning regimen damages host tissues and raises levels of proinflammatory cytokines such as interleukin-1 (IL-1), IL-6, tumor necrosis factor α (TNFα), and interferon-γ (IFN-γ), thus activating host antigen-presenting cells, which stimulate donor T cells. In the second phase, this interaction induces proliferation and differentiation of donor T cells, which in turn leads to rapid intracellular biochemical cascades that induce transcription of genes for many proteins (including cytokines TNFα, IFN-γ, and IL-2) and promote cellular activity. The third effector phase is a complex cascade of both cellular mediators and soluble inflammatory mediators such as TNFα, IFN-γ, IL-1, and nitric oxide, resulting in tissue injury. Although the pathophysiology of cGVHD is less known, significant advances in our understanding have been made in recent years, and it is now evident that the clinical manifestations result from a complex immune disease involving both donor B cells and T cells.[3] The long-standing hypothesis is that cGVHD is similar to an autoimmune disorder.[4] It is well established that the most important risk factor for the development of GVHD is the degree of HLA matching between the recipient and the donor,[5,6] although a significant proportion of patients undergoing transplantation with HLA-identical grafts develop aGVHD[1] and/or cGHVD.[7] Consequently, other non-HLA factors contribute to the development of this complication. Major clinical factors associated with GVHD include patient age, sex of donor/recipient,[8] stem-cell source,[9] GVHD prophylaxis, underlying disease, conditioning regimen,[10] and, for cGVHD, a history of aGVHD.

Genetic differences in non-HLA genes between recipients and donors are also important,[2] and the role of polymorphisms in human minor histocompatibility antigens,[11,12] innate immunity genes,[13-15] genes involved in drug metabolism, and proinflammatory cytokines must be taken into account.[16,17] During the past decade, single-nucleotide polymorphisms (SNPs) have been identified in genes involved in innate and adaptive immune responses, such as cytokines and their receptors, which have a role in the classic cytokine storm of GVHD.[18-21] However, information regarding the diagnostic, prognostic, and predictive significance of these molecules in GVHD is limited. Although clinically useful biomarkers are available, no particular biomarker alone is generally satisfactory in terms of sensitivity or specificity for the diagnosis or prediction of a disease. Therefore, it is important to build biomarker panels and risk models for GVHD. In recent years, many groups have been working in this field. Kim et al[22,23] built a risk model incorporating SNPs and clinical markers to stratify patients and more accurately predict the risk of GVHD in specific organs, Paczesny et al[24] developed protein panels that provide meaningful information to confirm the diagnosis of GVHD in patients at the onset of clinical symptoms of GVHD and provide useful data for prognosis.[25]

After genotyping a panel of polymorphisms in cytokine genes that had been previously associated with aGVHD or cGVHD,[7,18,26] we applied a complex estimation method, the least absolute shrinkage and selection operator (LASSO) procedure,[27] which is able to group optimal predictors from a large set of potential clinical and genetic predictor variables, improving their clinical utility.

Methods

Study design

This retrospective study included 509 patients with hematological malignancies from the Spanish Group for Hematopoietic Transplantation (GETH) who underwent conventional HLA-identical sibling-donor allo-SCT between 1997 and 2010 at 11 Spanish institutions (the mean number of patients from each center was 46.3 [range, 25-134]; supplemental Table 1). The median follow-up for living patients was 14.7 months (range, 2-105.4 months).

Only patients for whom all clinical and genetic data were available (a prerequisite of the LASSO procedure) were finally included in the analysis (n = 359; Table 1). Patients who died before day +100 without aGVHD (n = 96) or day +200 without cGVHD (n = 154) were excluded from the LASSO multivariate analyses, which therefore included 263 patients for aGVHD and 207 patients for cGVHD modeling.

The study was approved by the ethics committee of Hospital General Universitario Gregorio Marañón, and all recipients and donors provided written informed consent according to the Declaration of Helsinki.

Polymorphism genotyping

DNA was obtained from EDTA anticoagulated peripheral blood samples collected at the pretransplantation evaluation included in the GETH DNA bank.

A total of 25 SNPs in 12 cytokine genes (supplemental Table 2) were selected for their potential role in the pathogenesis of GVHD or in any autoimmune disease in other studies.[7,15,20] SNPs were genotyped using the MALDI-TOF MassARRAY iPLEX Gold platform (Sequenom, San Diego, CA) at CeGen (National Genotyping Centre, Santiago de Compostela, Spain).

Clinical and genetic variables

Three predictive models were constructed for each of the outcomes considered (grade 2-4 aGVHD, grade 3-4 aGVHD, cGVHD, extensive cGVHD, and NRM): 1 with clinical variables alone (C-M), 1 with genetic variables alone (G-M), and 1 with both clinical and genetic variables (CG-M).

Clinical variables included were donor and recipient sex, recipient age, female donor/male recipient, stem-cell source, conditioning regimen, total body irradiation (TBI)–containing regimen, and disease. Previous grade 2 to 4 aGVHD was included in the analysis of cGVHD. Genetic variables (supplemental Table 1) were assessed for donors and recipients, and 4 different models of transmission were considered (recessive, dominant, codominant, and additive). Therefore, 8 genetic variables were built for each SNP.

GVHD classification and clinical data collection were performed at the moment of GVHD diagnosis by the attending physician following the 1994 Consensus Conference on aGVHD grading[28] and the National Institutes of Health criteria for diagnosis and staging of cGVHD.[29]

Table 1. Clinical characteristics of patients

Characteristic	All cohort (n = 359)	aGVHD cohort (n = 263)	cGVHD cohort (n = 207)
Follow-up, mo			
Median (range)	32 (4-105)	23 (1-104)	33 (4-104)
Age, y			
Median (range)	45 (0-68)	44 (0-68)	42 (0-68)
Patient sex, n (%)			
Male	225 (63)	166 (63)	131 (63)
Female	134 (37)	97 (37)	76 (37)
Donor sex, n (%)			
Male	201 (56)	147 (56)	115 (55)
Female	158 (44)	116 (44)	92 (45)
Donor/recipient sex, n (%)			
Female donor/male recipient	91 (25)	69 (26)	53 (26)
Disease, n (%)			
AML	116 (32)	73 (28)	52 (25)
ALL	49 (13.5)	37 (14)	30 (14.5)
NHL, HD	54 (15)	39 (15)	29 (14)
Myelofibrosis, MDS	34 (9.5)	25 (9.5)	19 (9)
MM	33 (9)	25 (9.5)	18 (9)
Other (CML, AA, etc)	73 (20)	25 (9.5)	59 (28.5)
Conditioning regimen, n (%)			
Myeloablative	253 (70)	170 (65)	144 (70)
Reduced intensity	106 (30)	93 (35)	63 (30)
With TBI	94 (26)	71 (27)	59 (28.5)
GVHD prophylaxis, n (%)			
CsA + MTX	290 (81)	220 (83.5)	165 (79.7)
CsA ± steroids	8 (2)	4 (1.5)	3 (1.3)
Others	61 (17)	39 (15)	39 (18)
Stem-cell source, n (%)			
Peripheral blood	250 (69.6)	178 (68)	133 (64)
Bone marrow	109 (30.4)	85 (32)	74 (36)
aGVHD grade, n (%)			
2-4	115 (32)	74 (28)	48 (23)
3-4	50 (14)	29 (11)	21 (10)
cGVHD, n (%)			
Global	161 (45)	142 (54)	109 (53)
Extensive	100 (28)	84 (32)	63 (30)
Mortality, n (%)			
Incidence	86 (24)	27	31
Relapse, n (%)			
Incidence	105 (30)	74 (28)	53 (26)
Death, n (%)			
Incidence	154 (43)	96 (37)	54 (26)

Table lists characteristics used to build the predictive models with LASSO (aGVHD, n = 263; cGVHD, n = 207; NRM, n = 359).

AA, aplastic anemia; ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; CML, chronic myeloid leukemia; CsA, cyclosporin A; HD, Hodgkin disease; MDS, myelodysplastic syndrome; MM, multiple myeloma; MTX, methotrexate; NHL, non-Hodgkin lymphoma;

Table 2. Summary of the most important variables to predict GVHD and NRM in univariate analysis

Variable	Polymorphism	D/R	OR*				
			aGVHD 2-4	aGVHD 3-4	cGVHD	Extensive cGVHD	NRM
Clinical							
Recipient age >45 y					H	H	
Recipient sex, male vs female						H	
Female donor/male recipient							H
MDS					H		
Conditioning, RIC vs MAC					H	H	
Conditioning regimen, no TBI vs TBI				H			
Stem-cell source, PB vs BM					H	H	
aGVHD, grade 2-4					H	H	
Genetic							
IL-1A	rs1800587	D			H		
IL-1B	rs1143627	R	H				
		D			L		
	rs16944	R	H				
		D			H		
	rs1143634	R			H		
		D			H		
IL-2	rs2069762	R				H	
IL-6	rs1800795	R		H			
IL-17A	rs8193036	D	H				
	rs4711998	R	H			H	
		D				H	
	rs2275913	R					H
	rs3819024	D					L
		R					L
IL-23R	rs6687620	R			H	H	
		D				H	
	rs11209026	R			H		
TGFβ	rs1800469	D				H	
INF-γ	rs2069705	R			H		

All variables are statistically significant ($P < .05$).
BM, bone marrow; D, donor; H, higher risk; L, lower risk; MAC, myeloablative conditioning; OR, odds ratio; PB, peripheral blood; R, recipient; RIC, reduced-intensity conditioning.
*OR = 1 (neutral); OR < 1 (L); OR > 1 (H).

Limitations

Extensive cGVHD is considered in the present study, which includes patients from 1997, although it is no longer used as an end point in clinical practice.

Statistical analysis

The descriptive analysis of the SNPs was performed using the SNPassoc R package (version 1.5-8). Univariate regression analysis was performed using logistic regression with the SNPassoc R package for SNPs and with IBM SPSS Statistics for Windows (version 21.0; IBM Corp, Armonk, NY). $P < .05$ was considered significant.

Multivariate regression analysis was performed using the LASSO procedure, which is being increasingly applied to overcome the challenges posed by high-dimensional data.

LASSO is an innovative estimation method for linear regression models developed in 1996 by Tibshirani,[27] which is able to select a set of optimal predictors from a large set of potential predictor variables. This method constrains the sum of absolute values of the regression coefficients by means of a smoothing parameter (λ), shrinking the estimated coefficients toward 0. Because of this, it is considered a powerful method for variable selection, providing more interpretable models. The idea of LASSO is quite general and can be applied in other statistical models, such as the generalized linear models.

In this study, the response variable Y is a binary variable that denotes whether the patient is affected by GVHD/NRM or not (Y = 1 and Y = 0, respectively). In that sense, LASSO was considered a variable selection method under the estimation of a Logit regression model (which is a particular type of generalized linear model).

Table 3. LASSO multivariate analysis of the association between clinical variables and GVHD/NRM

	OR*				
Clinical variable	**aGVHD 2-4**	**aGVHD 3-4**	**cGVHD**	**Extensive cGVHD**	**NRM**
Recipient sex, female vs male	0.872			0.684	0.805
Recipient age > median 45 y			1.095	1.043	2.122
Female donor/male recipient					1.178
Diagnosis					
AML		0.876			0.856
ALL					
MDS			1.204		2.244
Lymphoma	1.181				1.232
Myelofibrosis	0.490				
MM	0.657				
Conditioning, RIC vs MAC	0.957	0.760			
Conditioning regimen, no TBI vs TBI		0.838			1.311
Stem-cell source, PB vs BM	1.151		1.470	2.360	
aGVHD 2-4			1.284	1.253	

*OR values for the variables selected by the LASSO procedure (described in "Methods"). OR values were obtained with the exponential function of β coefficient, which compares the strength of the effect of each individual independent variable with the dependent variable. The higher the absolute value of β, the stronger the effect. $\beta = 0$, OR = 1 (neutral); $\beta < 0$, OR < 1 (L); $\beta > 0$, OR > 1 (H).

In this model, the strength of the penalty term is controlled by a smoothing parameter (λ), so that the larger λ is, the more parsimonious the model is (if $\lambda = 0$, then all the predictors are considered in the final model). Because of this, it is important to find the optimal parameter λ, which provides the best predictive model to anticipate GVHD and NRM. To this end, λ was chosen by adhering to the principle of parsimony and maximizing the area under the receiver operating characteristic curve (AUC) and the correct classification rates (CCRs; global CCR); for patients who do not develop GVHD, CCR0; for patients who develop GVHD, CCR1) associated with the fitted models for a grid of 100 values of λ (supplemental Figure 1). Theoretically, the AUC takes values between 0 and 1, although the practical lower bound is 0.5. A perfect classifier has an AUC of 1. All clinical and genetic variables were included in the LASSO multivariate analysis independently of the *P* value.

The statistical model was fitted (goodness-of-fit assessment) by randomly selecting 85% of the data (training set: 85% of cases and 85% of controls, because sets were representative of our initial sample), and the predictive ability was computed with the remaining 15% (test set). To evaluate the performance and the predictive ability of each model, training and testing samples were randomly selected 100 times. The distribution of the CCR and the AUC over the 100 iterations was shown by means of box plots and a statistical summary of the results. LASSO multivariate regression analysis was shown as odds ratio, which is the exponential function of the β coefficient of LASSO (odds ratio = exp[β coefficient]).

Finally, for prediction purposes, the cutoff point between low and high risk was based on the proportion of patients who developed (Y = 1) and did not develop (Y = 0) GVHD or NRM.[30] These proportions of Y = 1 were 0.28 for grades 2 to 4 aGVHD, 0.11 for grade 3 to 4 aGVHD, 0.53 for cGVHD, 0.30 for extensive cGVHD, and 0.24 for NRM.

Predictive models

On the basis of LASSO multivariate analyses, a risk score was calculated for grade 2 to 4 and grade 3 to 4 aGVHD, for cGVHD and extensive cGVHD, and for NRM. To build the predictive model, GVHD and NRM risk scores were weighted by the size of the effect on the β coefficient of each variable and a constant obtained by the LASSO procedure, within a risk score equation.[31] Such risk scores were used to calculate the risk for each patient who was classified as low risk (when the risk score fell below the cutoff point) and high risk (risk score above the cutoff point).

Results

Descriptive analysis

Results of the descriptive analysis including genotype frequencies, Hardy-Weinberg equilibrium, and minor allele frequency of the 25 SNPs in donors and recipients are summarized in supplemental Table 3. Genotype frequencies were similar to those of the 1000 Genomes Project for the Spanish population and were in accordance with Hardy-Weinberg equilibrium, except for the frequencies of the IL-10 SNPs (rs1800871, rs1800872, and rs1800896), which were in linkage disequilibrium.

Univariate and multivariate analyses

The association between clinical and genetic variables in donors and in recipients with the development of aGVHD, cGVHD, and NRM was investigated using univariate analysis (summarized in Table 2; detailed in supplemental Table 4) and LASSO multivariate analysis (Tables 3-6).

Univariate analysis allowed us to identify statistically significant ($P < .05$) clinical variables and cytokine gene polymorphisms associated with the development of aGVHD, cGVHD, and NRM. Clinical variables seemed to have a reduced influence on the

Table 4. LASSO multivariate analysis of the association between genetic variables and GVHD/NRM

Genetic variable	Polymorphism	Genotype	D/R	Model of transmission	OR*				
					aGVHD 2-4	aGVHD 3-4	cGVHD	Extensive cGVHD	NRM
IL-1A	rs1800587	CT	D	Cod			0.750		1.383
IL-1B	rs1143627	TT	R	Dom	0.809				
		CT	R	Cod		1.037			
		TT	D	Dom			1.472	1.135	
		TT	D	Cod			1.113		
		CC<CT<TT	R	Add					0.95
		CC<CT<TT	D	Add			1.202	1.060	
	rs16944	GG	R	Dom	0.893				
		CC	D	Dom				1.020	
		AA<AG<GG	D	Add		1.128			
	rs1143634	CC	R	Dom		0.831	0.847		
		CC	R	Cod		0.952	0.987		
		TT<CT<CC	R	Add			0.993		
		TT	D	Rec		2.763	1.071	1.604	0.681
		TT<CT<CC	D	Add			0.747		
IL-2	rs2069762	GG	R	Rec		0.278	0.329	0.795	0.904
		GT	R	Cod		1.153	1.014		
		GG<GT<TT	D	Add			0.736		1.032
IL-6	rs1800795	GG	R	Dom		1.471			
		GG	R	Cod		0.587			0.758
		CC	R	Rec				0.768	
		CG	R	Cod		1.025			
		GG	D	Dom		0.877			
		CC	D	Rec			0.922		
IL-7R	rs1494555	CT	R	Cod			1.316		
		CC	R	Rec		1.703			
		CT	D	Cod		0.813		1.596	0.843
		TT	D	Dom			1.071		
		CC<CT<TT	D	Add			1.166		
IL-10	rs1800871	TT	R	Rec				0.702	
		CT	R	Cod		3.351	1.374		
		CT	D	Cod	1.124				
		TT	D	Rec		1.019			
		TT<CT<CC	D	Add				1.104	
	rs1800872	AC	R	Cod			1.004		
		AA	D	Rec		1.005			
	rs1800896	AG	R	Cod	1.246				
		AG	D	Cod	1.139				
		GG<AG<AA	D	Add			0.794	0.953	
		GG	D	Rec			1.419		
IL-17A	rs8193036	CC	R	Rec		1.397	2.395	1.329	
		CT	R	Cod			0.919	0.903	
		CC<CT<TT	R	Add		0.833			
		CC	D	Rec	.,419	2.677			

Add, additive; Cod, codominant; Dom, dominant; Rec, recessive.

Values (OR) for the variables were selected by the LASSO procedure (described in "Methods").

*OR values for the variables selected by the LASSO procedure (described in "Methods"). OR values were obtained with the exponential function of β coefficient, which compares the strength of the effect of each individual independent variable with the dependent variable. The higher the absolute value of β, the stronger the effect. $\beta = 0$, OR = 1 (neutral); $\beta < 0$, OR < 1 (L); $\beta > 0$, OR > 1 (H).

Table 4. (continued)

Genetic variable	Polymorphism	Genotype	D/R	Model of transmission	OR*				
					aGVHD 2-4	aGVHD 3-4	cGVHD	Extensive cGVHD	NRM
		CC<CT<TT	D	Add				0.858	
		CT	D	Cod		0.705			
	rs3819024	AA	R	Dom			1.149	1.254	
		GG	R	Cod			1.002		
		AG	R	Cod			0.747		
		GG	R	Rec		0.869			
		GG	D	Rec	1.126	1.386			
	rs4711998	AG	R	Cod			1.092	1.071	
		GG	R	Cod				0.990	
		AA	R	Rec			0.746		
		GG	R	Dom				0.712	
		AA<AG<GG	D	Add		0.837			
		GG	D	Dom			0.894	0.674	
		AG	D	Cod	1.377		1.013		
		GG	D	Cod			0.918		
	rs2275913	AG	R	Cod			1.340		
		AA<AG<GG	R	Add					0.835
		AA	R	Rec		0.077			1.617
		AA	D	Rec		1.424			1.446
IL-17F	rs763780	TT	R	Dom			0.723		
IL-23R	rs6687620	TT	R	Rec		0.539		8.736	
		CC	R	Dom			0.564		
		TT<CT<CC	R	Add					1.620
		TT	D	Rec		0.691	3.402	1.013	1.661
		TT<CT<CC	D	Add		1.358			0.639
		CT	D	Cod				0.514	
	rs11209026	GG	R	Dom			0.495	0.994	1.740
		AA<AG<GG	R	Add			0.776	0.943	
		AA	D	Rec			0.181		1.740
		AG	D	Cod			1.110		
		AA<AG<GG	D	Add		1.030			
INF-γ	rs2430561	AT	R	Cod			1.537	0.806	
		AA<AT<TT	R	Add		0.958			
		AA	R	Rec		1.778			
		AT	D	Cod		0.771			
	rs2069705	CT	R	Cod			0.641	0.755	0.966
		TT	R	Rec					1.114
		CT	D	Dom	0.880				
		TT<CT<CC	D	Cod		1.296			
		CT	D	Add				0.810	
TGFβ	rs2241716	AA<AG<GG	R	Add		0.990			
		GG	R	Dom		0.979			
		GG	D	Dom	0.359	0.154			
		AA<AG<GG	D	Add		0.667			

Add, additive; Cod, codominant; Dom, dominant; Rec, recessive.

Values (OR) for the variables were selected by the LASSO procedure (described in "Methods").

*OR values for the variables selected by the LASSO procedure (described in "Methods"). OR values were obtained with the exponential function of β coefficient, which compares the strength of the effect of each individual independent variable with the dependent variable. The higher the absolute value of β, the stronger the effect. $\beta = 0$, OR = 1 (neutral); $\beta < 0$, OR < 1 (L); $\beta > 0$, OR > 1 (H).

Table 4. (continued)

Genetic variable	Polymorphism	Genotype	D/R	Model of transmission	OR*				
					aGVHD 2-4	aGVHD 3-4	cGVHD	Extensive cGVHD	NRM
	rs1800469	TT	R	Rec		1.566	3.320		
		CC	R	Dom	0.875				
		TT<CT<CC	R	Add	0.994				
		CT	D	Cod	1.008	0.956			
		TT	D	Rec			0.280	0.722	
		TT<CT<CC	D	Add				1.101	
TNFα	rs1799964	CC	R	Rec	0.968				
		TT	R	Dom			1.237		
		CC	D	Rec		1.556			
		CC<CT<TT	D	Add		0.939			
	rs1800629	AG	R	Cod		2.820			
		AA	D	Rec			0.938		
	rs1800610	AA	R	Rec			0.807	0.632	
		AA<AG<GG	R	Add			1.128		
		TT	R	Rec			8.854		
		AA<AG<GG	D	Add					1.139
		TT	D	Rec		0.751			

Add, additive; Cod, codominant; Dom, dominant; Rec, recessive.

Values (OR) for the variables were selected by the LASSO procedure (described in "Methods").

*OR values for the variables selected by the LASSO procedure (described in "Methods"). OR values were obtained with the exponential function of β coefficient, which compares the strength of the effect of each individual independent variable with the dependent variable. The higher the absolute value of β, the stronger the effect. $\beta = 0$, OR = 1 (neutral); $\beta < 0$, OR < 1 (L); $\beta > 0$, OR > 1 (H).

development of aGVHD. In contrast, IL-1B and IL-17A were the most important cytokines for the development of grade 2 to 4 aGVHD, as IL-6 was for grade 3 to 4 aGVHD. Clinical variables (age, conditioning regimen, stem-cell source, and previous development of aGVHD) seemed to have a greater influence on the development of cGVHD. Likewise, the most important cytokines were IL-1A, IL-1B, IL-23R, and INF-γ for cGVHD and IL-2, IL-17A, IL-23R, and TGFβ for extensive cGVHD. Only sex mismatch and IL-17A were associated with the occurrence of NRM (Table 2; supplemental Table 4).

The LASSO approach (supplemental Figure 1) allowed us to obtain the best models for predicting GVHD (aGVHD and cGVHD) and NRM. For anticipating grade 2 to 4 aGVHD, cGVHD, and NRM, none of the models was good enough for stratifying patients (Figure 1). For grade 3 to 4 aGVHD, the best C-M included 3 variables: conditioning, TBI, and disease (Table 3; Figure 1), which rendered a CCR1 of 50% and AUC of 0.6. The negative predictive value (NPV) of this model was 91.8%.

The best G-M included 10 cytokines (IL-1B, IL-2, IL-6, IL-7R, IL-10, IL-17A, IL-23R, INF-γ, TGFβ, and TNFα). This model obtained a CCR1 of 88%, with an AUC of 0.8 and NPV of 96% (Table 4; Figure 1). In contrast, the model that included both clinical and genetic variables retained the same clinical variables from C-M and added 11 cytokines (IL-1A, IL-1B, IL-2, IL-6, IL-7R, IL-10, IL-17A, IL-23R, INF-γ, TGFβ, and TNFα). This model obtained a CCR1 of 100%, AUC of 0.9, and NPV of 98.6% (Tables 5 and 6; Figure 1). Interestingly, 9 SNPs were selected by LASSO in both models (grades 2 to 4 and 3 to 4), in genes that could be relevant for aGVHD pathophysiology. However, 13 SNPs were selected only in the grade 3 to 4 model in genes that may be related to the severity of the complication.

The best clinical model for predicting extensive cGVHD included age, sex, stem-cell source, and previous aGVHD (Table 3; Figure 1), with a CCR1 of 66.7%, AUC of 0.7, and NPV of 82.9%. The best genetic model included 10 cytokines (IL-1B, IL-2, IL-6, IL-7R, IL-10, IL-17A, IL-23R, INF-γ, TGFβ, and TNFα). This model obtained a CCR1 of 80%, AUC of 0.8, and NPV of 81% (Table 4; Figure 1).

When both genetic and clinical variables were included, the same clinical variables persisted, and 8 cytokines were added (IL-1B, IL-2, IL-7R, IL-10, IL-17A, IL-23R, INF-γ, and TGFβ), improving the results of C-M, with a CCR1 of 80%, AUC of 0.8, and NPV of 85.1% (Tables 5 and 6; Figure 1).

A detailed explanation of the SNPs selected in CG-M in light of previously reported results is included in Table 6 and in supplemental material.

On the basis of the β results from LASSO, risk scores were calculated for aGVHD and cGVHD as well as for NRM. Patients were categorized into 2 groups: low risk (below the cutoff value) and high risk (above the cutoff). Final risk scores with C-M, G-M, and CG-M are summarized in supplemental Tables 5-7, respectively.

Overall, prediction of grade 3 to 4 aGVHD was significantly better using CG-M ($P < .001$) than using C-M or G-M (Figures 1 and 2). However, similar results were obtained when predicting extensive cGVHD with both CG-M and G-M, and both performed better than

Table 5. Multivariate analysis by LASSO of clinical and genetic variables and GVHD/NRM

Variable	Polymorphism	Genotype	D/R	Model of transmission	OR*				
					aGVHD 2-4	aGVHD 3-4	cGVHD	Extensive cGVHD	NRM
Clinical									
Recipient age >45 y							1.080		1.570
Recipient sex, male vs female								0.693	0.974
Female donor/male recipient									1.318
AML						0.858			
ALL						1.614			
MDS							1.260		2.374
Lymphoma						1.330			
Myelofibrosis						0.640			
MM						1.827			
Conditioning, RIC vs MAC						0.679			
Conditioning regimen, TBI vs no TBI						0.575			
Stem-cell source, PB vs BM							1.557	2.281	
aGVHD, grade 2-4							1.343	1.231	
Genetic									
IL-1A	rs1800587	TC	D	Add					1.508
		TT	R	Rec		1.058			
IL-1B	rs1143627	TT	R	Dom	0.809				
		CC<CT<TT	D	Add			1.079	1.009	
	rs16944	GG	R	Dom	0.893				
		AA<AG<GG	D	Add		1.095			
	rs1143634	TT<CT<CC	R	Add		0.927			
		TT	D	Rec		2.968		1.016	0.664
IL-2	rs2069762	GT	D	Cod					1.162
		GG	R	Rec		0.248	0.848	1.107	0.800
		GT	R	Cod		1.241	1.092		
IL-6	rs1800795	GG	R	Cod		1.065			
		GG	R	Dom		1.664			
		CG	R	Cod		1.557			
		CC<CG<GG	R	Add		0.671			
		GG	D	Cod		0.996			
		GG	D	Dom		0.886			
IL-7R	rs1494555	CC	R	Rec		2.007			
		CT	D	Cod		0.792		1.161	0.845
IL-10	rs1800871	CT	R	Cod		3.046			
		TT	D	Rec		1.213			
		CT	D	Cod	1.124				
		TT<CT<CC	D	Add				1.002	
	rs1800872	AC	R	Cod		1.088			
	rs1800896	AG	R	Cod	1.246				
		AG	D	Cod	1.139				
IL-17A	rs8193036	CC	R	Rec		1.062	1.019		
		CC<CT<TT	R	Add		0.642			
		CC	D	Rec	1.419	2.274			
		CT	D	Cod		0.611			

*OR values for the variables selected by the LASSO procedure (described in "Methods"). OR values were obtained with the exponential function of β coefficient, which compares the strength of the effect of each individual independent variable with the dependent variable. The higher the absolute value of β, the stronger the effect. $\beta = 0$, OR = 1 (neutral); $\beta < 0$, OR < 1 (L); $\beta > 0$, OR > 1 (H).

Table 5. (continued)

Variable	Polymorphism	Genotype	D/R	Model of transmission	OR*				
					aGVHD 2-4	aGVHD 3-4	cGVHD	Extensive cGVHD	NRM
	rs3819024	GG	R	Rec		0.797			
		GG	D	Rec	1.126	1.222			
		GG<AG<AA	D	Add		0.917			0.906
	rs4711998	AG	R	Cod	1.377				
		GG	R	Dom				1.346	
		AA<AG<GG	D	Add		0.894			
	rs2275913	AA	R	Rec		0.085			1.267
		AA<AG<GG	R	Add					0.710
		AA	D	Rec		1.220			
		AG	D	Cod		0.963			
IL-23R	rs6687620	TT<CT<CC	R	Add			0.851		
		CT	R	Cod					1.671
		TT	R	Rec		0.702		3.827	
		CC	R	Dom			0.928		
		TT	D	Rec		0.396			
		TT<CT<CC	D	Add		1.267			
		CT	D	Cod				0.752	0.647
	rs11209026	GG	R	Dom			0.983		1.896
		AA<AG<GG	R	Add			0.967		
		AA	D	Rec					2.712
IFN-γ	rs2430561	AT	R	Cod				0.975	
		AA	R	Rec		1.745			
		AT	D	Cod		0.751			
	rs2069705	TT	R	Rec					0.804
		CT	R	Cod			0.997		
		TT<CT<CC	R	Add		0.929			
		CC	D	Dom	0.880				
		CT	D	Cod		1.304			1.162
		TT	D	Rec				1.017	
TGFβ	rs2241716	GG	D	Dom	0.359	0.140			
		GG	D	Cod		0.839			
		AA<AG<GG	D	Add		0.581			
	rs1800469	TT	R	Rec		1.201			
		CC	R	Dom	0.875				
		TT<CT<CC	R	Add	0.994				
		TT	D	Rec			0.865	0.908	
		CT	D	Cod	1.008				
TNFα	rs1799964	CC	R	Rec	0.968				
		CC	D	Rec		1.613			
		CC<CT<TT	D	Add		0.930			
	rs1800629	AG	R	Cod		3.422			
	rs361525	AG	R	Cod		1.140			
		GG	D	Dom		0.948			

*OR values for the variables selected by the LASSO procedure (described in "Methods"). OR values were obtained with the exponential function of β coefficient, which compares the strength of the effect of each individual independent variable with the dependent variable. The higher the absolute value of β, the stronger the effect. $\beta = 0$, OR = 1 (neutral); $\beta < 0$, OR < 1 (L); $\beta > 0$, OR > 1 (H).

Table 6. Summary of genetic variables selected by LASSO multivariate analysis and their influence on GVHD/NRM development compared with previous publications

Genetic variable	Polymorphism	Genotype	D/R	OR*					Inflammatory	Physiology	Reference
				aGVHD 2-4	aGVHD 3-4	cGVHD	Extensive cGVHD	NRM			
IL-1A	rs1800587	TT	R		H			H	Pro	T/T genotype: increased promoter activity and IL-1A production, higher incidence of cGVHD	51,52
IL-1B	rs1143627	TT	R	L					Pro	C allele: high levels of IL-1	53
		CC<CT<TT	D			H	H			TT: higher risk of aGVHD	54
	rs16944	CC	R	L						T allele: high levels of IL-1	53
		TT<TC<CC	D		H						
	rs1143634	TT<CT<CC	R		L		H			T allele: high IL-1 production and more aGVHD; TT genotype: correlated with protector for infections	55-57
		TT	D		H†		H	L			
IL-2	rs2069762	GT	D					H	Pro	GG: lower levels of IL-2 and lower GVHD	58,59
		GG	R		L†	L	H	L			
		GT	R		H	H					
IL-6	rs1800795	GG/GC	R		H				Pro	Allele G: higher levels of IL-6, higher risk of aGVHD	60,61
		GG	D		L						
IL-7R	rs1494555	GG	R		H†				Pro	GG in donor: higher severe GVHD and NRM	62
		GA	D		L		H	L			
IL-10	rs1800871	TC	R		H†				Anti	Allele C: high IL-10 production, Lower GVHD	63,64
		TT/TC	D	H							
	rs1800872	AC	R		H					Allele C: high IL-10 production, lower GVHD	
	rs1800896	AG	R	H						Allele G: high IL-10 production, lower GVHD	
		AG	D	H							
IL-17A	rs8193036	CC	R		H	H			Pro	Genotype CC: higher GVHD	38
		CC	D	H	H†						
		CT	D		L						
	rs2275913	AA	R		L†			H		Allele A: correlated with higher levels of IL-17A and higher GVHD in donor	65,66
		AA	D		H						
		AG	D		L						

*OR values for the variables selected by the LASSO procedure (described in "Methods"). OR values were obtained with the exponential function of β coefficient, which compares the strength of the effect of each individual independent variable with the dependent variable. The higher the absolute value of β, the stronger the effect. $\beta = 0$, OR = 1 (neutral); $\beta < 0$, OR < 1 (L); $\beta > 0$, OR > 1 (H).

†H, OR > 2; L, OR < 0.5.

Table 6. (continued)

Genetic variable	Polymorphism	Genotype	D/R	OR* aGVHD 2-4	OR* aGVHD 3-4	OR* cGVHD	OR* Extensive cGVHD	OR* NRM	Inflammatory	Physiology	Reference
	rs3819024	GG	R		L					Allele A: correlated with lower levels of IL-17A and lower GVHD	
		GG	D	H	H						
	rs4711998	GG/AG	R				H				
		AA<AG<GG	D		L						
IL-23R	rs6687620	TT<CT<CC	R			L			Pro		
		CT	R					H			
		TT	R		L		H†				
		TT/CT	D		L†		L				
	rs11209026	GG	R			L				No correlation was obtained	67
IFN-γ	rs2430561	AT	R				L		Pro/anti	Genotype TT: correlated with higher levels; T allele: correlated with moderate-severe GVHD	68,69
		TT	R					L			
		AA	R		H						
		AT	D		L						
	rs2069705	TT<CT<CC	R		L					Genotype TT: correlated with higher levels; CC genotype: correlated with low incidence of cGVHD	70,71
		CC	D	L							
		CT/TT	D		H		H	H			
TGFβ	rs2241716	GG	D	L†	L†				Anti	Genotype AA: high incidence of extensive cGVHD	71
	rs1800469	CC	R	L						Allele C: correlated with higher probability to develop aGVHD and lower levels	71,72
		CT	D	H							
		TT	D			L	L				
TNFα	rs1799964	CC	R	L					Pro	TT carriers: lower levels than CC genotype	73,74
		CC	D		H						
	rs1800629	AG	R		H†					AA genotype: correlated with higher levels	75,76
	rs361525	AG	R		H					GG genotype: low production of TNFα; GA genotype: higher rates of cGVHD	77,78
		GG	D			L					

*OR values for the variables selected by the LASSO procedure (described in "Methods"). OR values were obtained with the exponential function of β coefficient, which compares the strength of the effect of each individual independent variable with the dependent variable. The higher the absolute value of β, the stronger the effect. $\beta = 0$, OR = 1 (neutral); $\beta < 0$, OR < 1 (L); $\beta > 0$, OR > 1 (H).

†H, OR > 2; L, OR < 0.5.

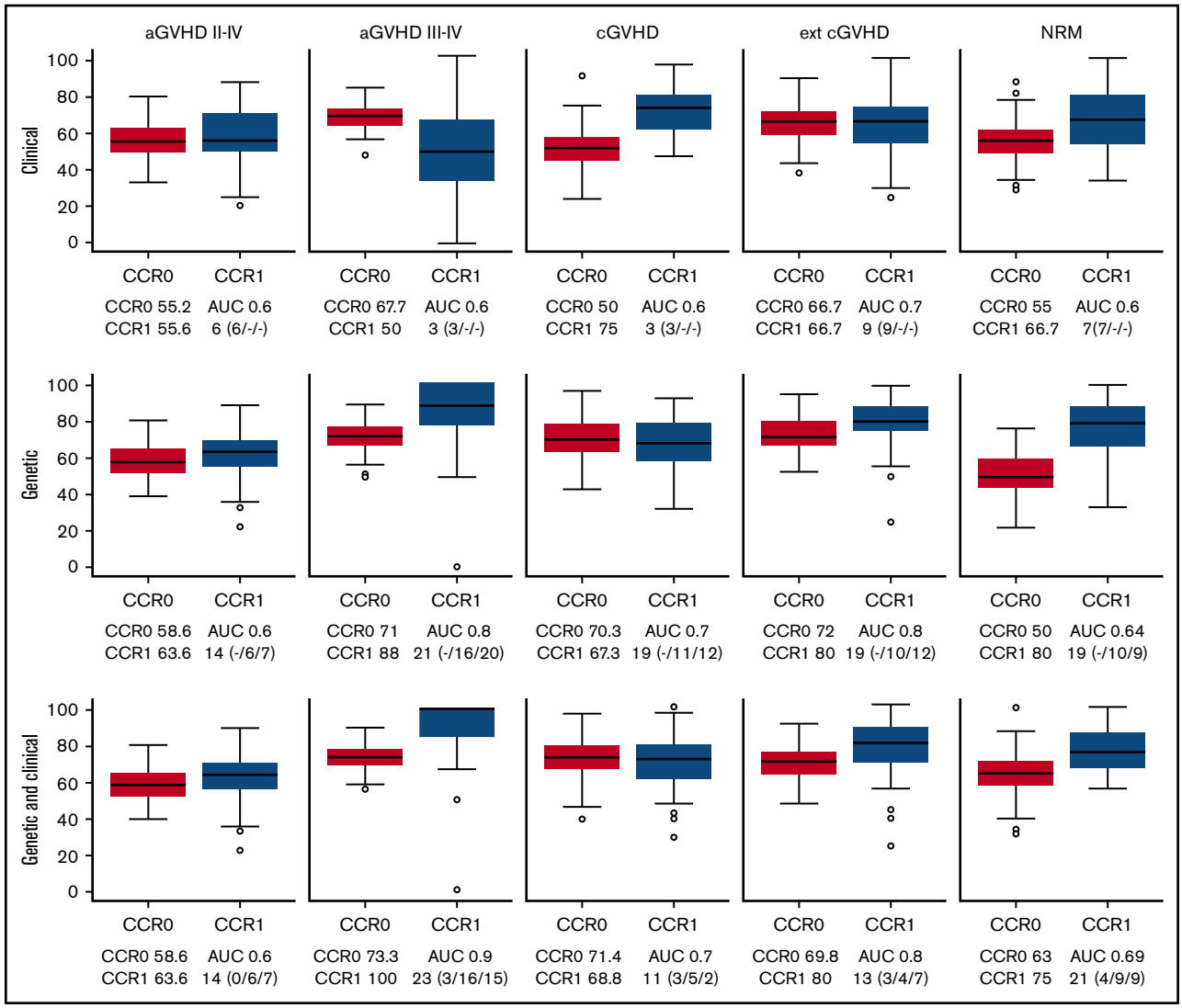

Figure 1. Box plots of CCRs for patients who develop (CCR1) and do not develop (CCR0) GVHD/NRM as predicted with the different models. The AUC and the number of variables used are shown in each case. The predictive ability of each model, built using 85% of the samples (training set), was computed with the remaining 15% of the samples (test set). To evaluate the performance and predictive ability of each model, training and testing samples were randomly selected and the procedure repeated 100 times. The distribution of the CCR and AUC over the 100 iterations is shown by means of box plots. CCRs for the development of aGVHD, cGVHD, and NRM obtained using the predictive model including only clinical variables (upper panels), the model including only genetic variables (middle panels), and the model including both clinical and genetic variables (lower panels) are shown. Number of clinical/recipient SNPs/donor SNPs variables is indicated in parenthesis.

C-M (Figures 1 and 3). When NRM was considered, both C-M and CG-M performed better than G-M (supplemental Figure 5).

Also, we calculated GVHD risk scores for the patients most recently undergoing transplantation (2005-2012) to test the usefulness of the models in a subset of patients treated following current practices. Interestingly, the results obtained in this subgroup of patients were similar to those reported for the whole cohort (supplemental Figures 3 and 4).

Finally, we calculated the incidence of aGVHD also including censored patients (total, n = 359), which interestingly rendered similar results compared with those of the previous analysis (supplemental Figure 4). Incidence of cGVHD could not be calculated because of a lack of cGVHD onset time point information.

Discussion

Despite advances in the knowledge of the pathophysiology of GVHD during the last 2 decades, 30% to 50% of patients undergoing allo-SCT develop this complication,[32] which leads to high morbidity, reduces quality of life, and is associated with a significantly higher risk of treatment-related mortality and poorer overall survival.[3] Therefore, it is essential to identify biomarkers that can help to estimate the risk of GVHD. Biomarkers may also identify patients who will not respond to traditional treatments,[33] making it possible to implement more stringent monitoring and specific preventive care or modify treatment. Moreover, the ability to anticipate the risk of subsequent morbidity and mortality could facilitate personalized

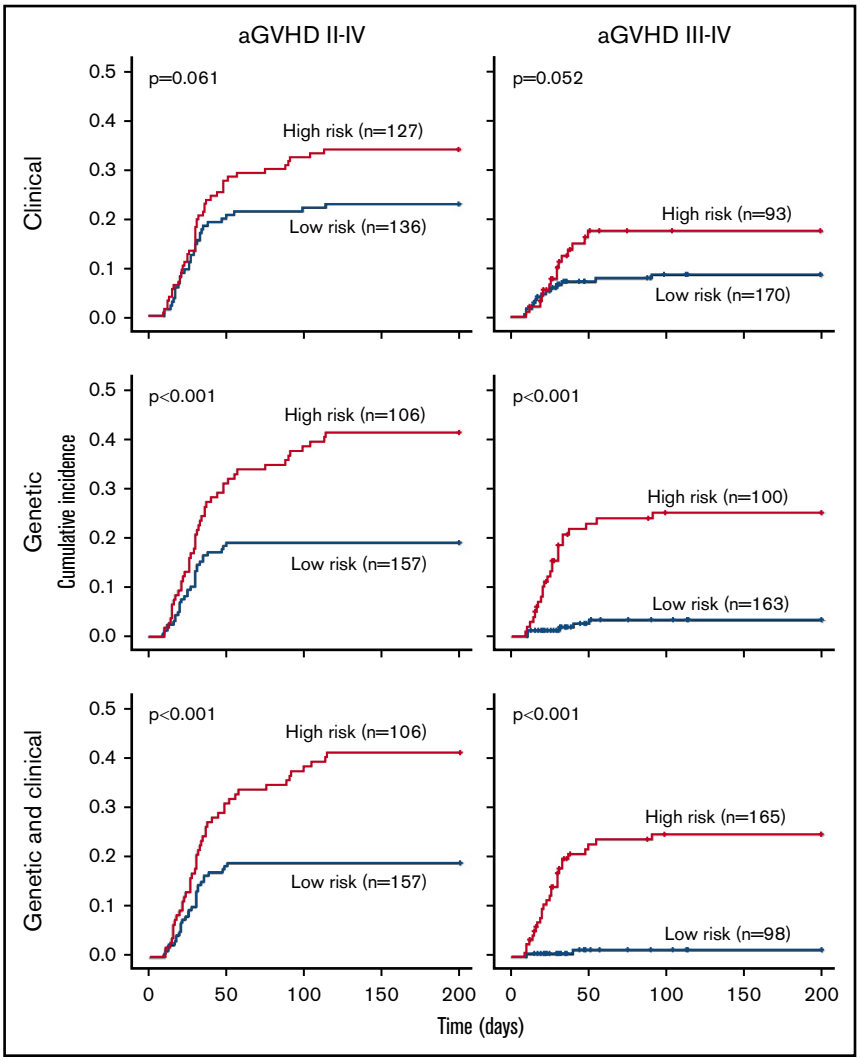

Figure 2. Stratification of the whole cohort of patients (n = 263) according to the risk of developing acute GVHD. Risk was calculated using the proposed predictive model including clinical variables (upper panels), genetic variables (middle panels), or both clinical and genetic variables (lower panels). Cumulative incidence curves are shown for the development of grade 2 to 4 aGVHD (left panels) and grade 3 to 4 aGVHD (right panels). The cutoff used was 0.28 for grade 2 to 4 aGVHD and 0.11 for grade 3 to 4 aGVHD.

treatment plans, including additional immunosuppressive therapies introduced early for high-risk patients or reduced-intensity approaches in low-risk patients.

Classically, GVHD has been estimated based almost entirely on the presence of clinical symptoms; indeed, over the last 15 years, several groups,[2,11,13,15,16] including ours,[12,14,34] have demonstrated that non-HLA SNPs can be used as biomarkers to anticipate GVHD. Although some of these reports identified individual SNPs, in most cases, no single SNP is sufficient for prognosis. Thus, the simultaneous use of several SNPs may increase specificity and predictability. In any case, there are currently no validated laboratory tests to predict the risk of GVHD or patient survival.

Given that this study was performed in a large and homogeneous cohort, our results suggest that SNPs in cytokine genes, in combination with clinical factors, could predict severe GVHD (grade 3-4 aGVHD and extensive cGVHD). One of the limitations of our retrospective study is that the end point of extensive cGVHD, which is no longer used in clinical practice, was considered because patients from 1997 were included. Therefore, current clinical applicability of such results is limited, and results must be validated in an independent study considering in-use end points of mild, moderate, and severe cGVHD.

As described before, the LASSO procedure autonomously selected clinical variables that were previously known to influence GVHD and NRM development, such as older patient age, peripheral blood as stem-cell source, female donor/male recipient, and previous aGVHD,[8-10] confirming the robustness of the approach.

Other characteristics (reduced-intensity conditioning and not having received TBI), for which reported data were more controversial, were also selected by the LASSO procedure as associated with GVHD,[10,32] probably because less intense regimens tend to be offered to older and more heavily treated patients. Gene variants have been shown to alter the expression or function of the proteins responsible for immune response,[4] and there is growing evidence to support the importance of genetic variability (gene polymorphisms) for predicting the risk of GVHD in individual recipients. In the present study, the LASSO procedure selected polymorphisms in known cytokine genes such as *IL1B*,[35,36] *IL6*,[26] *IL10*,[37] *IL17A*,[38-40] *IL23R*,[41] *INFγ* ,[42] *TGFβ* ,[43] and *TNFα*,[44,45] which were confirmed to correlate with the risk of severe aGVHD, and polymorphisms in *IL1B*,[36] *IL2*, *IL7R*, *IL17A*, *IL23R*, *INFγ* ,[42] and

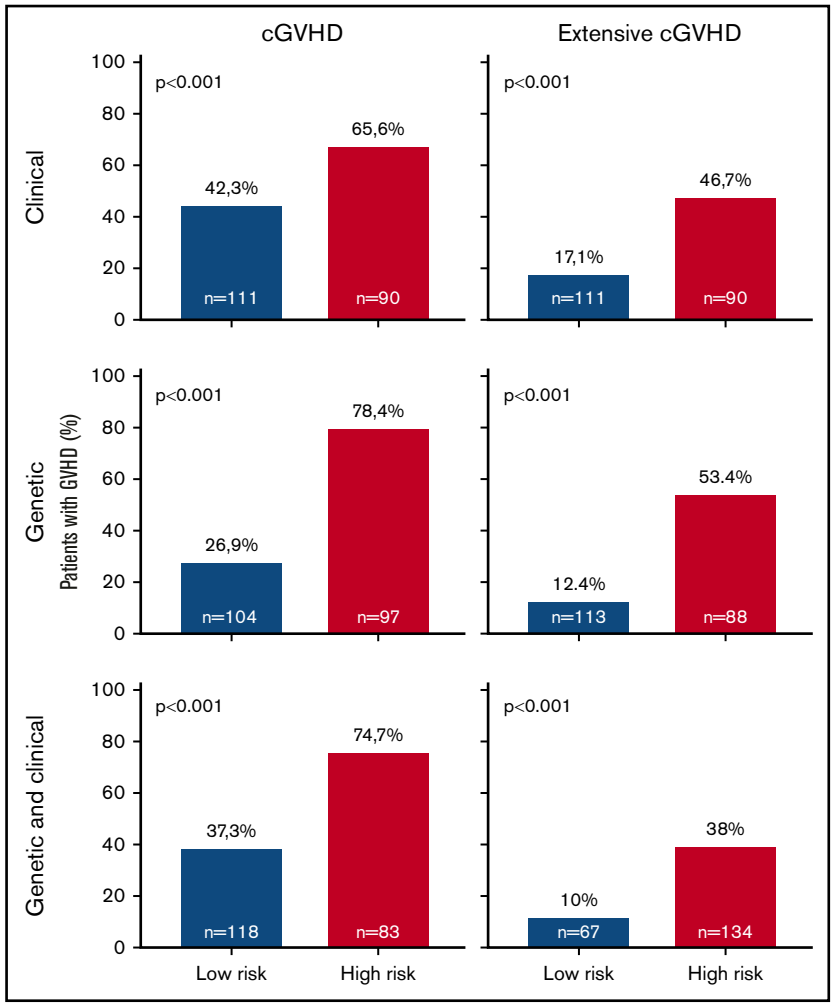

Figure 3. Stratification of the whole cohort of patients (n = 201) according to the risk of developing chronic GVHD. Risk was calculated using the proposed predictive model including clinical variables (upper panels), genetic variables (middle panels), or both clinical and genetic variables (lower panels). Because the time of onset after transplantation was not available to build cumulative incidence curves, bar charts are shown for cGVHD (right panels) and extensive cGVHD (left panels). The cutoff used was 0.53 for cGVHD and 0.3 for extensive cGVHD.

TGFβ were found to play an important role in the risk of extensive cGVHD. Furthermore, LASSO identified an association between various genes (*IL2*[46,47] and *IL7R*[48]) and GVHD, which have remained controversial in the literature. Of note, the approach generates complex models to predict severe GVHD that include a high number of genetic variables, probably derived from the fact that GVHD is a complex entity with different phases and cell types involved.

The main strength of our study is the development of a predictive model that combines clinical and genetic variables in a large cohort of homogeneous patients undergoing the same type of transplantation (HLA-identical sibling donor). These findings may not apply to patients undergoing transplantation with unrelated or non–HLA-identical sibling donors. The inclusion of SNPs as markers, together with clinical variables in the risk model, significantly improves the CCRs of patients with severe aGVHD in comparison with the models based only on clinical or genetic data. In fact, the best model for the anticipation of severe aGVHD was CG-M, with a high CCR1 of 100%; CG-M and G-M performed similarly, with a CCR of 80%. These results demonstrate the clinical usefulness of including genetic variables, in addition to the available clinical variables, in the predictive models. Such models are of clinical utility because they consistently identify patients who will develop GVHD. In any case, it is also important to identify those patients who will not develop severe GVHD. Interestingly, CG-M provided an NPV of 98.6% for severe aGVHD and 85.1% for extensive cGVHD. In contrast, the NPVs obtained for the other models were slightly worse (severe aGVHD: C-M, 91% and G-M, 96%; extensive cGVHD: C-M, 82.6% and G-M, 81%).

Interestingly, the models proposed here can be applied by other centers using the mathematical formulas shown in supplemental Tables 5-7.

In light of these results, it could be argued that patients who are classified as high risk for the development of severe GVHD, mainly aGVHD, would still receive standard-of-care immunosuppression. However, patients classified as low risk according to the model, and who therefore will most probably not develop severe GVHD, could benefit from modification of immunosuppressive therapy, thus preserving the graft-versus-leukemia effect. This would be of special relevance in those patients with persistent minimal residual disease before transplantation.[49] Of course, before making any recommendations, these findings must be validated in prospective studies with large cohorts to demonstrate their clinical utility.

As previously mentioned, it is clear that no single SNP is sufficient for prognosis, but the use of simultaneous SNPs may increase specificity and predictability. Thus, other authors have also developed GVHD predictive models. Kim et al[22] proposed SNP-based risk models, also including clinical and genetic variables, associated with transplantation outcomes, which allowed stratification of patients in terms of overall survival, relapse-free survival, NRM, and aGVHD, but not cGVHD. Hartwell et al[50] recently described an early-biomarker algorithm that predicted lethal GVHD and survival measuring 4 biomarkers (ST2, REG3a, TNFR1, and IL-2Rα) on plasma samples on day +7 after SCT in 1287 patients. This study included transplantations performed with various types of donors (unrelated or related), whereas ours included only HLA-identical transplantations. As in our study, this model was also capable of predicting the risk of severe GVHD after SCT before the onset of GVHD symptoms. Unlike our proposal, this algorithm only included the 4 biomarkers and did not consider clinical data. This approach could be combined in future studies with the 1 proposed here to further improve predictive models. Moreover, including genetic markers that help predict response to drugs could drive therapeutic interventions in the management of GVHD.[33]

Our main goal for the future would be to improve the model to make it useful for all patients. To this end, we are currently using next-generation sequencing to search for new polymorphisms in immune response–related genes, minor histocompatibility antigen genes, drug metabolism genes, and innate immunity genes. In conclusion, although prospective validation studies should be performed to confirm these results, the present study suggests a risk model using donor and recipient SNP markers and clinical variables that improves GVHD risk stratification, allowing optimized clinical management of patients undergoing transplantation.

Acknowledgments

The authors would like to thank the Centro Nacional de Genotipado for help with genotyping. They would also like to acknowledge the patients who participated in this study, as well as the staff of the Hematology Department, Hospital General Universitario Gregorio Marañón (Madrid, Spain), who made the study possible.

This study was partially supported by Ministry of Economy and Competitiveness ISCIII-FIS Grants PI08/1463, PI11/00708, PI14/01731, PI17/01880, and RD12/0036/0061 and cofinanced by the European Regional Development Fund from the European Commission, the "A way of making Europe" initiative, and grants from Fundación LAIR and Asociación Madrileña de Hematología y Hemoterapia.

Authorship

Contribution: C.M.-L., E.B., M.C.A.-M., J.L.D.-M., J.R., and I.B. were responsible for conception and design; N.S., B.M.-A., V.G., J.B.N., M.G., R.d.l.C., S.B., A.J.-V., I.E., C.V., A.S., D.S., M.K., J.G., Á.U.-I., C.S., D.G., and J.L.D.-M. provided patients and samples; C.M.-L., E.B., M.C.A.-M., A.P., R.L., and I.B. collected and assembled data; C.M.-L., E.B., M.C.A.-M., A.P., M.G.-R., R.L., J.M.B., P.B., J.L.D.-M., J.R., and I.B. were responsible for data analysis and interpretation; C.M.-L., E.B., M.C.A.-M., and I.B. wrote the manuscript; and all authors gave final approval of the manuscript and are accountable for all aspects of the work.

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

A complete list of the members of the GVHD/Immunotherapy Committee of the Spanish Group for Hematopoietic Transplantation appears in "Appendix."

Correspondence: Carolina Martínez-Laperche, Laboratorio Genética Hematológica, Edif Oncología PI−1, Servicio de Hematología, Hospital General Universitario Gregorio Marañón, C/Doctor Esquerdo 46, 28007 Madrid, Spain; e-mail: cmlaperchehgugm@gmail.com.

Appendix: study group members

The members of the GVHD/Immunotherapy Committee of the Spanish Group for Hematopoietic Transplantation are: C.M.-L., B.M.-A., J.B.N., M.G., R.d.l.C., S.B., I.E., C.V., A.S., D.S., M.K., J.G., P.B., Á.U.-I., C.S., D.G., J.L.D.-M., and I.B.

References

1. Ferrara JL, Levine JE, Reddy P, Holler E. Graft-versus-host disease. *Lancet*. 2009;373(9674):1550-1561.
2. Ferrara JL, Reddy P. Pathophysiology of graft-versus-host disease. *Semin Hematol*. 2006;43(1):3-10.
3. Socié G, Ritz J. Current issues in chronic graft-versus-host disease. *Blood*. 2014;124(3):374-384.
4. Shlomchik WD. Graft-versus-host disease. *Nat Rev Immunol*. 2007;7(5):340-352.
5. Kawase T, Morishima Y, Matsuo K, et al; Japan Marrow Donor Program. High-risk HLA allele mismatch combinations responsible for severe acute graft-versus-host disease and implication for its molecular mechanism. *Blood*. 2007;110(7):2235-2241.
6. Gallardo D, Brunet S, Torres A, et al. Hla-DPB1 mismatch in HLA-A-B-DRB1 identical sibling donor stem cell transplantation and acute graft-versus-host disease. *Transplantation*. 2004;77(7):1107-1110.
7. Dickinson AM, Norden J. Non-HLA genomics: does it have a role in predicting haematopoietic stem cell transplantation outcome? *Int J Immunogenet*. 2015;42(4):229-238.
8. Flowers MED, Inamoto Y, Carpenter PA, et al. Comparative analysis of risk factors for acute graft-versus-host disease and for chronic graft-versus-host disease according to National Institutes of Health consensus criteria. *Blood*. 2011;117(11):3214-3219.
9. Mohty M, Kuentz M, Michallet M, et al; Société Française de Greffe de Moelle et de Thérapie Cellulaire (SFGM-TC). Chronic graft-versus-host disease after allogeneic blood stem cell transplantation: long-term results of a randomized study. *Blood*. 2002;100(9):3128-3134.
10. Mielcarek M, Martin PJ, Leisenring W, et al. Graft-versus-host disease after nonmyeloablative versus conventional hematopoietic stem cell transplantation. *Blood*. 2003;102(2):756-762.

11. Dzierzak-Mietla M, Markiewicz M, Siekiera U, et al. Occurrence and impact of minor histocompatibility antigens' disparities on outcomes of hematopoietic stem cell transplantation from HLA-matched sibling donors. *Bone Marrow Res*. 2012;2012:257086.
12. Santos N, Rodríguez-Romanos R, Nieto JB, et al; GvHD/Immunotherapy Working Party of the Spanish Group of Hematopoietic Transplant (GETH). UGT2B17 minor histocompatibility mismatch and clinical outcome after HLA-identical sibling donor stem cell transplantation. *Bone Marrow Transplant*. 2016;51(1):79-82.
13. Behar E, Chao NJ, Hiraki DD, et al. Polymorphism of adhesion molecule CD31 and its role in acute graft-versus-host disease. *N Engl J Med*. 1996;334(5): 286-291.
14. Bosch-Vizcaya A, Pérez-García A, Brunet S, et al; GvHD/Immunotherapy committee of the Spanish Group for Hematopoietic Transplant (GETH). Donor CTLA-4 genotype influences clinical outcome after T cell-depleted allogeneic hematopoietic stem cell transplantation from HLA-identical sibling donors. *Biol Blood Marrow Transplant*. 2012;18(1):100-105.
15. Dickinson AM, Holler E. Polymorphisms of cytokine and innate immunity genes and GVHD. *Best Pract Res Clin Haematol*. 2008;21(2):149-164.
16. Dickinson AM, Middleton PG, Rocha V, Gluckman E, Holler E; Eurobank members. Genetic polymorphisms predicting the outcome of bone marrow transplants. *Br J Haematol*. 2004;127(5):479-490.
17. Rachakonda SP, Penack O, Dietrich S, et al. Single-nucleotide polymorphisms within the thrombomodulin gene (THBD) predict mortality in patients with graft-versus-host disease. *J Clin Oncol*. 2014;32(30):3421-3427.
18. Dickinson AM. Non-HLA genetics and predicting outcome in HSCT. *Int J Immunogenet*. 2008;35(4-5):375-380.
19. Kim DH, Jung HD, Lee NY, Sohn SK. Single nucleotide polymorphism of CC chemokine ligand 5 promoter gene in recipients may predict the risk of chronic graft-versus-host disease and its severity after allogeneic transplantation. *Transplantation*. 2007;84(7):917-925.
20. Elmaagacli AH, Koldehoff M, Landt O, Beelen DW. Relation of an interleukin-23 receptor gene polymorphism to graft-versus-host disease after hematopoietic-cell transplantation. *Bone Marrow Transplant*. 2008;41(9):821-826.
21. 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.
22. Kim DDH, Yun J, Won H-H, et al. Multiple single-nucleotide polymorphism-based risk model for clinical outcomes after allogeneic stem-cell transplantation, especially for acute graft-versus-host disease. *Transplantation*. 2012;94(12):1250-1257.
23. Kim D, Won H-H, Su S, et al. Risk stratification of organ-specific GVHD can be improved by single-nucleotide polymorphism-based risk models. *Bone Marrow Transplant*. 2014;49(5):649-656.
24. Paczesny S, Krijanovski OI, Braun TM, et al. A biomarker panel for acute graft-versus-host disease. *Blood*. 2009;113(2):273-278.
25. Yu J, Storer BE, Kushekhar K, et al. Biomarker panel for chronic graft-versus-host disease. *J Clin Oncol*. 2016;34(22):2583-2590
26. Chien JW, Zhang XC, Fan W, et al. Evaluation of published single nucleotide polymorphisms associated with acute GVHD. *Blood*. 2012;119(22): 5311-5319.
27. Tibshirani R. Regression shrinkage and selection via the Lasso. *J R Stat Soc B*. 1996;58(1):267-288.
28. Przepiorka D, Weisdorf D, Martin P, et al. 1994 consensus conference on acute GVHD grading. *Bone Marrow Transplant*. 1995;15(6):825-828.
29. Filipovich AH, Weisdorf D, Pavletic S, et al. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. diagnosis and staging working group report. *Biol Blood Marrow Transplant*. 2005;11(12):945-956.
30. Hosmer DW, Lemeshow S. Applied Logistic Regression. New York, NY: Wiley; 1989.
31. Salazar Couso A. Modelos de respuesta discreta en R y aplicación con datos reales. http://masteres.ugr.es/moea/pages/tfm1011/modelosderespuestadiscretaenryaplicacion/!. Accessed 9 July 2011.
32. Jacobsohn DA, Vogelsang GB. Acute graft versus host disease. *Orphanet J Rare Dis*. 2007;2:35.
33. Vander Lugt MT, Braun TM, Hanash S, et al. ST2 as a marker for risk of therapy-resistant graft-versus-host disease and death. *N Engl J Med*. 2013; 369(6):529-539.
34. Noriega V, Martínez-Laperche C, Buces E, et al; Spanish Hematopoietic Stem Cell Transplantation and Cell Therapy Group (GETH). The genotype of the donor for the (GT)n polymorphism in the promoter/enhancer of FOXP3 is associated with the development of severe acute GVHD but does not affect the GVL effect after myeloablative HLA-identical allogeneic stem cell transplantation. *PLoS One*. 2015;10(10):e0140454.
35. Buces E, Martínez-Laperche C, González-Rivera M, et al. Donor and recipient genotypes for interleukin 1 gene single nucleotide polymorphisms (SNPs) allow anticipation of acute graft versus host disease after HLA-identical allogeneic stem cell transplantation (allo-SCT). *Blood*. 2014;124(21):666.
36. Yin WT, Pan YP, Lin L. Association between IL-1α rs17561 and IL-1β rs1143634 polymorphisms and periodontitis: a meta-analysis. *Genet Mol Res*. 2016;15(1).
37. Chambrone L, Ascarza A, Guerrero M-E, et al. Association of -1082 interleukin-10 gene polymorphism in Peruvian adults with chronic periodontitis. *Med Oral Patol Oral Cir Bucal*. 2014;19(6):e569-e573.
38. Carvalho A, Cunha C, Di Ianni M, et al. Prognostic significance of genetic variants in the IL-23/Th17 pathway for the outcome of T cell-depleted allogeneic stem cell transplantation. *Bone Marrow Transplant*. 2010;45(11):1645-1652.
39. Jin Y, Deng Z, Cao C, Li L. IL-17 polymorphisms and asthma risk: a meta-analysis of 11 single nucleotide polymorphisms. *J Asthma*. 2015;52(10): 981-988.
40. Li N, Zhu Q, Li Z, et al. IL17A gene polymorphisms, serum IL-17A and IgE levels, and hepatocellular carcinoma risk in patients with chronic hepatitis B virus infection. *Mol Carcinog*. 2014;53(6):447-457.
41. Gruhn B, Intek J, Pfaffendorf N, et al. Polymorphism of interleukin-23 receptor gene but not of NOD2/CARD15 is associated with graft-versus-host disease after hematopoietic stem cell transplantation in children. *Biol Blood Marrow Transplant*. 2009;15(12):1571-1577.

42. Leng R-X, Pan H-F, Liu J, et al. Evidence for genetic association of TBX21 and IFNG with systemic lupus erythematosus in a Chinese Han population. *Sci Rep*. 2016;6:22081.

43. Wan PQ, Wu JZ, Huang LY, Wu JL, Wei YH, Ning QY. TGF-β1 polymorphisms and familial aggregation of liver cancer in Guangxi, China. *Genet Mol Res*. 2015;14(3):8147-8160.

44. Goyal RK, Lin Y, Schultz KR, et al. Tumor necrosis factor-alpha gene polymorphisms are associated with severity of acute graft-versus-host disease following matched unrelated donor bone marrow transplantation in children: a Pediatric Blood and Marrow Transplant Consortium study. *Biol Blood Marrow Transplant*. 2010;16(7):927-936.e1.

45. Senhaji N, Serrano A, Badre W, et al. Association of inflammatory cytokine gene polymorphisms with inflammatory bowel disease in a Moroccan cohort. *Genes Immun*. 2016;17(1):60-65.

46. MacMillan ML, Radloff GA, Kiffmeyer WR, DeFor TE, Weisdorf DJ, Davies SM. High-producer interleukin-2 genotype increases risk for acute graft-versus-host disease after unrelated donor bone marrow transplantation. *Transplantation*. 2003;76(12):1758-1762.

47. Christensen U, Haagerup A, Binderup HG, Vestbo J, Kruse TA, Børglum AD. Family based association analysis of the IL2 and IL15 genes in allergic disorders. *Eur J Hum Genet*. 2006;14(2):227-235.

48. Azarpira N, Dehghani M, Aghdaie MH, Darai M. Interleukin-7 receptor-alpha gene polymorphisms in bone marrow transplant recipients. *Mol Biol Rep*. 2010;37(1):27-31.

49. Buckley SA, Wood BL, Othus M, et al. Minimal residual disease prior to allogeneic hematopoietic cell transplantation in acute myeloid leukemia: a meta-analysis. *Haematologica*. 2017;102(5):865-873.

50. Hartwell MJ, Özbek U, Holler E, et al. An early-biomarker algorithm predicts lethal graft-versus-host disease and survival. *JCI Insight*. 2017;2(3):e89798.

51. Mehta PA, Eapen M, Klein JP, et al. Interleukin-1 alpha genotype and outcome of unrelated donor haematopoietic stem cell transplantation for chronic myeloid leukaemia. *Br J Haematol*. 2007;137(2):152-157.

52. Cullup H, Dickinson AM, Cavet J, et al. Polymorphisms of interleukin-1alpha constitute independent risk factors for chronic graft-versus-host disease after allogeneic bone marrow transplantation. *Br J Haematol*. 2003;122(5):778-787.

53. Hall SK, Perregaux DG, Gabel CA, et al. Correlation of polymorphic variation in the promoter region of the interleukin-1 beta gene with secretion of interleukin-1 beta protein. *Arthritis Rheum*. 2004;50(6):1976-1983.

54. Cullup H, Stark G. Interleukin-1 polymorphisms and graft-vs-host disease. *Leuk Lymphoma*. 2005;46(4):517-523.

55. Kornman K, Duff G, Reilly P. Re: A critical assessment of interleukin-1 (IL-1) genotyping when used in a genetic susceptibility test for severe chronic periodontitis. Greenstein G, Hart TC (2002;73:231-247). *J Periodontol*. 2002;73(12):1553-1556.

56. Resende RG, Abreu MHNG, de Souza LN, et al. Association between IL1B (+3954) polymorphisms and IL-1β levels in blood and saliva, together with acute graft-versus-host disease. *J Interferon Cytokine Res Off J Int Soc Interferon Cytokine Res*. 2013;33(7):392-397.

57. Montoya-Ruiz C, Jaimes FA, Rugeles MT, López JÁ, Bedoya G, Velilla PA. Variants in LTA, TNF, IL1B and IL10 genes associated with the clinical course of sepsis. *Immunol Res*. 2016;64(5-6):1168-1178.

58. Jung JW, Han SJ, Song MK, et al. Tear cytokines as biomarkers for chronic graft-versus-host disease. *Biol Blood Marrow Transplant*. 2015;21(12):2079-2085.

59. Matesanz F, Fedetz M, Leyva L, et al. Effects of the multiple sclerosis associated -330 promoter polymorphism in IL2 allelic expression. *J Neuroimmunol*. 2004;148(1-2):212-217.

60. Fishman D, Faulds G, Jeffery R, et al. The effect of novel polymorphisms in the interleukin-6 (IL-6) gene on IL-6 transcription and plasma IL-6 levels, and an association with systemic-onset juvenile chronic arthritis. *J. Clin. Invest*. 1998;102(7):1369-1376.

61. Cavet J, Dickinson AM, Norden J, et al. Interferon-gamma and interleukin-6 gene polymorphisms associate with graft-versus-host disease in HLA-matched sibling bone marrow transplantation. *Blood*. 2001;98(5):1594-1600.

62. Shamim Z, Ryder LP, Christensen IJ, et al. Prognostic significance of interleukin-7 receptor-α gene polymorphisms in allogeneic stem-cell transplantation: a confirmatory study. *Transplantation*. 2011;91(7):731-736.

63. Ouma C, Davenport GC, Were T, et al. Haplotypes of IL-10 promoter variants are associated with susceptibility to severe malarial anemia and functional changes in IL-10 production. *Hum Genet*. 2008;124(5):515-524.

64. Tseng L-H, Storer B, Petersdorf E, et al. IL10 and IL10 receptor gene variation and outcomes after unrelated and related hematopoietic cell transplantation. *Transplantation*. 2009;87(5):704-710.

65. Karimi MH, Salek S, Yaghobi R, et al. Association of IL-17 gene polymorphisms and serum level with graft versus host disease after allogeneic hematopoietic stem cell transplantation. *Cytokine*. 2014;69(1):120-124.

66. Espinoza JL, Takami A, Nakata K, et al. A genetic variant in the IL-17 promoter is functionally associated with acute graft-versus-host disease after unrelated bone marrow transplantation. *PLoS One*. 2011;6(10):e26229.

67. Elmaagacli AH, Steckel N, Ditschkowski M, et al. Toll-like receptor 9, NOD2 and IL23R gene polymorphisms influenced outcome in AML patients transplanted from HLA-identical sibling donors. *Bone Marrow Transplant*. 2011;46(5):702-708.

68. Lio D, Marino V, Serauto A, et al. Genotype frequencies of the +874T→A single nucleotide polymorphism in the first intron of the interferon-gamma gene in a sample of Sicilian patients affected by tuberculosis. *Eur J Immunogenetics Off J Br Soc Histocompat Immunogenetics*. 2002;29(5):371-374.

69. Karimi MH, Daneshmandi S, Pourfathollah AA, et al. The IFN-gamma allele is correlated to moderate-to-severe acute graft-versus-host disease after allogeneic stem cell transplant. *Exp Clin Transplant Off J Middle East Soc Organ Transplant*. 2010;8(2):125-129.

70. Nasr A, Allam G, Hamid O, Al-Ghamdi A. IFN-gamma and TNF associated with severe falciparum malaria infection in Saudi pregnant women. *Malar J*. 2014;13:314.

71. Harkensee C, Oka A, Onizuka M, et al. Single nucleotide polymorphisms and outcome risk in unrelated mismatched hematopoietic stem cell transplantation: an exploration study. *Blood*. 2012;119(26):6365-6372.

72. Guo W, Dong Z, Guo Y, et al. Polymorphisms of transforming growth factor-β1 associated with increased risk of gastric cardia adenocarcinoma in north China. *Int J Immunogenet*. 2011;38(3):215-224.

73. Sandoval-Pinto E, Padilla-Gutiérrez JR, Valdés-Alvarado E, et al. Association of the −1031T>C polymorphism and soluble TNF-α levels with acute coronary syndrome. *Cytokine*. 2016;78(suppl C):37-43.

74. Nourian M, Chaleshi V, Pishkar L, et al. Evaluation of tumor necrosis factor (TNF)-α mRNA expression level and the rs1799964 polymorphism of the TNF-α gene in peripheral mononuclear cells of patients with inflammatory bowel diseases. *Biomed Rep*. 2017;6(6):698-702.

75. Umapathy D, Krishnamoorthy E, Mariappanadar V, Viswanathan V, Ramkumar KM. Increased levels of circulating (TNF-α) is associated with (-308G/A) promoter polymorphism of TNF-α gene in diabetic nephropathy. *Int J Biol Macromol*. 2018;107(Pt B):2113-2121.

76. Nordlander A, Uzunel M, Mattsson J, Remberger M. The TNFd4 allele is correlated to moderate-to-severe acute graft-versus-host disease after allogeneic stem cell transplantation. *Br J Haematol*. 2002;119(4):1133-1136.

77. Silkov AN, Sennikova NS, Goreva EP, Lopatnikova YA, Sennikov SV. Production of TNF-α and IL-1β by peripheral blood mononuclear cells in carriers of different allele variants of the gene. *Bull Exp Biol Med*. 2012;153(1):68-71.

78. Viel DO, Tsuneto LT, Sossai CR, et al. IL2 and TNFA gene polymorphisms and the risk of graft-versus-host disease after allogeneic haematopoietic stem cell transplantation. *Scand J Immunol*. 2007;66(6):703-710.