

CLINICAL TRIALS AND OBSERVATIONS

Plasma biomarkers of risk for death in a multicenter phase 3 trial with uniform transplant characteristics post–allogeneic HCT

Mohammad Abu Zaid,^{1,2} Juan Wu,³ Cindy Wu,³ Brent R. Logan,⁴ Jeffrey Yu,^{1,2} Corey Cutler,⁵ Joseph H. Antin,⁵ Sophie Paczesny,^{1,2,*} and Sung Won Choi^{6,*}

¹Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN; ²Melvin and Bren Simon Cancer Center, Indiana University, Indianapolis, IN; ³The Emmes Corporation, Rockville, MD; ⁴Division of Biostatistics, Medical College of Wisconsin, Milwaukee, WI; ⁵Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA; and ⁶Department of Pediatrics, University of Michigan Medical School, Ann Arbor, MI

Key Points

- High ST2 and TIM3 at day 28 after allogeneic HCT were associated with nonrelapse mortality and overall survival at 2 years.
- Low day 28 L-Ficolin was associated with VOD/SOS and high CXCL9 correlated with chronic GVHD.

A phase 3 clinical trial (BMT CTN 0402) comparing tacrolimus/sirolimus (Tac/Sir) vs tacrolimus/methotrexate (Tac/Mtx) as graft-versus-host disease (GVHD) prophylaxis after matched-related allogeneic hematopoietic cell transplantation (HCT) recently showed no difference between study arms in acute GVHD-free survival. Within this setting of a prospective, multicenter study with uniform GVHD prophylaxis, conditioning regimen, and donor source, we explored the correlation of 10 previously identified biomarkers with clinical outcomes after allogeneic HCT. We measured biomarkers from plasma samples collected in 211 patients using enzyme-linked immunosorbent assay (Tac/Sir = 104, Tac/Mtx = 107). High suppression of tumorigenicity-2 (ST2) and T-cell immunoglobulin mucin-3 (TIM3) at day 28 correlated with 2-year nonrelapse mortality in multivariate analysis ($P = .0050$, $P = .0075$, respectively) and in a proportional hazards model with time-dependent covariates (adjusted hazard ratio: 2.43 [1.49–3.95], $P = .0038$ and 4.87 [2.53–9.34], $P < .0001$, respectively). High ST2 and TIM3 correlated with overall survival.

Chemokine (C-X-C motif) ligand 9 (CXCL9) levels above the median were associated with chronic GVHD compared with levels below the median in a time-dependent proportional hazard analysis ($P = .0069$). Low L-Ficolin was associated with hepatic veno-occlusive disease ($P = .0053$, AUC = 0.80). We confirmed the correlation of plasma-derived proteins, previously assessed in single-center cohorts, with clinical outcomes after allogeneic HCT within this prospective, multicenter study. (*Blood*. 2017;129(2):162-170)

Introduction

Several plasma biomarkers that correlate with clinical outcomes after allogeneic hematopoietic cell transplantation (HCT) have been identified: suppression of tumorigenicity-2 (ST2) with therapy-resistant acute graft-versus-host disease (GVHD) and nonrelapse mortality (NRM)¹⁻³; regenerating islet-derived 3- α (Reg3 α) and T-cell immunoglobulin mucin-3 (TIM3) with gastrointestinal acute GVHD³⁻⁷; interleukin-6 (IL-6) with acute GVHD⁸; ST2, chemokine (C-X-C motif) ligand 9 (CXCL9), matrix metalloproteinase 3 (MMP3), and osteopontin (OPN) with chronic GVHD^{9,10}; and L-Ficolin, hyaluronic acid (HA), vascular cell adhesion molecule-1 (VCAM1), and ST2 with hepatic veno-occlusive disease (VOD) or sinusoidal obstruction syndrome (SOS).¹¹

The Blood and Marrow Transplant Clinical Trials Network (BMT CTN) 0402 study that prospectively compared tacrolimus/sirolimus (Tac/Sir) with tacrolimus/methotrexate (Tac/Mtx) GVHD prophylaxis found no difference in day 114 acute GVHD-free survival in HLA-matched related donor HCT.¹² In addition, there were no differences in grade 2 to 4 acute GVHD, chronic GVHD, relapse-free survival, and overall survival (OS) at 2 years between study arms. Therefore, we

investigated whether a selected set of previously validated plasma-derived biomarkers¹⁻¹¹ would correlate with clinical outcomes using samples collected from patients within this prospective, multicenter setting of uniform GVHD prophylaxis, conditioning regimen (full-intensity), and donor source (HLA-matched related).

Patients and methods

Study population

Peripheral blood samples were obtained from study participants at predetermined time points after HCT (days 28, 100, 180, and 365) in accordance with the BMT CTN 0402 protocol.¹² The study was an open-label, phase 3, multicenter, randomized trial that included eligible subjects <60 years of age and undergoing transplantation for acute leukemia in remission, myelodysplastic disorder, or chronic myeloid leukemia in chronic or accelerated phase. Enrollment began in November 2006 and ended in October 2011, and all subjects were followed for 2 years. The study was approved by the Protocol Review Committee and

Submitted 22 August 2016; accepted 26 October 2016. Prepublished online as *Blood* First Edition paper, 8 November 2016; DOI 10.1182/blood-2016-08-735324.

*S.P. and S.W.C. contributed equally to this study as joint senior authors.

The online version of this article contains a data supplement.

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

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

© 2017 by The American Society of Hematology

Table 1. Patient characteristics (N = 211)

Variable	Tac/Sir	Tac/Mtx	P value
Number of patients	104	107	
Underwent transplantation	104 (100)	107 (100)	
Age, median, y (range)	45 (19-59)	41 (13-58)	.079
Male sex	53 (51)	48 (45)	.41
Primary malignancy			.19
Acute myelogenous leukemia	46 (44)	40 (37)	
Acute lymphoblastic leukemia	38 (37)	53 (50)	
Chronic myelogenous leukemia	8 (8)	9 (8)	
Myelodysplastic syndrome	11 (11)	5 (5)	
Acute biphenotypic leukemia	1 (1)	0 (0)	
Disease status at transplantation			
Acute myelogenous leukemia			.58
1st complete remission	37 (80)	34 (85)	
2nd complete remission	9 (20)	6 (15)	
Acute lymphoblastic leukemia			.86
1st complete remission	31 (82)	44 (83)	
2nd complete remission	7 (18)	9 (17)	
Chronic myelogenous leukemia			.60
Chronic phase	7 (88)	7 (78)	
Accelerated phase	1 (13)	2 (22)	
Acute biphenotypic leukemia			
1st complete remission	1		
Karnofsky score			.97
90-100%	72 (69)	77 (72)	
<90%	32 (31)	30 (28)	
Recipient-donor CMV status			.028
+/+	41 (39)	33 (31)	
+/-	21 (20)	34 (32)	
-/+	9 (9)	18 (17)	
-/-	27 (26)	17 (16)	
Missing	6 (6)	5 (5)	
Recipient CMV status			.66
+	66 (63)	71 (66)	
-	38 (37)	36 (34)	
Donor-recipient sex match			.77
Female-male	29 (28)	30 (28)	
Conditioning regimen			
Cyclophosphamide/total body irradiation	85 (82)	84 (79)	.56
Etoposide/total body irradiation	19 (18)	23 (22)	

the Data Safety Review Committee of the National Heart, Lung, and Blood Institute and also by the Institutional Review Boards of all participating centers. All subjects provided written informed consent before enrollment. The study was conducted in accordance with the Declaration of Helsinki. All authors vouched for the accuracy and completeness of the reported data, analyses, and the adherence of the study protocol.

Sample preparation and ELISA

All blood samples (either serum or plasma) were prospectively collected and stored per institutional guidelines. The frozen samples were shipped to the Paczesny Laboratory at the University of Indiana (Indianapolis, Indiana) for analysis. ST2, IL-6, Reg3 α , and TIM3 were measured on days 28, 100, 180, and 365 as previously examined in the acute GVHD setting.¹⁻⁸ ST2, CXCL9, OPN, and MMP3 were measured at days 100, 180, and 365, as previously examined in the chronic GVHD setting.^{9,10} L-Ficolin, HA, and VCAM1 were measured on day 28 only, as previously examined in VOD/SOS.¹¹ All of these biomarkers were measured using sequential enzyme-linked immunosorbent assay (ELISA), as previously reported.¹³ The antibody pairs included Reg3 α (MBL International, Ab-Match Assembly Human PAP1 [Reg3 α] kit) and Ab-Match Universal kit), CXCL9 (RayBiotech, RayBio Human MIG ELISA Kit), L-Ficolin (Hycult Biotech, HK336 Human Ficolin-2 ELISA kit), and HA (Corgenix HA test kit). Duoset kits were used for IL-6, MMP3, TIM3, OPN, and VCAM1, and quantikine kit for ST2 (R&D Systems). All the kits permitted comparable measurements in plasma or serum; thus, the ST2 Duoset kit was not

used for this study. Samples were analyzed in duplicate, as previously described.¹³ Pipetting for the Reg3 α assay (384-well plate format) was performed using the EpMotion 4500 liquid handling system (Eppendorf) and for other assays (96-well plate format) by multichannel or the Multidrop 384 Reagent Dispenser (Thermo Scientific). All washes were performed using the Aquamax 2000 plate washer (Molecular Devices). Absorbance was measured immediately after termination of the substrate reaction using a SpectraMax Plus plate reader (Molecular Devices), and results were calculated using SoftMax Pro, version 6.2.2 (Molecular Devices). Laboratory investigators were blinded to all clinical information and transplant outcomes.

Statistical analysis

Demographics and baseline characteristics of study participants with available biomarker and clinical data were described by median and range for continuous variables and by frequency and percent for categorical variables. GVHD prophylaxis arms (Tac/Sir and Tac/Mtx), henceforward referred to as treatment arms or groups, were compared using Wilcoxon rank-sum test for the continuous variables and χ^2 test or Fisher's exact test for the categorical variables. Descriptive statistics were assessed for each biomarker at the different time points. Biomarker levels were log transformed to adjust for the nonnormality of the data distribution. Biomarker values were treated as continuous variables for ST2, Reg3 α , TIM3, OPN, and MMP3. The median levels of each marker at each time point was used as a predictor in the different analyses. IL-6 and CXCL9 were grouped into 2 categories (high vs low) due to high frequent values below the assay detection level. "Low" means below the median, which includes the undetectable values; "high" means above median. Wilcoxon rank-sum test was used to compare median biomarker values between the 2 treatment groups at each time point. Correlation among the biomarkers was explored using pairwise Spearman rank correlation coefficients. ST2, IL-6, Reg3 α , and TIM3 were used in the analysis for acute GVHD, NRM, OS, and relapse. ST2, MMP3, OPN, and CXCL9 were used in the analysis for chronic GVHD.¹⁰ ST2, L-Ficolin, HA, and VCAM at day 28 were used in the analysis for VOD/SOS.¹¹ Univariate logistic regression was used to evaluate the association between clinical outcomes and individual biomarker at each time point for the entire cohort. Multivariate analyses were performed, including the treatment group and 6 covariates: age (≥ 40 , < 40), sex, performance score (90% to 100%, $< 90\%$), conditioning regimen (cyclophosphamide/total body irradiation, etoposide/total body irradiation), primary malignancy (acute myeloid leukemia, acute lymphoblastic leukemia, or others), and recipient cytomegalovirus (CMV) status. Covariates associated with each outcome at a significance level of 0.1 were selected for the multivariate analysis. The treatment arm was always included in the model. For chronic GVHD, NRM, OS, and relapse outcomes, proportional hazards model with time-dependent covariates for single biomarker adjusting for treatment arm and selected covariates was explored. For GVHD outcomes, a landmark analysis was used, wherein only patients who are still at risk for GVHD were included to evaluate the association of GVHD outcome with biomarker at a particular time point. Receiver operating characteristics (ROC) curve from logistic regression models with area under the curve (AUC) was used to present the correlation of clinical outcomes and biomarkers. Cumulative incidence analysis was used to describe the association between NRM and biomarkers with relapse as a competing risk. P value $\leq .01$ was considered statistically significant. Clinical outcomes were defined as reported in the BMT CTN 0402 study.¹² All statistical analyses were done with SAS software (version 9.3) except that cumulative incidence analyses were done with R software (version 2.15.1).

Results

Patient and graft demographics

Biomarker data were available for a total of 211 patients of 304 enrolled on the BMT CTN 0402 study. A total of 714 blood samples ($> 80\%$ were plasma) were analyzed (day 28: N = 197; day 100: N = 197; day 180: N = 177; and day 365: N = 143). The characteristics of the 211 patients in these analyses are shown in Table 1. The median age of participants was 45 years (range, 13 to 59 years). There were no

Table 2. Summary of the clinical outcomes of study participants (N = 211)

Outcome	Event time	Count	Tac/Sir	Tac/Mtx	P value*
Acute GVHD grade 2-4	Never	141	71	70	.73
	By day 28	22	14	8	
	After day 28	48	19	29	
Chronic GVHD†	Never	76	32	44	.09
	By day 100	4	3	1	
	After day 100	131	69	62	
	After day 180	95	51	44	
	After day 365	19	10	9	
NRM	After day 730	1	0	1	.46
	Never	188	91	97	
	By day 28	0	0	0	
	After day 28	23	13	10	
	After day 100	23	13	10	
	After day 180	20	11	9	
Survival status by 2 y (day 730)	After day 365	9	5	4	.57*
	After day 730	1	0	1	
	Yes (alive)	157	76	81	
	No (dead)	54	28	26	
	Relapse/progression	159	80	79	
Relapse/progression	Never	159	80	79	.57
	By day 28	1	1	0	
	After day 28	51	23	28	
	After day 100	42	20	22	
	After day 180	27	14	13	
	After day 365	12	7	5	
VOD/SOS	After day 730	0	0	0	.08
	Never	197	94	103	
	By day 28	6	4	2	
Total no. of patients	After day 28	8	6	2	
	Biomarker data available	211	104	107	

*P value from Gray's test for the cumulative incidence of the outcome, except for OS where the P value is from a log-rank test.

†Incidence of chronic GVHD was slightly higher in this study compared with the original BMT CTN 0402 study (64% vs 49%).

differences between study arms with regards to age, sex, primary malignancy, stage of malignancy at transplantation, and performance status at transplantation, or conditioning regimen. Recipient-donor CMV serostatus combinations were slightly different between the 2 groups. A summary of the clinical outcomes of interest is shown in Table 2.

Biomarker levels by treatment arm

Correlation among the biomarkers was examined at day 28 using pairwise Spearman rank correlation coefficients (supplemental Table 1, available on the *Blood* Web site). TIM3 levels were significantly correlated with ST2, IL-6, and Reg3 α at day 28. There were significant differences in ST2 and Reg3 α levels between the treatment arms. At day 28, ST2 was significantly higher and Reg3 α significantly lower in the Tac/Sir group compared with the Tac/Mtx group ($P < .01$; supplemental Table 2; supplemental Figure 1). The multivariate models found no interactions between biomarkers levels and different treatment groups. L-Ficolin and HA were significantly different between the treatment arms (supplemental Table 2; supplemental Figure 1), possibly due to higher incidence of VOD/SOS events in the Tac/Sir arm (Table 2). Based on these differences in biomarker levels between the treatment arms, the treatment arm was included as a covariate in the multivariate model for all of the analyses.

Biomarker levels and nonrelapse mortality

In the multivariate logistic regression analysis for the entire cohort, high ST2 and TIM3 levels at day 28 were associated with significantly higher risk of NRM at 2 years ($P = .0050$ and $.0075$, respectively;

Figure 1A). A 4-biomarker panel, including ST2, TIM3, IL-6, and Reg3 α , adjusted for the clinical covariates had an area under the ROC curve of 0.80 for NRM at 2 years (Figure 1B). The cumulative incidence of NRM at 2 years for day 28 ST2 \geq median was 17.8% (95% confidence interval: 10.9%, 26.1%) compared with 5.2% (95% confidence interval: 1.9%, 11%) for day 28 ST2 $<$ median (Gray's test $P = .008$; Figure 2A). When a proportional hazards model with time-dependent covariates was explored, IL-6 and Reg3 α were also found to be associated with NRM. In this model, a 1-log increase in ST2 and TIM3 had a hazard ratio (HR) of 2.43 (1.49-3.95) ($P = .00038$) and 4.87 (2.53-9.34) ($P < .0001$) for NRM at 2 years, respectively (Table 3). Supplemental Table 3 shows the causes of NRM by 2 years by biomarker level. There were no significant differences in causes of NRM based on low- or high-biomarker level (ST2 or TIM3).

Biomarker levels and risk of relapse

Although a trend toward low ST2 and relapse ($P = .040$) was observed in the proportional hazard model (supplemental Table 4), none of the biomarkers were associated with relapse.

Biomarker levels and OS

In the adjusted proportional hazards model with time-dependent covariates, the 4 markers (ST2, TIM3, IL-6, and Reg3 α) were shown to correlate with OS at 2 years (Table 3). As most deaths occurred after day 180, we next evaluated ST2 value at day 180 for OS at 2 years. For patients with a high ST2 value, 2-year OS was 69.8%, whereas for those with a low ST2 value, the rate was 84.6% ($P = .032$; Figure 2B).

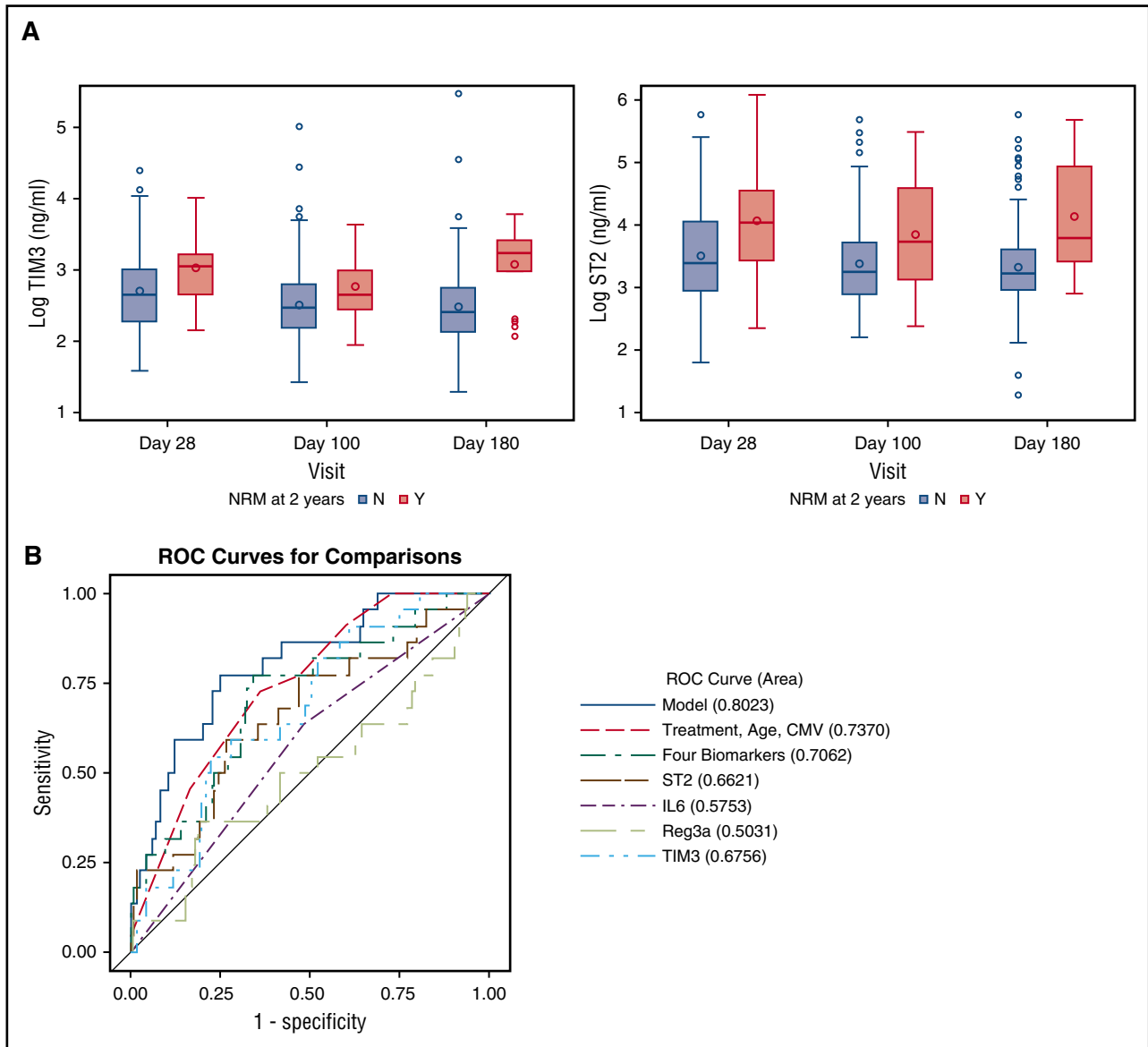


Figure 1. High ST2 and TIM3 are associated with NRM by 2 years. (A) Absolute values on a logarithmic scale for ST2 and TIM3 at days 28, 100, and 180 with their medians and interquartile range, $P < .01$ for all comparisons except for TIM3 on day 100 ($P = .037$). (B) ROC curve for NRM by 2 years with AUC for post-HCT day 28 ST2, IL-6, Reg3a, TIM3, the 4 biomarkers combined, and clinical covariates in the combined model. The biomarker panel improves the AUC from 0.74 for the clinical covariates to 0.80 in the combined model. N, patients who did not die from NRM; Y, patients who died from NRM.

Biomarker levels and acute GVHD

The earliest predetermined blood sample available was day 28, at which time 22 of 70 events of the acute GVHD already occurred and were therefore excluded from the analysis. None of the biomarkers evaluated in this study correlated with acute GVHD. There was a trend toward high ST2 at day 28 being correlated with grade 2 to 4 acute GVHD ($P = .09$) (supplemental Table 5). Different cutoffs were used based on previously published data² to show sensitivity and specificity of day 28 ST2 in detecting grade 2 to 4 acute GVHD by day 100 in this cohort (supplemental Table 6).

Biomarker levels and chronic GVHD

ST2, OPN, MMP3, and CXCL9 were previously shown to correlate with chronic GVHD.¹⁰ In this study, although ST2, OPN, and

MMP3 measured at day 100 did not correlate with chronic GVHD, day 100 CXCL9 values trended toward an association with chronic GVHD outcome ($P = .029$, AUC = 0.65). However, in an analysis where time-dependent biomarker levels were used in the proportional hazards model, which takes the most recent biomarker data available for the patient prior to the chronic GVHD event, CXCL9 levels that were above the median had a statistically significant HR of 1.60 (1.14, 2.26) ($P = .0069$) compared with those below the median and undetectable levels (Table 3). These data suggest that the CXCL9 value closer to chronic GVHD is more predictive of subsequent chronic GVHD.

Biomarker levels and VOD/SOS

Fourteen of the 197 evaluable study participants (7%) developed VOD/SOS (Tac/Sir: N = 10 and Tac/Mtx: N = 4). Six patients developed

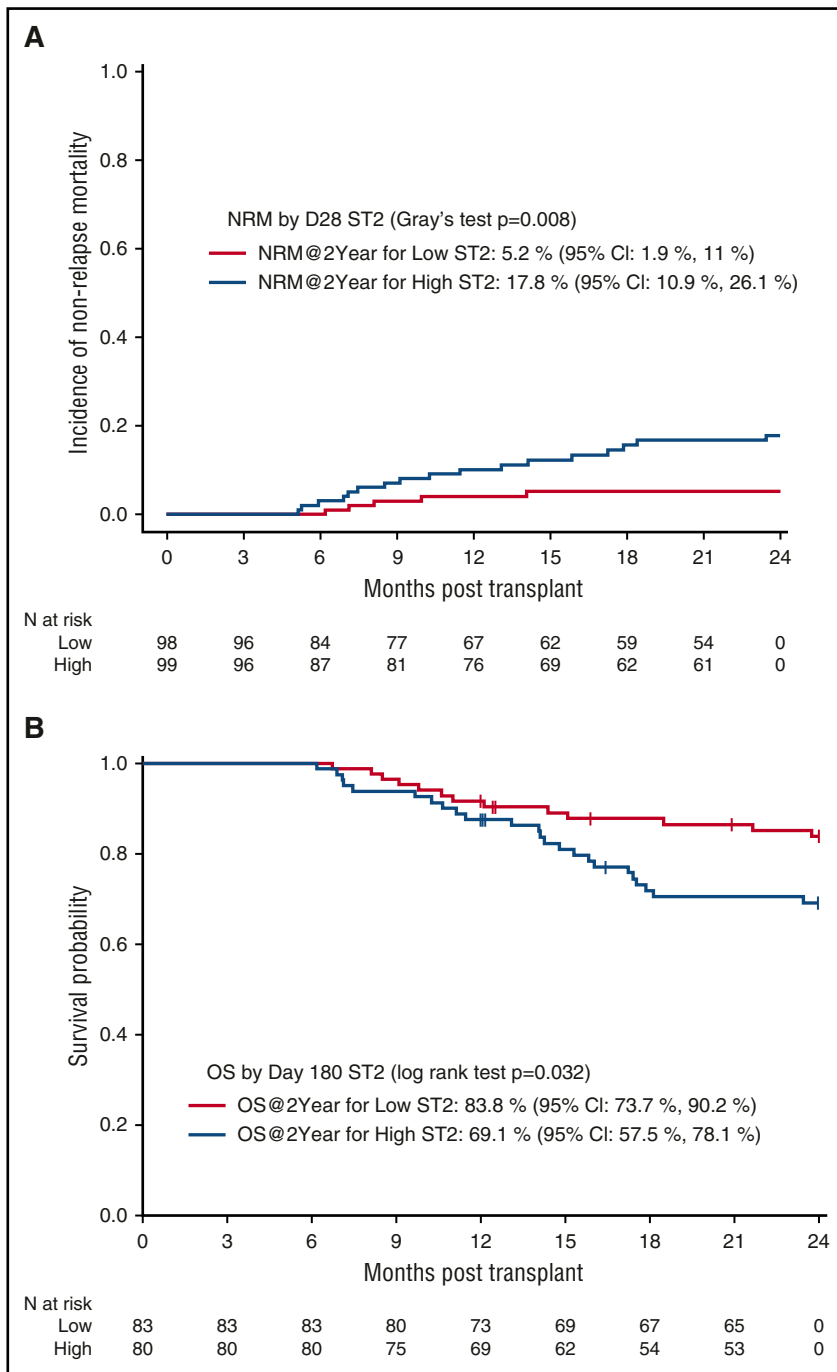


Figure 2. High ST2 is correlated with future NRM and OS. (A) The cumulative incidence of NRM by 2 years stratified by day 28 ST2 levels (high vs low, median cutoff of 31 ng/mL). (B) Kaplan-Meier curve for OS stratified by day 180 ST2 levels (high vs low, median cutoff of 26.2 ng/mL). CI, confidence interval; N, number.

VOD/SOS before day 28, while 8 patients developed VOD/SOS after day 28. In the multivariate analysis, low concentrations of L-Ficolin on day 28 were associated with VOD/SOS ($P = .0053$; Table 4) as recently published.¹¹ High concentrations of VCAM1 trended toward an association with VOD/SOS ($P = .017$). L-Ficolin levels for patients with and without VOD/SOS are shown as representative in Figure 3A. ST2 and HA were not significantly associated with VOD/SOS (Table 4). The 4-biomarker panel had an AUC of 0.81 for a diagnosis of VOD/SOS (Table 4; Figure 3B). For patients with a low L-Ficolin value, VOD/SOS cumulative incidence was 5.1%, whereas for patients with a high L-Ficolin value, the cumulative incidence was 2% (Figure 3C).

Discussion

Noninvasive peripheral blood biomarkers have the potential to improve diagnosis and facilitate therapeutic management of complications after allogeneic HCT. Despite significant progress made in recent years, biomarkers are not yet available for clinical use in the HCT setting. In this paper, we have shown that plasma biomarkers measured from prospective sample collections of a BMT CTN study with uniform HLA-matched related donor and myeloablative conditioning strategy were associated with transplant-related complications, including NRM, death, and VOD/SOS. Consistent with recent studies, high ST2 and

Table 3. Correlation between biomarkers and clinical outcomes

Biomarker	P value	Hazard ratio	HR lower CL	HR upper CL
Nonrelapse mortality by 2 y*				
ST2	.00038	2.43	1.49	3.95
IL-6†	.0090	3.80	1.39	10.33
Reg3α	.00073	2.35	1.43	3.87
TIM3	<.0001	4.87	2.53	9.34
Overall survival‡				
ST2	.0014	1.73	1.24	2.41
IL-6†	.0022	2.42	1.37	4.25
Reg3α	<.0001	2.00	1.46	2.73
TIM3	.00035	1.87	1.33	2.65
Chronic GVHD§				
ST2	.30	1.13	0.89	1.44
CXCL9†	.0069	1.60	1.14	2.26
OPN	.06	0.73	0.52	1.01
MMP3	.90	0.99	0.85	1.16

*Proportional hazards model with time-dependent covariates adjusting for treatment group, age, and CMV status.

†High vs low (see Patients and methods).

‡Proportional hazards model with time-dependent covariates adjusting for treatment regimen.

§Proportional hazards model with time-dependent covariates adjusting for treatment group and primary malignancy.

TIM3 values were strongly associated with NRM in this cohort as well. Although we observed a trend toward high ST2 and acute GVHD, it was not entirely surprising that ST2 and the other biomarkers did not correlate significantly with acute GVHD. The earliest sample collection in this study was on day 28, at which time nearly one-third of acute GVHD events had already occurred. Accordingly, these patients were excluded from the analysis, and possibly even diluted the power of detecting a meaningful difference. Moreover, in previously published studies of ST2,^{1,3} TIM3,³ and IL-6,⁸ these biomarkers were also measured at an earlier time point post-HCT (day 14). Another difference of the current study is that all patients with acute GVHD were included regardless of organ involvement, whereas some prior studies of TIM3³ and Reg3α⁴ focused on gastrointestinal acute GVHD.

In the time-dependent proportional hazard models, we found that high ST2, IL-6, Reg3α, and TIM3 were associated with increased risk of NRM at 2 years. Additional large studies involving many patients from multiple centers with longitudinal blood sample collections are needed to further examine the specificity and sensitivity of the biomarkers for GVHD prediction relative to other clinical outcomes, such as NRM and death, after allogeneic HCT.

Based on the recently identified association of a 4-biomarker panel (ST2, CXCL9, MMP3, and OPN) with chronic GVHD,¹⁰ and previous association of CXCL9, discovered through a proteomic approach, with chronic GVHD,⁹ we measured these biomarkers herein and found that high CXCL9 and not the other markers was strongly associated with chronic GVHD. This association was only significant when using a proportional hazards model where time-dependent biomarker levels were used while the correlation of CXCL9 at a single time point (ie, day 100) was moderate, which is in agreement with the recently published study where the panel measured at day 100 could only predict chronic GVHD occurring within the next 3 months.¹⁰

We found that low L-Ficolin was strongly associated with VOD/SOS, consistent with recently published findings.¹¹ L-Ficolin is a complement-activating pattern-recognition lectin involved in the innate immune response and has recently been shown to be involved in

homeostatic clearance of mitochondria in the liver.¹⁴ In VOD/SOS patients, the concentrations of L-Ficolin were decreased, suggesting that this homeostatic clearance may no longer happen efficiently, and possibly implicating pathways involved in VOD/SOS other than those related to hemostasis and endothelial injury.

A major strength of our study is that the study population was relatively homogenous with regards to important variables known to affect evaluation of biomarkers. Namely, all patients received a myeloablative regimen, had an HLA-matched sibling donor, and received tacrolimus (with sirolimus or methotrexate) as their GVHD prophylaxis. The other strength was the measurement of biomarkers at fixed time points after transplant within the setting of routine clinical care from multiple centers.

The markers investigated in this study have been shown to be of biological and possibly therapeutic importance in GVHD. ST2 is the receptor of IL-33 and is present in 2 main isoforms: a membrane-bound form (mST2) and a soluble form (sST2).¹⁵ sST2 sequesters IL-33, limiting its availability to T cells expressing mST2 (T-helper 2 cells and ST2⁺FoxP3⁺ regulatory T cells). In a recent work by our group, we found that blockade of sST2 in the peritransplant period with a neutralizing monoclonal antibody reduced GVHD severity and mortality in a GVHD murine model.¹⁶ ST2 blockade also reduced sST2 production by IL-17-producing T cells while maintaining protective mST2-expressing T cells, increasing the frequency of intestinal myeloid-derived suppressor cells and decreasing the frequency of intestinal CD103 dendritic cells. Finally, ST2 blockade preserved graft-versus-leukemia activity in a model of MLL-AF9 acute myeloid leukemia. ST2 could be a therapeutic target for severe GVHD. Similarly, TIM3 has 2 forms, a soluble form and a membrane-bound form.⁷ It is speculated that plasma TIM3 may exacerbate the severity of acute GVHD by blocking the interaction between TIM3 and its ligand, thereby abrogating the regulatory activity of the TIM3 pathway. CXCL9 is an interferon-γ-inducible chemokine that binds to CXCR3, its only known receptor, and promotes lymphocyte migration to inflamed tissues.^{17,18} CXCR3 has also been shown to be critical for the recruitment of alloreactive T cells in acute GVHD,^{19,20} whereas CXCL9 has been shown to be elevated in tissue samples from patients with oral,²¹ ocular,²² and cutaneous²³ chronic GVHD.

Nonetheless, we recognize the limitations of our analyses herein. First, day 28 post-HCT samples were the first available that excluded nearly one-third of the patient population available for acute GVHD analyses. An a priori power calculation was not conducted for this study, as the power was limited by the available samples from the main clinical trial. We conducted a post-hoc assessment of power to detect differences in acute GVHD incidence. The overall incidence of acute GVHD grade 2 to 4 in this cohort among those still at risk at day 28 is 26%. The sample size of 176 evaluable patients would have at least 80% power to detect an odds ratio (OR) of 2.78 (corresponding to an approximate difference in incidence of 19%), between 2 equally sized

Table 4. Biomarkers at day 28 and VOD/SOS occurrence

Biomarker	Day 28 (14 events/197 used)			
	Multivariate logistic regression analysis*		Biomarker panel*	
	P value	AUC	P value	AUC
ST2	.064	0.75	.54	0.81
L-Ficolin	.0053	0.80	.012	
HA	.17	0.73	.18	
VCAM	.017	0.74	.13	

*Adjusted for treatment group, performance score, and CMV status.

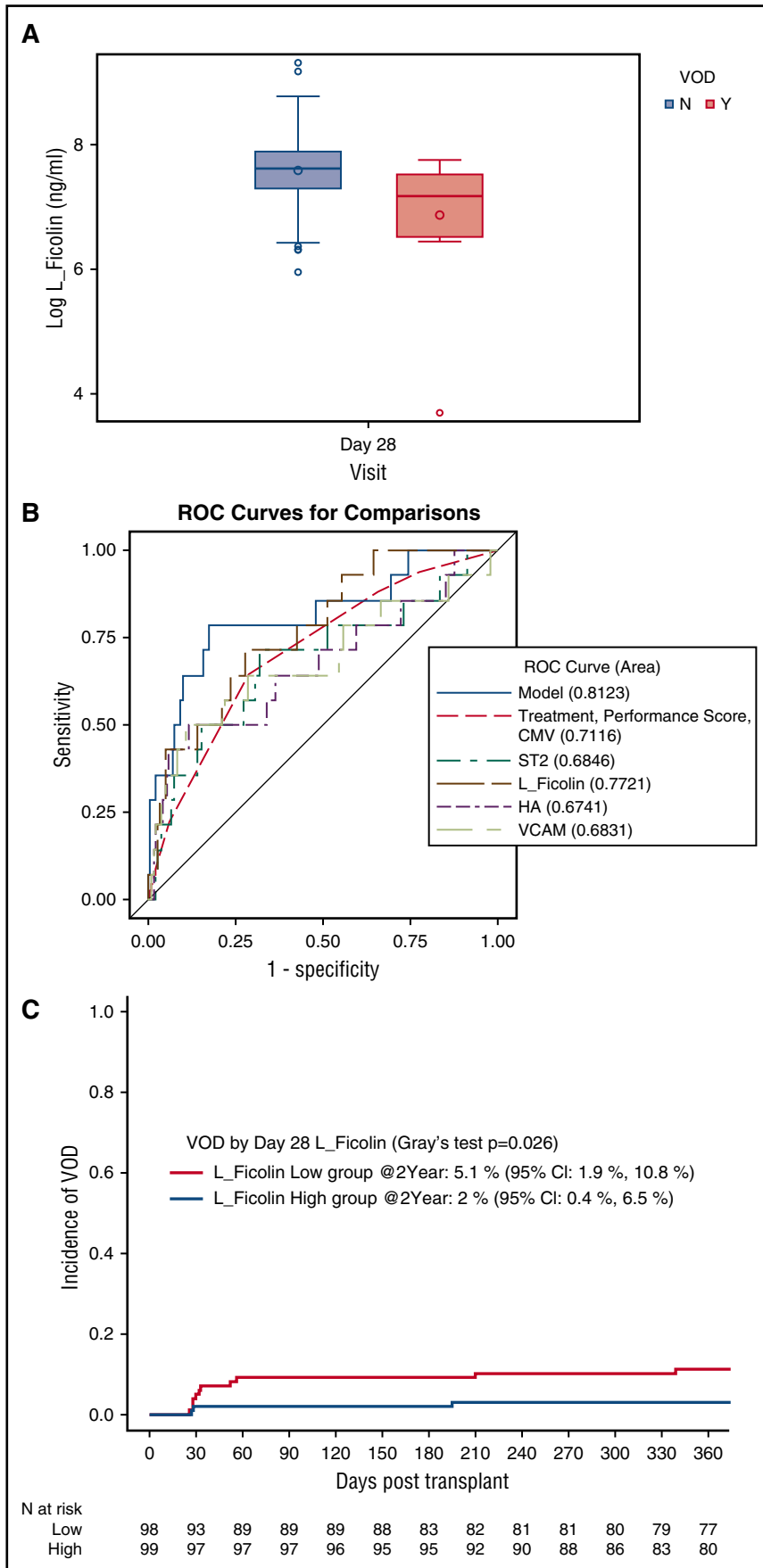


Figure 3. Low L-Ficolin is correlated with hepatic VOD. (A) Concentrations of day 28 L-Ficolin (median and interquartile range) for patients who developed VOD (Y) and those who did not (N) ($P = .005$). (B) ROC curves for day 28 L-Ficolin, ST2, HA, and VCAM, the clinical covariates, and the 4-biomarker panel (model). (C) Cumulative incidence of VOD stratified by day 28 L-Ficolin (high vs low).

biomarker groups defined by above vs below the median. The detectable difference compares favorably with what was reported previously (OR for high vs low panel of 2.7, OR for high vs low ST2 of 3.7).¹ The ideal timing for sampling would be earlier, in our opinion (day 7 to day 21), for purposes of acute GVHD prediction and probably even earlier for VOD/SOS prediction (day 0 to day 10). Biomarkers measured pretransplant on day of stem cell infusion (day 0) are unlikely to be informative of an allo-reaction, which is consistent with our experience with earlier studies.¹ More frequent sample collection early after transplant that is balanced with clinical feasibility and cost-effectiveness would be crucial for future biomarker studies. Indeed, this is actively being done in the BMT CTN 1202 multicenter biomarker study, which is collecting samples weekly starting at day 0 until day 28 with additional samples on day 42, 56, and 91. This study recently completed accrual and it is hoped will help advance the biomarker field for posttransplant complications.

Second, there were relatively few grade 2 to 4 acute GVHD events in each treatment arm, in general, that possibly limited a more robust analysis. However, we did adjust for the treatment arm in the multivariate analysis for all of the analyses. When a subgroup analysis was attempted, the direction of the relationship between the markers and outcomes was in the same direction, but did not uniformly reach statistical significance. Third, we were limited by the lack of clinical data regarding specific acute GVHD organ involvement as well as degree of severity of chronic GVHD. Unfortunately, analyses including these variables were not possible. Fourth, we recognize that other biomarkers, such as TNFR1,²⁴ BAFF,²⁵ and CXCL10,^{26,27} have been reported in acute and chronic GVHD, respectively, which were not explored in this study. We specifically sought to confirm the validity of biomarkers that we have previously validated in at least 2 cohorts at the time of the analysis.

Significant developments in the field of biomarkers for post-HCT outcomes have been made in recent years. Many studies have shown different markers to be associated with 1 or several important outcomes. The question remains on how to move these findings to the clinic in real time. One approach is to use these biomarkers in a preemptive fashion that can guide more aggressive immunosuppressive therapy for patients at high risk for GVHD. In such cases, health care providers may want to use a higher cutoff for the biomarkers (eg, ST2 > 50 ng/mL). Moreover, in the case where a more rapid taper of immunosuppressive agents is favored, such as in patients at high risk of relapse, low biomarker levels (ST2 < 35 ng/mL) may possibly guide treatment decisions that favor lower risk of GVHD and NRM. In this study, we specifically used thresholds to correlate with clinical outcomes. Similarly, CXCL9 levels on day 100 or day 180 could guide preemptive therapy for those at higher risk for chronic GVHD based on biomarker levels. Also, we are in the midst of proposing a clinical trial utilizing a VOD/SOS biomarker panel to guide preemptive defibrotide, a recently approved therapy for established VOD/SOS,²⁸ prior to onset of clinical features. We recognize that further refinement may be needed for specific thresholds depending on the HCT setting (GVHD prophylaxis, donor source, and/or intensity). Nonetheless, it is likely that this approach may be more feasible to implement across providers, multicenter institutions, and larger clinical populations,

which could prove useful in guiding therapy for patients with clinical features in between “low” and “very high” risk for mortality.²⁹ Alternatively, another potential approach is to include biomarkers in risk-stratification models.^{24,30,31} For example, a recent study developed a biomarker-based algorithm to predict NRM within 6 months of diagnosis of acute GVHD.²⁴ The challenge of incorporating this complex model system will be tested in a future BMT CTN prospective multicenter clinical trial (BMT CTN 1501). It is clear that further studies will be needed to determine which approach will be more feasible and more accurate in defining risk. Importantly, whether therapy dictated by such approaches leads to improved clinical outcomes remains an unanswered question that will require well-designed, prospective, multicenter collaborative trials.

Acknowledgments

The authors would like to thank all the patients and investigators on the BMT CTN 0402 study.

This work was supported by grants from the National Institutes of Health (NIH), National Cancer Institute (R01CA168814) (S.P.) and (K23AI091623) (S.W.C.), the A. Alfred Taubman Medical Research Institute Emerging Scholar/Edith Briskin and Shirley K. Schlafer Foundation Award (S.W.C.), the Leukemia and Lymphoma Society Scholar Award (1293-15) (S.P.), and the Lilly Physician Scientist Initiative Award (S.P.). Support for this study was provided by grant U10HL069294 to the BMT CTN from the NIH, National Heart, Lung, and Blood Institute and the NIH, National Cancer Institute, along with contributions by Wyeth Pharmaceuticals Inc.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the abovementioned parties.

Authorship

Contribution: M.A.Z., J.Y., and S.P. performed experiments; J.W., C.W., and B.R.L. analyzed results and made the figures; S.P., S.W.C., J.H.A., and C.C. designed the research; M.A.Z., S.P., and S.W.C. wrote the manuscript, and all authors reviewed and approved the manuscript.

Conflict-of interest disclosure: S.P. has a patent on “Methods of detection of graft-versus-host-disease” licensed to Viracor-IBT Laboratories. The remaining authors declare no competing financial interests.

Correspondence: Sung Won Choi, University of Michigan Medical School, 1500 East Medical Center Dr, D4118 Medical Professional Building, Ann Arbor, MI 48109-5718; e-mail: sungchoi@med.umich.edu; and Sophie Paczesny, Indiana University School of Medicine, Departments of Pediatrics and Immunology, 1044 West Walnut St, Room 425, Indianapolis, IN 46202; e-mail: sophpacz@iu.edu.

References

- 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.
- Ponce DM, Hilden P, Mumaw C, et al. High day 28 ST2 levels predict for acute graft-versus-host disease and transplant-related mortality after cord blood transplantation. *Blood*. 2015;125(1):199-205.
- McDonald GB, Tabellini L, Storer BE, Lawler RL, Martin PJ, Hansen JA. Plasma biomarkers of acute GVHD and nonrelapse mortality: predictive value of measurements before GVHD onset and treatment. *Blood*. 2015;126(1):113-120.
- Ferrara JL, Harris AC, Greenon JK, et al. Regenerating islet-derived 3-alpha is a biomarker of gastrointestinal graft-versus-host disease. *Blood*. 2011;118(25):6702-6708.

5. Levine JE, Logan BR, Wu J, et al. Acute graft-versus-host disease biomarkers measured during therapy can predict treatment outcomes: a Blood and Marrow Transplant Clinical Trials Network study. *Blood*. 2012;119(16):3854-3860.
6. Harris AC, Ferrara JLM, Braun TM, et al. Plasma biomarkers of lower gastrointestinal and liver acute GVHD. *Blood*. 2012;119(12):2960-2963.
7. Hansen JA, Hanash SM, Tabellini L, et al. A novel soluble form of Tim-3 associated with severe graft-versus-host disease. *Biol Blood Marrow Transplant*. 2013;19(9):1323-1330.
8. Kennedy GA, Varelias A, Vuckovic S, et al. Addition of interleukin-6 inhibition with tocilizumab to standard graft-versus-host disease prophylaxis after allogeneic stem-cell transplantation: a phase 1/2 trial. *Lancet Oncol*. 2014;15(13):1451-1459.
9. Kitko CL, Levine JE, Storer BE, et al. Plasma CXCL9 elevations correlate with chronic GVHD diagnosis. *Blood*. 2014;123(5):786-793.
10. Yu J, Storer BE, Kushekhar K, et al. Biomarker panel for chronic graft-versus-host disease. *J Clin Oncol*. 2016;34(22):2583-2590.
11. Akil A, Zhang Q, Mumaw CL, et al. Biomarkers for diagnosis and prognosis of sinusoidal obstruction syndrome after hematopoietic cell transplantation. *Biol Blood Marrow Transplant*. 2015;21(10):1739-1745.
12. Cutler C, Logan B, Nakamura R, et al. Tacrolimus/sirolimus vs tacrolimus/methotrexate as GVHD prophylaxis after matched, related donor allogeneic HCT. *Blood*. 2014;124(8):1372-1377.
13. Fiema B, Harris AC, Gomez A, et al. High throughput sequential ELISA for validation of biomarkers of acute graft-versus-host disease. *J Vis Exp*. 2012(68).
14. Brinkmann CR, Jensen L, Dagnæs-Hansen F, et al. Mitochondria and the lectin pathway of complement. *J Biol Chem*. 2013;288(12):8016-8027.
15. Iwahana H, Yanagisawa K, Ito-Kosaka A, et al. Different promoter usage and multiple transcription initiation sites of the interleukin-1 receptor-related human ST2 gene in UT-7 and TM12 cells. *Eur J Biochem*. 1999;264(2):397-406.
16. Zhang J, Ramadan AM, Griesenauer B, et al. ST2 blockade reduces sST2-producing T cells while maintaining protective mST2-expressing T cells during graft-versus-host disease. *Sci Transl Med*. 2015;7(308):308ra160.
17. Flier J, Boorsma DM, van Beek PJ, et al. Differential expression of CXCR3 targeting chemokines CXCL10, CXCL9, and CXCL11 in different types of skin inflammation. *J Pathol*. 2001;194(4):398-405.
18. Menke J, Zeller GC, Kikawada E, et al. CXCL9, but not CXCL10, promotes CXCR3-dependent immune-mediated kidney disease. *J Am Soc Nephrol*. 2008;19(6):1177-1189.
19. Piper KP, Horlock C, Curnow SJ, et al. CXCL10-CXCR3 interactions play an important role in the pathogenesis of acute graft-versus-host disease in the skin following allogeneic stem-cell transplantation. *Blood*. 2007;110(12):3827-3832.
20. Duffner U, Lu B, Hildebrandt GC, et al. Role of CXCR3-induced donor T-cell migration in acute GVHD. *Exp Hematol*. 2003;31(10):897-902.
21. Imanguli MM, Swaim WD, League SC, Gress RE, Pavletic SZ, Hakim FT. Increased T-bet+ cytotoxic effectors and type I interferon-mediated processes in chronic graft-versus-host disease of the oral mucosa. *Blood*. 2009;113(15):3620-3630.
22. Westekemper H, Meller S, Citak S, et al. Differential chemokine expression in chronic GVHD of the conjunctiva. *Bone Marrow Transplant*. 2010;45(8):1340-1346.
23. Croudace JE, Inman CF, Abbotts BE, et al. Chemokine-mediated tissue recruitment of CXCR3+ CD4+ T cells plays a major role in the pathogenesis of chronic GVHD. *Blood*. 2012;120(20):4246-4255.
24. Levine JE, Braun TM, Harris AC, et al; Blood and Marrow Transplant Clinical Trials Network. A prognostic score for acute graft-versus-host disease based on biomarkers: a multicentre study. *Lancet Haematol*. 2015;2(1):e21-e29.
25. Sarantopoulos S, Stevenson KE, Kim HT, et al. Altered B-cell homeostasis and excess BAFF in human chronic graft-versus-host disease. *Blood*. 2009;113(16):3865-3874.
26. Karimnia A, Holtan SG, Ivison S, et al. Heterogeneity of chronic graft-versus-host disease biomarkers: association with CXCL10 and CXCR3+ NK cells. *Blood*. 2016;127(24):3082-3091.
27. Paczesny S, Abu Zaid M. CXCL10: most consistent cGVHD biomarker? *Blood*. 2016;127(24):2950-2951.
28. Richardson PG, Riches ML, Kernan NA, et al. Phase 3 trial of defibrotide for the treatment of severe veno-occlusive disease and multi-organ failure. *Blood*. 2016;127(13):1656-1665.
29. McDonald GB. How I treat acute graft-versus-host disease of the gastrointestinal tract and the liver. *Blood*. 2016;127(12):1544-1550.
30. MacMillan ML, Robin M, Harris AC, et al. A refined risk score for acute graft-versus-host disease that predicts response to initial therapy, survival, and transplant-related mortality. *Biol Blood Marrow Transplant*. 2015;21(4):761-767.
31. Holtan SG, MacMillan ML. A risk-adapted approach to acute GVHD treatment: are we there yet? *Bone Marrow Transplant*. 2016;51(2):172-175.