ARTICLE

Cell Therapy & Immunotherapy

Plasma-derived proteomic biomarkers in human leukocyte antigen-haploidentical or human leukocyte antigen-matched bone marrow transplantation using post-transplantation cyclophosphamide

Christopher G. Kanakry,[1] Giorgos Bakoyannis,[2] Susan M. Perkins,[2] Shannon R. McCurdy,[1] Ante Vulic,[1] Edus H. Warren,[3] Etienne Daguindau,[4,5] Taylor Olmsted,[4,5] Christen Mumaw,[4,5] Andrea M.H. Towlerton,[3] Kenneth R. Cooke,[1] Paul V. O'Donnell,[3] Heather J. Symons,[1] Sophie Paczesny[4,5] and Leo Luznik[1]

[1]Sidney Kimmel Comprehensive Cancer Center, The Johns Hopkins University School of Medicine, Baltimore, MD; [2]Department of Biostatistics, Indiana University School of Medicine, Indianapolis, IN; [3]Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA; [4]Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN; [5]Department of Immunology, Indiana University School of Medicine, Indianapolis, IN, USA

Correspondence:

luznile@jhmi.edu or
sophpacz@iu.edu

ABSTRACT

Recent studies have suggested that plasma-derived proteins may be potential biomarkers relevant for graft-*versus*-host disease and/or non-relapse mortality occurring after allogeneic blood or marrow transplantation. However, none of these putative biomarkers have been assessed in patients treated either with human leukocyte antigen-haploidentical blood or marrow transplantation or with post-transplantation cyclophosphamide, which has been repeatedly associated with low rates of severe acute graft-*versus*-host disease, chronic graft-*versus*-host disease, and non-relapse mortality. We explored whether seven of these plasma-derived proteins, as measured by enzyme-linked immunosorbent assays, were predictive of clinical outcomes in post-transplantation cyclophosphamide-treated patients using plasma samples collected at serial predetermined timepoints from patients treated on prospective clinical studies of human leukocyte antigen-haploidentical (n=58; *clinicaltrials.gov Identifier: 00796562*) or human leukocyte antigen-matched-related or -unrelated (n=100; *clinicaltrials.gov Identifiers: 00134017 and 00809276*) T-cell-replete bone marrow transplantation. Day 30 levels of interleukin-2 receptor α, tumor necrosis factor receptor 1, serum STimulation-2 (IL1RL1 gene product), and regenerating islet-derived 3-α all had high areas under the curve of 0.74-0.97 for predicting non-relapse mortality occurrence by 3 months post-transplant in both the human leukocyte antigen-matched and human leukocyte antigen-haploidentical cohorts. In both cohorts, all four of these proteins were also predictive of subsequent non-relapse mortality occurring by 6, 9, or 12 months post-transplant and were significantly associated with non-relapse mortality in univariable analyses. Furthermore, day 30 elevations of interleukin-2 receptor α were associated with grade II-IV and III-IV acute graft-*versus*-host disease occurring after day 30 in both cohorts. These data confirm that plasma-derived proteins previously assessed in other transplantation platforms appear to retain prognostic and predictive utility in patients treated with post-transplantation cyclophosphamide.

Introduction

High-dose, post-transplantation cyclophosphamide (PTCy) provides effective graft-*versus*-host disease (GvHD) prophylaxis after allogeneic blood or marrow transplantation (alloBMT).[1] This approach has facilitated the safe performance of T-cell-replete human leukocyte antigen (HLA)-haploidentical alloBMT[1-3] and can function as single-agent GvHD prophylaxis after myeloablative conditioning and HLA-matched bone marrow allografting.[1,4-6] Despite these clinical successes with low rates of severe acute GvHD, chronic GvHD, and non-relapse mortality (NRM), biomarkers prognostic for GvHD or predictive for NRM occurring despite the use of PTCy have not been explored. Such biomarkers could potentially help guide treatment decisions and direct more intensive clinical surveillance of patients at high-risk for poor outcomes. Furthermore, they may provide biologic insight into immunologic pathways that could be targeted to prevent adverse clinical events.

A number of candidate plasma-derived biomarkers have been examined in other alloBMT platforms,[7-23] and several have repeatedly been found to be associated with clinical outcomes. In the study herein, we focused on seven particularly promising proteins. These proteins are all biologically plausible molecules either related directly to the inflammatory response thought to mediate GvHD (interleukin [IL]-2 receptor α [IL-2Rα],[24] IL-6,[25] tumor necrosis factor receptor 1 [TNFR-1],[26] serum STimulation-2 (IL1RL1 gene product) [ST2],[27,28] and chemokine [C-X-C motif] ligand 9 [CXCL9][29]), or are released from tissue directly damaged by GvHD (lower gastrointestinal tract [regenerating islet-derived 3-α, REG3α][30-32] and skin [elafin][33]). When measured at the start of clinical acute GvHD, elevated plasma levels of IL-2Rα, IL-6, TNFR-1, ST2, REG3α, and elafin have been associated with the presence and predicted severity of acute GvHD, response to immunosuppressive therapy, and/or risk for NRM.[7-9,12-17,21] Plasma levels of CXCL9, IL-2Rα, and elafin measured at the time of onset of chronic GvHD have been associated with chronic GvHD diagnosis; CXCL9 had the highest predictive accuracy and was also associated with chronic GvHD severity.[18]

A few studies have examined the utility of biomarker measurements at pre-determined timepoints. When measured at day 7 after myeloablative alloBMT, elevated plasma levels of TNFR-1 were associated with an increased incidence of NRM and an increased incidence and severity of acute GvHD.[10] Elevated plasma levels of ST2 at days 14 or 21 after myeloablative alloBMT or day 14 after reduced-intensity conditioning alloBMT were associated with an increased risk of NRM.[17] In patients receiving double umbilical cord blood transplantation (dUCBT), elevated day 28 plasma levels of ST2 were associated with an increased incidence of NRM and grade II-IV and III-IV acute GvHD occurring after that timepoint; elevated day 28 plasma levels of TNFR-1 and REG3α, while not associated with acute GvHD, were associated with an increased risk of NRM.[20] Finally, in patients treated with myeloablative HLA-matched-related alloBMT using either tacrolimus/sirolimus or tacrolimus/methotrexate for GvHD prophylaxis, elevated day 28 plasma levels of ST2, REG3α, and IL-6 were all associated with an increased risk of NRM, but were not associated with GvHD risk.[23]

With the increasing use of PTCy worldwide and the clinical observations (e.g., low incidence of chronic GvHD) suggesting that PTCy may modulate GvHD risk in a dissimilar way to other T-cell-replete alloBMT approaches,[1] we explored whether these seven promising candidate biomarkers had utility in prognosticating clinical outcomes for PTCy-treated patients. We assessed plasma levels using blood previously collected at serial pre-determined timepoints from 158 patients treated on one of three prospective clinical studies of PTCy as GvHD prophylaxis after myeloablative alloBMT using either HLA-matched or HLA-haploidentical donors.[4,6,34] We hypothesized that plasma elevations in these proteins might be associated with negative clinical outcomes, particularly NRM, for PTCy-treated patients.

Methods

Study design

This study was designed to assess whether plasma-derived proteins, measured at specific post-transplant timepoints, are predictive of NRM occurrence or prognostic for the development of acute or chronic GvHD after alloBMT using PTCy as GvHD prophylaxis. The sample was based on the number of available plasma specimens previously collected at pre-determined timepoints from patients treated on one of three prospective clinical studies (*clinicaltrials.gov Identifiers: 00134017, 00809276,* and *00796562*) (*Online Supplementary Figure S1*). Due to differences in the donor sources and adjunct GvHD prophylaxis, patients receiving HLA-matched *versus* HLA-haploidentical alloBMT were analyzed in separate cohorts for all analyses.

Patients and samples

All three studies exclusively employed myeloablative conditioning and T-cell-replete bone marrow allografts. Two of the clinical trials, (*clinicaltrials.gov Identifiers: 00134017* and *00809276*), were both prospective studies of HLA-matched-related or -unrelated donor alloBMT using PTCy as single-agent GvHD prophylaxis for adult patients with advanced hematologic malignancies.[4,6] One of these, (*clinicaltrials.gov Identifier: 00134017*), was a single-institutional study using busulfan/cyclophosphamide conditioning (n=122);[4] although this study spanned 2004-2009, plasma was only cryopreserved for a group of 35 patients transplanted in 2007 and 2008. The other, (*clinicaltrials.gov Identifier: 00809276*), was a multi-institutional study using busulfan/fludarabine conditioning (n=92);[6] as part of that protocol, plasma samples from all 80 patients treated at two of the three participating centers (Johns Hopkins Hospital and Fred Hutchinson Cancer Research Center) were collected. The third trial, (*clinicaltrials.gov Identifier: 00796562*), was a single-institutional prospective study of busulfan/cyclophosphamide conditioning, HLA-haploidentical donor alloBMT, and GvHD prophylaxis using PTCy, mycophenolate mofetil (MMF), and tacrolimus for adult or pediatric patients with advanced hematologic malignancies (n=95).[34] Plasma samples were collected for 92 patients. As this latter trial enrolled patients with chemorefractory hematologic malignancies and the study did not mandate specimen collection once a patient relapsed, patients on this trial that relapsed less than 6 months post-transplant (n=23) were not included in this analysis.

Blood was collected only at pre-determined timepoints. Patients who experienced graft failure (n=10), suffered NRM prior to day 30 (n=4), or survived to day 30 and had sustained engraftment but did not have a day 30 plasma sample collected (n=12) were excluded (*Online Supplementary Figure S1*). Both studies of HLA-matched alloBMT (*clinicaltrials.gov Identifiers: 00134017* and

Table 1. Patient, donor, and transplantation characteristics.

Characteristics	HLA-Matched (n=100)	HLA-Haploidentical (n=58)	All Patients (n=158)
Donor, n (%)			
Related	42 (42%)	58 (100%)	100 (63%)
Unrelated	58 (58%)	N/A	58 (37%)
Conditioning			
Bu/Cy	30 (30%)	58 (100%)	88 (56%)
Bu/Flu	70 (70%)	0 (0%)	70 (44%)
GvHD prophylaxis	PTCy alone	PTCy, MMF, tacro	
Age, y, median (range)	51 (22-65)	42 (2-64)	49 (2-65)
Sex, n (%) female	57 (57%)	27 (47%)	84 (53%)
Diagnosis, n (%)			
AML/MDS/CMML	73 (73%)	33 (57%)	106 (67%)
ALL	16 (16%)	8 (14%)	24 (15%)
NHL/CLL	3 (3%)	9 (16%)	12 (8%)
Hodgkin lymphoma	0 (0%)	4 (7%)	4 (3%)
CML	6 (6%)	4 (7%)	10 (6%)
Multiple myeloma	2 (2%)	0 (0%)	2 (1%)
Remission status at alloBMT, n (%)			
CR without MRD	52 (52%)	31 (53%)	83 (53%)
CR with MRD	23 (23%)	9 (16%)	32 (20%)
Not in CR	25 (25%)	18 (31%)	43 (27%)
HCT-CI score, median (range)	2 (0-8)	1 (0-5)	2 (0-8)
Donor age, y, median (range)	41 (17-67)	36 (11-69)	40 (11-69)
Donor sex, n (%) female	56 (56%)	34 (59%)	90 (57%)
Female donor into male recipient, n (%)	23 (23%)	18 (31%)	41 (26%)
CMV serostatus			
D-/R-	31 (31%)	28 (48%)	59 (37%)
D+/R-	16 (16%)	10 (18%)	26 (16%)
D-/R+	30 (30%)	9 (16%)	39 (25%)
D+/R+	22 (22%)	11 (19%)	33 (21%)
D^{unk}/R+	1 (1%)	0 (0%)	1 (1%)
Allograft cell counts, # (range)/kg recipient weight			
Total nucleated cells, x 10^8	4.09 (1.3-8.82)	4.84 (2.63-11.2)	4.31 (1.3-11.2)
CD34$^+$ cells, x 10^6	3.72 (1.0-9.87)	4.16 (1.68-9.73)	3.82 (1.0-9.87)
CD3$^+$ cells, x 10^7	3.8 (0.6-7.75)	4.67 (1.09-11.07)	4.11 (0.6-11.07)

HLA: human leukocyte antigen; Bu: busulfan; Cy: cyclophosphamide; Flu: fludarabine; GvHD: graft-*versus*-host disease; tacro: tacrolimus; PTCy: post-transplantation cyclophosphamide; MMF: mycophenolate mofetil; AML: acute myeloid leukemia; MDS: myelodysplastic syndrome; CMML: chronic myelomonocytic leukemia; ALL: acute lymphoblastic leukemia; NHL: non-Hodgkin lymphoma; CLL: chronic lymphocytic leukemia; CML: chronic myelogenous leukemia; alloBMT: allogeneic bone marrow transplantation; CR: morphologic complete remission; MRD: minimal residual disease; HCT-CI, hematopoietic cell transplantation-comorbidity index; CMV: cytomegalovirus; D: donor; R: recipient; unk: unknown.

00809276) were combined together as a single cohort for analysis as these patients were identically treated except for the type of myeloablative conditioning (cyclophosphamide *versus* fludarabine in addition to busulfan). Ultimately, 100 patients receiving HLA-matched alloBMT and 58 patients receiving HLA-haploidentical alloBMT were included in this study (Table 1; *Online Supplementary Figure S1*). Plasma samples from 33 healthy donors of bone marrow for patients treated on the two latter mentioned trials were also included in these analyses. All patients and healthy donors provided written informed consent on institutional review board-approved protocols prior to specimen collection.

Proteomic analysis

Enzyme-linked immunosorbent assays (ELISAs) for IL-2Rα, IL-6, TNFR-1, ST2, elafin, REG3α, and CXCL9 were performed in batches on cryopreserved plasma.[12-14,16-18,20] All of these biomarkers were measured using sequential ELISA as previously reported.[35] These methods are described in detail in the *Online Supplementary Methods*.

Statistical analysis

Statistical comparisons between donors and recipients regarding biomarker levels at various timepoints were based on the nonparametric Mann-Whitney test. In order to account for multiple hypothesis testing due to the multiple timepoints, we applied the Bonferroni type I error adjustment. All other outcomes were analyzed under the competing risks framework.[36] Definitions of endpoints and competing risks are defined in the *Online Supplementary Methods*. To explicitly quantify the predictive accuracy of each biomarker at 30 days for the different outcomes of the study, and to account for both the time-dependency of the outcomes and the competing risks, time-dependent receiver operating characteristic (ROC) curves for competing endpoints were computed.[37] Predictive accuracy was estimated based on the area under the ROC curve at 3, 6, 9, and 12 months. To assess the association of each biomarker with outcomes, the linear effect of the biomarker (treated as a continuous variable), either alone or adjusted for the potential confounders of recipient age and recipient cytomegalovirus serostatus, on the cumulative incidence was esti-

Table 2. Univariable associations of day 30 biomarker levels with acute GvHD and non-relapse mortality.

Biomarker	Grade II-IV Acute GvHD SHR (95% CI)	P	Grade III-IV Acute GvHD SHR (95% CI)	P	Non-Relapse Mortality SHR (95% CI)	P
IL-2Rα (per 500 units)						
HLA-matched	1.27 (1.06-1.53)	0.01	1.47 (1.18-1.83)	0.001	1.55 (1.17-2.05)	0.002
HLA-haploidentical	1.66 (1.04-2.64)	0.03	1.61 (1.08-2.41)	0.021	1.88 (1.28-2.76)	0.001
TNFR-1 (per 1000 units)						
HLA-matched	1.04 (0.96-1.13)	0.33	1.09 (1.00-1.19)	0.05	1.22 (1.07-1.38)	0.002
HLA-haploidentical	1.00 (0.91-1.10)	0.99	1.01 (0.88-1.16)	0.93	1.28 (1.20-1.36)	<0.001
ST2 (per 20 units)						
HLA-matched	1.04 (0.95-1.15)	0.39	1.06 (0.88-1.27)	0.56	1.12 (1.02-1.23)	0.013
HLA-haploidentical	1.10 (0.92-1.33)	0.30	1.10 (0.89-1.37)	0.36	1.41 (1.26-1.56)	<0.001
REG3α (per 20 units)						
HLA-matched	1.00 (0.98-1.02)	0.60	1.01 (0.99-1.03)	0.24	1.05 (1.04-1.07)	<0.001
HLA-haploidentical	0.99 (0.98-1.00)	0.11	0.99 (0.97-1.01)	0.17	1.01 (1.01-1.01)	<0.001
Elafin (per 1000 units)						
HLA-matched	1.02 (0.98-1.07)	0.32	1.00 (0.94-1.06)	0.98	1.00 (0.93-1.06)	0.88
HLA-haploidentical	1.06 (0.99-1.13)	0.09	1.03 (0.98-1.08)	0.31	1.11 (1.06-1.17)	<0.001
IL-6 (per 10 units)						
HLA-matched	1.00 (0.99-1.01)	0.98	1.00 (0.99-1.02)	0.69	1.00 (0.98-1.02)	0.91
HLA-haploidentical	0.97 (0.83-1.13)	0.68	0.12 (0.02-0.80)	0.03	1.03 (1.02-1.04)	<0.001
CXCL9 (per 5 units)						
HLA-matched	1.02 (0.98-1.06)	0.24	1.03 (0.99-1.07)	0.17	1.03 (0.98-1.08)	0.25
HLA-haploidentical	1.50 (0.92-2.44)	0.10	2.70 (1.80-4.04)	<0.001	1.23 (0.73-2.09)	0.44

For grade II-IV acute GvHD, the numbers of events and competing risks were 31 and 23 for the HLA-matched cohort and 6 and 19 for the HLA-haploidentical cohort, respectively. For grade III-IV acute GvHD, the numbers of events and competing risks were 16 and 39 for the HLA-matched cohort and 2 and 23 for the HLA-haploidentical cohort, respectively. For non-relapse mortality, the numbers of events and competing risks were 14 and 37 for the HLA-matched cohort and 10 and 12 for the HLA-haploidentical cohort, respectively. Biomarkers for these analyses were assessed as continuous variables. The SHR listed for each biomarker/outcome is the risk per given number of biomarker units shown. GvHD: graft-*versus*-host disease; IL-2Rα: interleukin-2 receptor α; TNFR-1: tumor necrosis factor receptor-1; ST2: serum STimulation-2, IL1RL1 gene product; REG3α: regenerating islet-derived 3-α; IL-6: interleukin-6; CXCL9: chemokine [C-X-C motif] ligand 9; HLA: human leukocyte antigen; SHR: subdistribution hazard ratio; CI: confidence interval.

mated under the semiparametric proportional subdistribution hazards model (Fine-Gray model).[38] Additionally, day 30 biomarker values were dichotomized (higher than or equal to median value *vs.* lower than median value), and cumulative incidence functions, estimated using the nonparametric Aalen-Johansen estimator,[39] were plotted. Statistical comparisons between the cumulative incidence curves for the outcome by the biomarker level categories were undertaken using Gray's nonparametric test.[40] All analyses were performed separately for the HLA-matched and HLA-haploidentical cohorts.
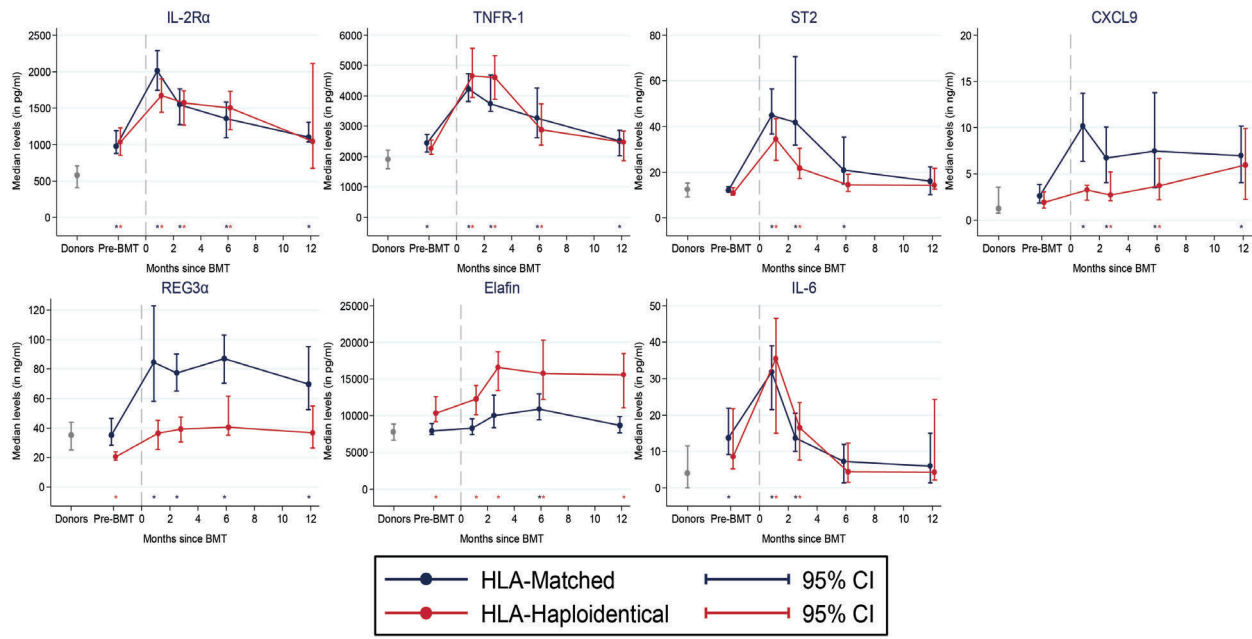
Results

Plasma-derived putative proteomic biomarkers are elevated post-transplant

Plasma levels of all seven tested proteins were significantly elevated in patients post-transplant compared with plasma levels of their healthy bone marrow donors (Figure 1). Levels of IL-2Rα, TNFR-1, ST2, and IL-6 were highest at 30 days post-transplant, declined at subsequent post-transplant measurements, and had similar temporal kinetics between the HLA-matched and HLA-haploidentical cohorts (Figure 1). Levels of the other three proteins were relatively stable during the first post-transplant year, but differed between the two cohorts: REG3α and CXCL9 levels were higher in the HLA-matched cohort, while

elafin levels were higher in the HLA-haploidentical cohort (Figure 1). The differences in REG3α levels between the two cohorts may be related to the much higher incidence of gastrointestinal acute GvHD in the HLA-matched cohort (39% *versus* 12%) (*Online Supplementary Table S1*). The difference in elafin levels between cohorts was not directly attributable to varying GvHD rates, as the incidence of cutaneous acute GvHD was higher in the HLA-matched cohort (*Online Supplementary Table S1*).

Day 30 IL-2Rα, TNFR-1, ST2, and REG3α are consistently predictive of the occurrence of non-relapse mortality

Using the first post-transplant timepoint (day 30) samples, levels of each of the seven proteins were tested for associations with subsequent occurrence of NRM. First, time-dependent ROC curves, which relate a biomarker's sensitivity and specificity, were generated to assess the overall accuracy (area under the ROC curve) of each biomarker for predicting NRM. Separate ROC curves were generated for each cohort (HLA-matched and HLA-haploidentical) (Figure 2). In both cohorts, IL-2Rα, TNFR-1, ST2, and REG3α all had high area under the curve (AUC) values of 0.74-0.97, consistent with high degrees of predictive accuracy for NRM occurrence by 3 months post-transplant (Figure 2). In the HLA-haploidentical cohort, elafin and IL-6 also had high AUC values of 0.72-0.83 (Figure 2).

Figure 1. Dynamics of biomarker levels over time. Plasma biomarker levels are shown for healthy bone marrow donors (n=33) and for patients at pre-transplant and serial pre-defined (1, 2, 6, and 12 months) post-transplant timepoints. The number of samples for the HLA-haploidentical cohort is as follows: pre-transplant n=53, 1 month n=58, 2 months n=54, 6 months n=41, 12 months n=18. The number of samples for the HLA-matched cohort is as follows: pre-transplant n=93, 1 month n=100, 2 months n=90, 6 months n=33, 12 months n=36. Samples for the HLA-matched and HLA-haploidentical cohorts were drawn at the same timepoints, but are shown slightly staggered for the sake of clarity. Asterisks show statistically significant differences for each cohort compared with donor levels. IL-2Rα: interleukin-2 receptor α; TNFR-1: tumor necrosis factor receptor 1; ST2: serum STimulation-2, IL1RL1 gene product; CXCL9: chemokine [C-X-C motif] ligand 9; REG3α: regenerating islet-derived 3-α; IL-6: interleukin-6; BMT: allogeneic bone marrow transplantation; HLA: human leukocyte antigen; CI: confidence interval.

In both cohorts, the AUC values remained high for each of these biomarkers when assessing NRM occurring by 6, 9, or 12 months post-transplant, suggesting that early elevations in these biomarkers may be predictive of both shorter and longer term NRM risk.

In univariable analyses for association with NRM, high day 30 levels of IL-2Rα, TNFR-1, ST2, and REG3α were significantly associated with greater cumulative incidence of NRM within each cohort (Table 2). High elafin and IL-6 levels were also associated with NRM in the HLA-haploidentical cohort only (Table 2), while CXCL9 was not associated with NRM in either cohort. As prior analyses had shown significant associations of recipient age and recipient cytomegalovirus serostatus with NRM,[5] biomarker analyses were subsequently adjusted for these two factors and showed similar results as the unadjusted univariable analyses (*Online Supplementary Table S2*).

To provide results comparable with one of the most analogous alloBMT biomarker studies,[20] the association of each protein with NRM was also assessed for each cohort by dichotomizing each biomarker at the median and non-parametrically estimating the cumulative incidence functions of each outcome (*Online Supplementary Tables S3 and S4*). Of note, in all univariable analyses, there were no significant associations between day 30 levels for any of the seven proteins and subsequent relapse of malignant disease.

Day 30 IL-2Rα levels are consistently prognostic for the occurrence of acute GvHD

The timing of sample collection was not ideal for assess-

ing the potential association of these biomarkers with acute GvHD as 31 patients (19.6%) had onset of grade II-IV acute GvHD prior to day 30. Despite this limitation, and in order to better understand the associations with subsequent NRM, we did assess whether these biomarkers were prognostic for acute GvHD occurring after day 30. Importantly, the NRM for patients with grade II-IV acute GvHD onset prior to day 30 was not higher than the NRM of patients who developed grade II-IV acute GvHD after day 30 (time-dependent proportional cause-specific hazard ratio (HR), grade II-IV acute GvHD developing after day 30 compared with before day 30, HR 1.14 (95% confidence interval (CI) 0.41-3.16); *P*=0.81). Furthermore, there was no evidence that the risk of developing grade III-IV acute GvHD was different in those who developed grade II-IV acute GvHD after day 30 compared with before day 30 based on Fisher's exact test (*P*=0.68) or the Cox proportional cause-specific hazards model (HR 1.3 (95% CI 0.5-3.35), *P*=0.59).

In testing associations with grade II-IV acute GvHD using the day 30 protein levels, ROC curves showed high AUC values of >0.7 in both cohorts for IL-2Rα, TNFR-1, and CXCL9, consistent with high predictive accuracy for grade II-IV acute GvHD occurring after day 30 (*Online Supplementary Figures S2 and S3*). In the HLA-haploidentical cohort, ST2 and REG3α also had high AUC values of approximately 0.9 for grade II-IV acute GvHD, although the number of events (n=6) was low (*Online Supplementary Figure S3*). In testing associations with grade III-IV acute GvHD development after day 30, day 30 IL-2Rα, CXCL9, and REG3α levels all had high AUC values (>0.73 for the
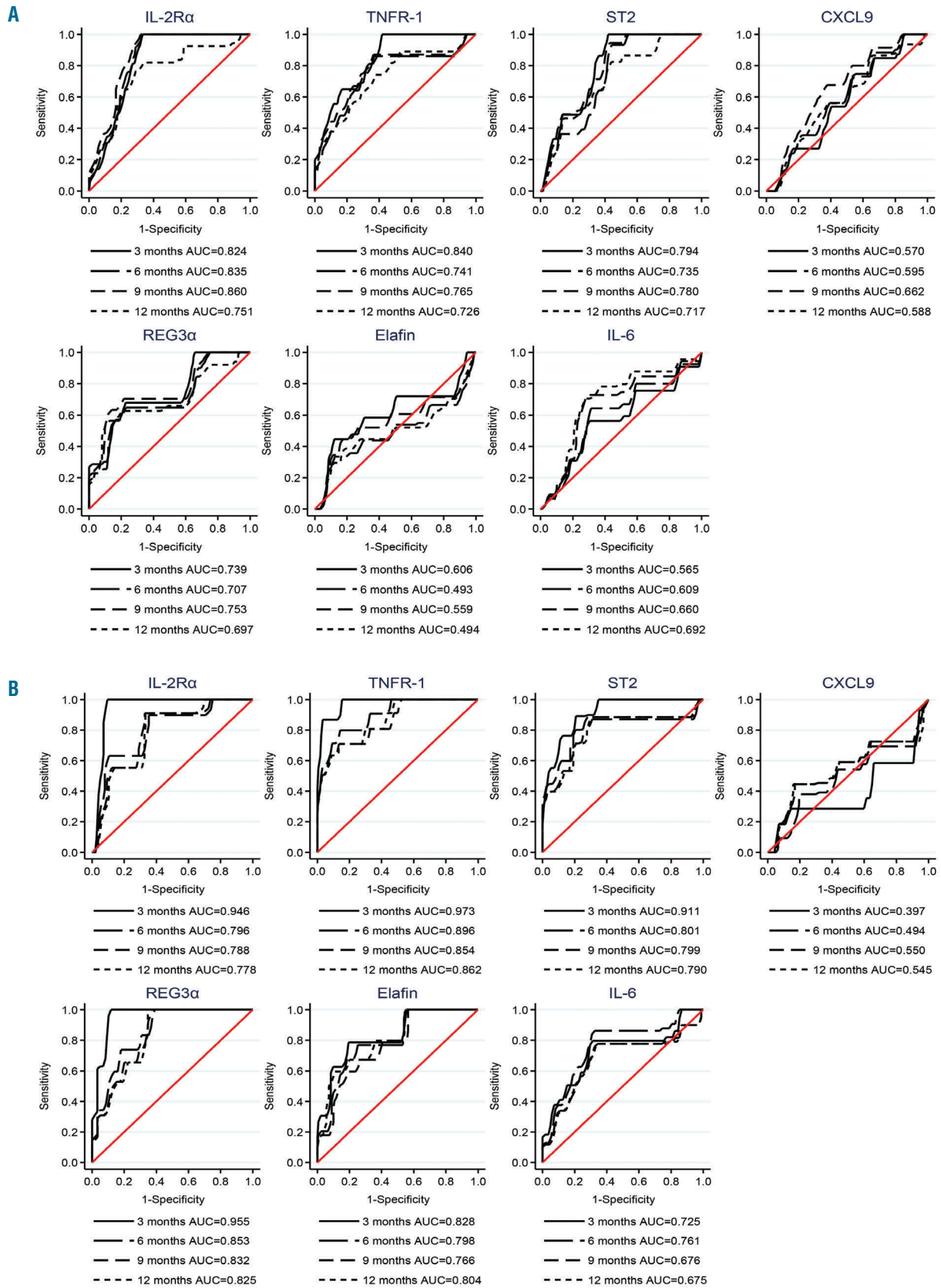
Figure 2. Time-dependent ROC curves for NRM separated by donor type. Using day 30 biomarker levels, ROC curves for predicting subsequent NRM occurrence by 3, 6, 9, or 12 months post-transplant are shown separately for the (A) HLA-matched and (B) HLA-haploidentical cohorts. The analysis was performed using a competing risks framework with relapse as a competing risk for NRM. IL-2Rα: interleukin-2 receptor α; TNFR-1: tumor necrosis factor receptor 1; ST2: serum STimulation-2, IL1RL1 gene product; CXCL9: chemokine [C-X-C motif] ligand 9; REG3α: regenerating islet-derived 3-α; IL-6: interleukin-6; AUC: area under the curve.

HLA-matched cohort and >0.9 for the HLA-haploidentical cohort) (*Online Supplementary Figures S4* and *S5*). Within the HLA-haploidentical cohort, day 30 TNFR-1, ST2, and elafin levels also had high AUC values (>0.8) for grade III-IV acute GvHD, while IL-6 had a low AUC value (0.18), although again the event number was quite low (n=2). In univariable analyses, high day 30 IL-2Rα levels were consistently associated with a greater cumulative incidence of both grade II-IV and grade III-IV acute GvHD (Table 2 and *Online Supplementary Table S2; Online Supplementary Figures S2-S5*).

Six-month plasma protein levels were used to evaluate for an association with chronic GvHD development. However, only 71 of the 125 patients (56.8%) that were chronic GvHD-free and still at risk for chronic GvHD at 6 months had 6-month samples collected. Univariable analyses with chronic GvHD development did not show any statistically significant associations, although CXCL9 trended towards significance in the HLA-haploidentical cohort (HLA-haploidentical cohort, HR 1.49 (95% CI 0.93-2.40), *P*=0.096; HLA-matched cohort, HR 1.12 (95% CI 0.95-1.32), *P*=0.17).

Discussion

In the study herein, we explored whether putative proteomic biomarkers, previously assessed using other transplantation platforms, were applicable to PTCy-treated patients. Among its novel features, our study is the first to examine the relevance of promising biomarkers in PTCy-treated patients or patients receiving HLA-haploidentical alloBMT. All patients were uniformly treated with myeloablative conditioning, bone marrow allografts, and PTCy for GvHD prophylaxis, and we found effects that appear to be conserved across different donor sources. Furthermore, we have used novel statistical methods to address our scientific aims. These include the Fine-Gray model to directly estimate the effect of an independent variable on the cumulative incidence function of the outcome of interest in the presence of competing risks and also time-dependent ROC curve methodology for competing risks to directly quantify the predictive accuracy of each biomarker.

We found that, despite low rates of NRM and severe GvHD in PTCy-treated patients, elevated levels of several of these plasma proteins remained prognostic for adverse outcomes. Consistently in both cohorts, elevations in IL-2Rα, TNFR-1, ST2, and REG3α at 30 days post-transplant each were significantly associated with a greater cumulative incidence of NRM. Furthermore, in both cohorts, elevations in IL-2Rα at 30 days post-transplant were associated with subsequent grade II-IV and III-IV acute GvHD occurrence.

Prior studies of putative proteomic biomarkers in other transplantation platforms, using plasma measurements at pre-determined timepoints during the first post-transplant month, have shown associations between levels of TNFR-1, ST2, and REG3α and subsequent occurrence of NRM and/or acute GvHD.[10,17,20,23] Most consistent with our results, a recent large study of patients treated with HLA-matched-related alloBMT showed that elevated day 28 ST2 and REG3α levels were associated with greater risk for NRM, but did not have a significant relationship with acute or chronic GvHD.[23] Also compatible with our

results, a study of patients undergoing dUCBT found that elevations in TNFR-1, ST2, and REG3α at day 28 post-transplant were associated with NRM at 180 days.[20] In contrast to their results, but consistent with the HLA-matched-related alloBMT study, we did not find a statistically significant association between ST2 levels and acute GvHD. Furthermore, in the dUCBT study, elevations above the median in IL-2Rα were not associated with GvHD or NRM, whereas our study showed strong associations of IL-2Rα with grade II-IV acute GvHD, grade III-IV acute GvHD, and NRM. The strong relationship we found between day 30 IL-2Rα levels and acute GvHD was consistent with another study that demonstrated a peak in IL-2Rα levels at 14 days post-transplant in patients who would subsequently develop acute GvHD.[7] Moreover, our results are consistent with other studies which have shown associations between IL-2Rα levels measured at the start of acute GvHD and subsequent NRM.[12,16] In contrast with prior studies that showed that day 14 IL-6 levels were prognostic of subsequent grade III-IV acute GvHD occurrence or that day 28 IL-6 levels were prognostic of NRM,[21,23] we did not find a consistent relationship between IL-6 and outcomes between our two cohorts.

Our study has several limitations. Importantly, we do not have an independent verification cohort as we utilized all plasma samples available from three prospective studies of PTCy. Furthermore, cross-validation was not performed to evaluate the out-of-sample predictive ability of our model. This is because the goal of the present study was neither to propose a specific model for the alloBMT-related outcomes nor to provide a specific classification/prediction rule. Our goal was to evaluate the predictive potential of a set of biomarkers, and therefore we quantified this using time-dependent ROC curves for competing risks. Even so, we found consistent results between the two cohorts in terms of four predictors of NRM and one predictor of acute GvHD, suggesting that our results may have some external validity. Differences in plasma protein levels and associations for three of the tested biomarkers (CXCL9, REG3α, and elafin) between the HLA-matched and HLA-haploidentical cohorts could reflect differences in the biology of donor HLA disparity, but more likely reflect the differences in the immunosuppressants used (the former used PTCy alone, whereas the latter also incorporated MMF and tacrolimus), as demonstrated in a recent publication regarding REG3α,[23] in addition to the differing incidences of acute GvHD (lower grade II-IV and III-IV acute GvHD rates for the HLA-haploidentical cohort) probably due to this adjunct immunosuppression therapy. Unfortunately, our limited number of events for NRM and GvHD outcomes made it statistically infeasible to perform multivariable analyses to attempt to thoroughly dissect the impact of other potential confounders. Another limitation of this study is that we did not adjust for type I error inflation due to multiple statistical testing in the univariable models. However, the main purpose of this work was to estimate the predictive accuracy for a set of biomarkers that had already been reported to be associated with NRM and GvHD in previous reports using other transplantation platforms. The evaluation of this predictive accuracy was not based on hypothesis testing but on the calculation of ROC AUCs. Furthermore, hypothesis tests that were conducted were specified *a priori* and were confirmatory in nature rather than exploratory.

Lastly, our study is limited by the sample collection performed. Our samples were collected at uniform pre-determined timepoints, but samples from patient-specific timepoints (e.g., at GvHD diagnosis or longitudinally in response to GvHD treatment) were not collected. Furthermore, the first post-transplant plasma samples were taken at day 30; 20% of patients had experienced the onset of grade II-IV acute GvHD prior to that timepoint and had to be excluded from the acute GvHD analyses. Thus, our results for acute GvHD are only relevant for patients surviving to day 30 without yet developing acute GvHD. The apparent lack of statistical association of some of these proteins (e.g., ST2, REG3α, or elafin for acute GvHD) with clinical outcomes in our patients could in part be explained by these temporal and contextual differences in sampling; had samples been collected earlier post-transplant or at the time of diagnosis of acute GvHD, these proteins may have had a more apparent prognostic import. While the HLA-haploidentical cohort also suffered from exclusion of patients who relapsed within 6 months post-transplant, those patients would not have remained at risk for NRM or GvHD in any case due to the occurrence of a competing risk. Thus, the main findings of this manuscript would likely remain unchanged even if samples had been universally available for those patients. Moreover, many of the results in the HLA-haploidentical cohort were confirmed in the HLA-matched cohort, in which there was no bias in sample availability.

Even though PTCy is broadly and highly effective in preventing chronic GvHD in patients undergoing alloBMT,[1,2,4-6] our study has not revealed any factor that can prognosticate which of our patients will still develop chronic GvHD despite PTCy treatment. It is possible that the lack of any significant findings was due in part to the low incidence of chronic GvHD in our patients and thus low statistical power to detect true differences that may exist; however, the number of chronic GvHD events (n=24) was identical to the number of NRM events (n=24) and more than the number of grade III-IV acute GvHD events (n=18). Further insight may require a better understanding of the underlying biology of chronic GvHD, which remains poorly defined.[41] Immunologic parameters could potentially be of interest given the recently demonstrated role for regulatory T cells in GvHD prevention by PTCy,[42,43] but such studies would require a large number of patients for adequate statistical power given the low incidence of chronic GvHD development in PTCy-treated patients.

Although not prognostic for chronic GvHD, elevations in IL-2Rα, TNFR-1, ST2, and REG3α at day 30 post-transplant were consistently predictive of NRM occurrence in our patients. While IL-2Rα was also prognostic for acute GvHD occurrence in both cohorts, the majority of our patients did not die from GvHD-related causes (*Online*

Supplementary Table S5). This apparent disconnect between predictors of GvHD and NRM was also seen in a recent study of HLA-matched alloBMT in which day 28 plasma biomarker assessments were studied.[23] Thus, whether elevations in these proteins are specific for GvHD is unclear. It is possible that GvHD and many cases of NRM are both direct sequelae of a dysregulated immune system, which is what many of these proteins are directly measuring. Further dissection of what elevations in each of these biomarkers mean, particularly in cases where they are not correlated with each other, would require complex prospective assessment of a host of clinical and immunologic factors and would require a large number of patients in addition to pre-clinical supporting studies.

Beyond the immediate risk for NRM, it is disconcerting that day 30 elevations of these proteins are associated with an increased risk for NRM in our patients, which remains present for at least the first post-transplant year. Given this apparent longer term risk, a fundamental outstanding question is whether this risk is modifiable with more intensive monitoring, immunosuppressive treatment, or other novel interventions. Conversely, perhaps low levels of these plasma proteins could justify less intensive monitoring in such patients (i.e., risk-adapted therapy). If the risk is modifiable, what is the earliest post-transplant timepoint at which this risk could be determined such that an intervention could be expeditiously implemented? Prospective studies are required to further clarify the role of biomarkers in facilitating personalized, risk-adapted strategies for monitoring patients after alloBMT. This may be particularly relevant as older and less fit patients are being transplanted.[44,45] Moreover, the optimal cutoffs of "high" *versus* "low" would need to be better defined and standardized before broader clinical application of these results. Despite the need for further research, our work supports the fact that subsequent studies, such as the BMT Clinical Trials Network study 1202,[46] hold high promise for resulting in validated plasma-derived biomarkers that could have prognostic and predictive utility broadly applicable to patients undergoing alloBMT, irrespective of the particular transplantation platform employed.

References

1. Kanakry CG, Fuchs EJ, Luznik L. Modern approaches to HLA-haploidentical blood or marrow transplantation. Nat Rev Clin Oncol. 2016;13(2):132.

2. Luznik L, O'Donnell PV, Symons HJ, et al. HLA-haploidentical bone marrow transplantation for hematologic malignancies using nonmyeloablative conditioning and high-dose, posttransplantation cyclophosphamide. Biol Blood Marrow Transplant. 2008;14(6):641-650.

3. McCurdy SR, Kanakry JA, Showel MM, et al. Risk-stratified outcomes of nonmyeloablative HLA-haploidentical BMT with high-dose posttransplantation cyclophosphamide. Blood. 2015;125(19):3024-3031.

4. Luznik L, Bolanos-Meade J, Zahurak M, et al. High-dose cyclophosphamide as single-

agent, short-course prophylaxis of graft-versus-host disease. Blood. 2010; 115(16):3224-3230.

5. Kanakry CG, Tsai HL, Bolanos-Meade J, et al. Single-agent GVHD prophylaxis with posttransplantation cyclophosphamide after myeloablative, HLA-matched BMT for AML, ALL, and MDS. Blood. 2014; 124(25):3817-3827.

6. Kanakry CG, O'Donnell PV, Furlong T, et al. Multi-institutional study of post-transplantation cyclophosphamide as single-agent graft-versus-host disease prophylaxis after allogeneic bone marrow transplantation using myeloablative busulfan and fludarabine conditioning. J Clin Oncol. 2014; 32(31):3497-3505.

7. Miyamoto T, Akashi K, Hayashi S, et al. Serum concentration of the soluble interleukin-2 receptor for monitoring acute graft-versus-host disease. Bone Marrow Transplant. 1996;17(2):185-190.

8. Foley R, Couban S, Walker I, et al. Monitoring soluble interleukin-2 receptor levels in related and unrelated donor allogenic bone marrow transplantation. Bone Marrow Transplant. 1998;21(8):769-773.

9. Grimm J, Zeller W, Zander AR. Soluble interleukin-2 receptor serum levels after allogeneic bone marrow transplantations as a marker for GVHD. Bone Marrow Transplant. 1998;21(1):29-32.

10. Kitko CL, Paczesny S, Yanik G, et al. Plasma elevations of tumor necrosis factor-receptor-1 at day 7 postallogeneic transplant correlate with graft-versus-host disease severity and overall survival in pediatric patients. Biol Blood Marrow Transplant. 2008;14(7):759-765.

11. Choi SW, Kitko CL, Braun T, et al. Change in plasma tumor necrosis factor receptor 1 levels in the first week after myeloablative allogeneic transplantation correlates with severity and incidence of GVHD and survival. Blood. 2008;112(4):1539-1542.

12. Paczesny S, Krijanovski OI, Braun TM, et al. A biomarker panel for acute graft-versus-host disease. Blood. 2009;113(2):273-278.

13. Paczesny S, Braun TM, Levine JE, et al. Elafin is a biomarker of graft-versus-host disease of the skin. Sci Transl Med. 2010; 2(13):13ra12.

14. Ferrara JL, Harris AC, Greenson JK, et al. Regenerating islet-derived 3-alpha is a biomarker of gastrointestinal graft-versus-host disease. Blood. 2011;118(25):6702-6708.

15. Harris AC, Ferrara JL, Braun TM, et al. Plasma biomarkers of lower gastrointestinal and liver acute GVHD. Blood. 2012; 119(12):2960-2963.

16. 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.

17. 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.

18. Kitko CL, Levine JE, Storer BE, et al. Plasma CXCL9 elevations correlate with chronic GVHD diagnosis. Blood. 2014;123(5):786-793.

19. Holtan SG, Verneris MR, Schultz KR, et al. Circulating angiogenic factors associated with response and survival in patients with acute graft-versus-host disease: results from Blood and Marrow Transplant Clinical Trials Network 0302 and 0802. Biol Blood Marrow Transplant. 2015; 21(6):1029-1036.

20. 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.

21. 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.

22. Holler E, Kolb HJ, Moller A, et al. Increased serum levels of tumor necrosis factor alpha precede major complications of bone marrow transplantation. Blood. 1990;75(4):1011-1016.

23. Abu Zaid M, Wu J, Wu C, et al. Plasma biomarkers of risk for death in a multicenter phase 3 trial with uniform transplant characteristics post-allogeneic HCT. Blood. 2016;129(2):162-170.

24. Rubin LA, Kurman CC, Fritz ME, et al. Soluble interleukin 2 receptors are released from activated human lymphoid cells in vitro. J Immunol. 1985;135(5):3172-3177.

25. Coghill JM, Sarantopoulos S, Moran TP, Murphy WJ, Blazar BR, Serody JS. Effector CD4+ T cells, the cytokines they generate, and GVHD: something old and something new. Blood. 2011;117(12):3268-3276.

26. Holler E, Kolb HJ, Hintermeier-Knabe R, et al. Role of tumor necrosis factor alpha in acute graft-versus-host disease and complications following allogeneic bone marrow transplantation. Transplant Proc. 1993;25(1 Pt 2):1234-1236.

27. Matta BM, Lott JM, Mathews LR, et al. IL-33 is an unconventional Alarmin that stimulates IL-2 secretion by dendritic cells to selectively expand IL-33R/ST2+ regulatory T cells. J Immunol. 2014;193(8):4010-4020.

28. Schiering C, Krausgruber T, Chomka A, et al. The alarmin IL-33 promotes regulatory T-cell function in the intestine. Nature. 2014;513(7519):564-568.

29. Groom JR, Luster AD. CXCR3 ligands: redundant, collaborative and antagonistic functions. Immunol Cell Biol. 2011; 89(2):207-215.

30. Cash HL, Whitham CV, Behrendt CL, Hooper LV. Symbiotic bacteria direct expression of an intestinal bactericidal lectin. Science. 2006;313(5790):1126-1130.

31. Zheng Y, Valdez PA, Danilenko DM, et al. Interleukin-22 mediates early host defense against attaching and effacing bacterial pathogens. Nat Med. 2008;14(3):282-289.

32. Ogawa H, Fukushima K, Naito H, et al. Increased expression of HIP/PAP and regenerating gene III in human inflammatory bowel disease and a murine bacterial reconstitution model. Inflamm Bowel Dis. 2003; 9(3):162-170.

33. Alkemade JA, Molhuizen HO, Ponec M, et al. SKALP/elafin is an inducible proteinase inhibitor in human epidermal keratinocytes. J Cell Sci. 1994;107(Pt 8):2335-2342.

34. Symons HJ, Chen A, Gamper C, et al. Haploidentical BMT using fully myeloablative conditioning, T cell replete bone marrow grafts, and post-transplant cyclophosphamide (PT/Cy) has limited toxicity and promising efficacy in largest reported experience with high risk hematologic malignancies. Biol Blood Marrow Transplant. 2015;21(2):S29-S29.

35. 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. pii 4247.

36. Bakoyannis G, Touloumi G. Practical methods for competing risks data: a review. Stat Methods Med Res. 2012;21(3):257-272.

37. Saha P, Heagerty PJ. Time-dependent predictive accuracy in the presence of competing risks. Biometrics. 2010;66(4):999-1011.

38. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. Stat Med. 1999;94(446):496-509.

39. Aalen OO, Johansen S. An empirical transition matrix for non-homogeneous Markov chains based on censored observations. Scand J Stat. 1978;5:141-150.

40. Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. Ann Stat. 1988;16:1141-1154.

41. Blazar BR, Murphy WJ, Abedi M. Advances in graft-versus-host disease biology and therapy. Nat Rev Immunol. 2012;12(6):443-458.

42. Kanakry CG, Ganguly S, Zahurak M, et al. Aldehyde dehydrogenase expression drives human regulatory T cell resistance to posttransplantation cyclophosphamide. Sci Transl Med. 2013;5(211):211ra157.

43. Ganguly S, Ross DB, Panoskaltsis-Mortari A, et al. Donor CD4+ Foxp3+ regulatory T cells are necessary for posttransplantation cyclophosphamide-mediated protection against GVHD in mice. Blood. 2014; 124(13):2131-2141.

44. Kasamon YL, Bolanos-Meade J, Prince GT, et al. Outcomes of nonmyeloablative HLA-haploidentical blood or marrow transplantation with high-dose post-transplantation cyclophosphamide in older adults. J Clin Oncol. 2015; 33(28):3152-3161.

45. McClune BL, Ahn KW, Wang HL, et al. Allotransplantation for patients age >/=40 years with non-Hodgkin lymphoma: encouraging progression-free survival. Biol Blood Marrow Transplant. 2014;20(7):960-968.

46. US National Library of Medicine. ClinicalTrials.gov [online]; 2013. Available from: https: // clinicaltrials.gov/ct2/show/NCT01879072. Last Accessed: 21st December 2016.