Graft-Versus-Host Disease–Free Antitumoral Signature After Allogeneic Donor Lymphocyte Injection Identified by Proteomics and Systems Biology

Xiaowen Liu, PhD[1,2]; Zongliang Yue, MS[3]; Yimou Cao, MS[1]; Lauren Taylor, MD[2]; Qing Zhang, MS[4]; Sung W. Choi, MD[5]; Samir Hanash, MD, PhD[6]; Sawa Ito, MD[7]; Jake Y. Chen, PhD[3]; Huanmei Wu, PhD[1]; and Sophie Paczesny, MD, PhD[2]

PURPOSE As a tumor immunotherapy, allogeneic hematopoietic cell transplantation with subsequent donor lymphocyte injection (DLI) aims to induce the graft-versus-tumor (GVT) effect but often also leads to acute graft-versus-host disease (GVHD). Plasma tests that can predict the likelihood of GVT without GVHD are still needed.

PATIENTS AND METHODS We first used an intact-protein analysis system to profile the plasma proteome post-DLI of patients who experienced GVT and acute GVHD for comparison with the proteome of patients who experienced GVT without GVHD in a training set. Our novel six-step systems biology analysis involved removing common proteins and GVHD-specific proteins, creating a protein-protein interaction network, calculating relevance and penalty scores, and visualizing candidate biomarkers in gene networks. We then performed a second proteomics experiment in a validation set of patients who experienced GVT without acute GVHD after DLI for comparison with the proteome of patients before DLI. We next combined the two experiments to define a biologically relevant signature of GVT without GVHD. An independent experiment with single-cell profiling in tumor antigen–activated T cells from a patient with post–hematopoietic cell transplantation relapse was performed.

RESULTS The approach provided a list of 46 proteins in the training set, and 30 proteins in the validation set were associated with GVT without GVHD. The combination of the two experiments defined a unique 61-protein signature of GVT without GVHD. Finally, the single-cell profiling in activated T cells found 43 of the 61 genes. Novel markers, such as RPL23, ILF2, CD58, and CRTAM, were identified and could be extended to other antitumoral responses.

CONCLUSION Our multiomic analysis provides, to our knowledge, the first human plasma signature for GVT without GVHD. Risk stratification on the basis of this signature would allow for customized treatment plans.

INTRODUCTION

Allogeneic hematopoietic cell transplantation (HCT) is one of the most effective forms of tumor immunotherapy available to date. The lymphocytes in the donor graft recognize and eliminate residual tumoral cells through the graft-versus-tumor (GVT) effect, and thus, donor lymphocyte injection (DLI) often is used at the time of relapse post-HCT to induce GVT. However, GVT can occur in parallel with lymphocyte reactivity to normal host tissues, which gives rise to graft-versus-host disease (GVHD). Despite the correlation of GVT and GVHD, indirect evidence for a GVT reaction separate from GVHD has been reported in large cohorts of HCT patients or when DLI is administered to induce remission in HCT patients who have experienced a relapse.[1,2] GVT activity can be increased by targeted therapy, as has been shown with sorafenib in FLT3 internal tandem duplication–mutant leukemia cells.[3]

The GVT effect is mediated by minor histocompatibility antigens (miHAgs) on recipient leukemic cells that are recognized by donor CD4[+] and CD8[+] T cells.[4] miHAgs can mediate GVT without inducing GVHD if they are only expressed by the recipient hematopoietic cells (ie, minor H antigen [HA]-1, HA-2, BCL2A1, and HB-1).[5,6] Nonhistocompatibility antigen proteins expressed by tumor cells, called tumor-associated antigens (survivin, Wilms tumor 1, proteinase 3) also can mediate GVT activity.[7] Cytokines (ie, interleukin-15 [IL-15])[3]; checkpoint proteins, such as PD-1/PD-L1[8]; and T-cell trafficking modulators[9] are other possible mediators of GVT.

Our study aim was to develop a proteomic signature to identify GVT without GVHD after allogeneic DLI. We used an intact-protein analysis system (IPAS) coupled with protein tagging as previously reported by us[10] first in a training set and then in a validation set and added a novel systems biology pipeline to identify a signature

CONTEXT SUMMARY

Key Objective

Separating graft-versus-host disease (GVHD) from graft-versus-tumor (GVT) effect after hematopoietic cell transplantation (HCT) has proven difficult. We attempted to do so through plasma proteomics and systems biology analyses of patients in relapse after HCT who received donor lymphocytes as immunotherapy.

Knowledge Generated

Our novel six-step systems biology analysis involved the removal of common proteins and GVHD-specific proteins, creation of a protein-protein interaction network, calculation of relevance and penalty scores, and visualization of candidate biomarkers in gene networks to define a unique, biologically relevant 61-protein signature of GVHD-free GVT. Forty-three of the 61 genes also were found in an independent experiment using massive single-cell profiling of tumor antigen–activated T cells from a patient who experienced post-HCT relapse.

Relevance

This multiomic analysis provides the first human plasma signature for GVHD-free GVT to our knowledge. Novel, biologically relevant markers were identified and could be extended to other antitumoral responses. Risk stratification on the basis of the GVT without GVHD protein signature would allow for customized treatment plans after HCT.

of 61 proteins that are significantly expressed in the plasma of HCT patients who received DLI for tumor relapse. Furthermore, 43 (70%) of the 61 genes were found in tumor antigen (PRAME)–activated T cells from a patient with post-HCT relapse.

PATIENTS AND METHODS

All patients or their legal guardians provided written informed consent to participate in this study, and the collection of samples for studying post-HCT complications was approved by the institutional review board of the University of Michigan. The methodology for this study is described in the Data Supplement.

RESULTS

#### Plasma Proteomics of GVHD-Free GVT in a Training Set

Our initial approach to identifying GVT-specific biomarkers was to undertake a proteomics analysis. We hypothesized that proteomes of GVT are distinct from those of GVHD. We reasoned that the best setting to observe a GVT effect without any effect of chemotherapy or preparative conditioning regimen would be after DLI is given for relapse of malignant disease. Our first proteomics experiment (IPAS-01) was designed to distinguish proteomes that predict GVHD-free GVT from those that predict GVT with GVHD after DLI in a training set. We compared a pool of plasma samples taken approximately 30 days after DLI from five patients without relapse and without GVHD (GVT-positive, GVHD-negative) labeled with the heavy isotope and compared them with a pool of plasma samples at matched time points from 11 patients without relapse but with concomitant GVHD (GVT-positive, GVHD-positive) labeled with the light isotope (Fig 1). In this design, only proteins with the heavy/light isotope ratio greater than 1.2 (upregulated) will be GVHD free and considered for additional analyses. The patient and DLI characteristics are listed in the Data

Supplement. No significant differences between groups were observed for patient age, disease status (all but one with morphologic relapse, and only one patient did not receive chemotherapy), DLI donor type, and DLI dose, and all patients were off immunosuppression. All patients were evaluated for remission using morphologic evaluation.

A total of 825 proteins were confidently identified and quantified using the SwissProt database (Data Supplement). Of these proteins, 218 were upregulated with a heavy/light ratio greater than 1.2 in GVT-positive, GVHD-negative samples (and either not identified in GVT-positive, GVHD-positive samples or identified with a ratio 1.2 or less). Similar results were obtained using the UniProt human proteome database (Data Supplement). A total of 755 proteins were identified, including 327 that were upregulated by more than 1.2-fold in GVT-positive, GVHD-negative samples. The threshold of 1.2-fold also was used in our most recent GVHD clinical proteomics, which has allowed the discovery of biologically relevant biomarkers.[11]

Because the ratio-based approach is imperfect and can still capture some GVHD proteins, we performed additional steps in the analysis to remove common proteins and GVHD-specific proteins as well as to enrich for biologically relevant proteins with our novel systems biology analysis briefly described in the next sentence and in detail in the Data Supplement. Our computational biology approach involved six processing steps that were modified from our previous workflow[12,13] to fit the GVT experiment and includes generating abundance ratios from comparisons of plasma samples taken from GVT without GVHD, filtering out GVHD-specific proteins identified in previous experiments, generating protein-protein interaction (PPI) pairs with and without outer gene networks, selecting GVT-relevant proteins by relevance scores and penalty scores for unspecific proteins found in other diseases, and using a visual analytic

FIG 1. Pool selection and top-down tandem mass spectrometry (MS/MS) workflow to keep intact proteins. The intact-protein analysis system (IPAS) compared graft-versus-tumor (GVT)–positive and graft-versus-host disease (GVHD)–negative post–donor lymphocyte injection (DLI; heavy isotope) with GVT-positive and GVHD-positive post-DLI (light isotope) samples in a training set (IPAS-01), and GVT-positive and GVHD-negative post-DLI (heavy isotope) with GVT-positive and GVHD-negative pre-DLI (light isotope) samples in a validation set (IPAS-02).
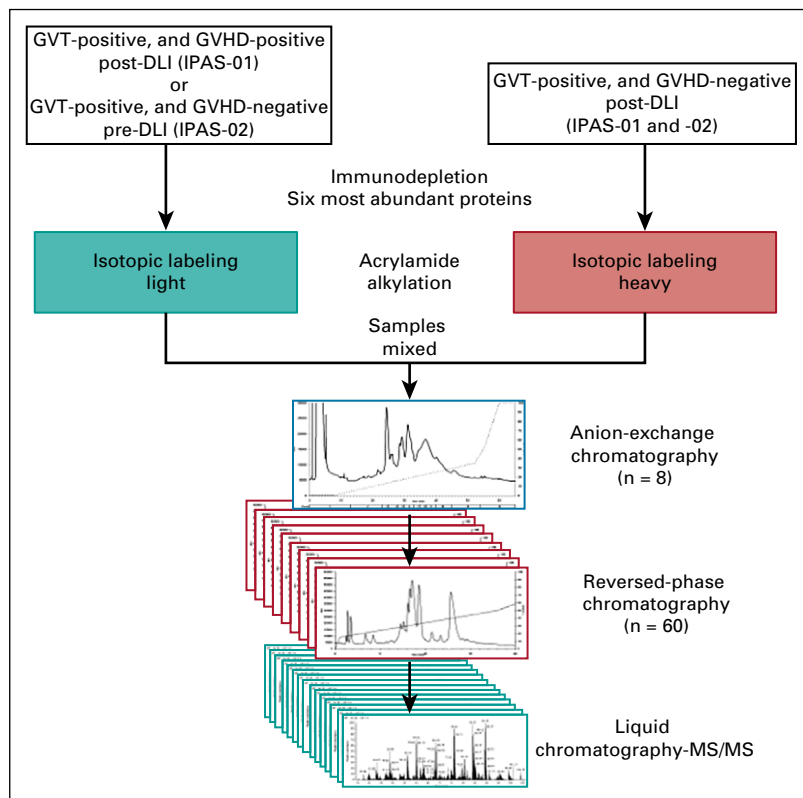
software tool (Data Supplement). After this data processing, the IPAS-01 experiment provided a list of 46 proteins associated with a GVT signature without GVHD post-DLI in the training set (Table 1). Although we started with IPAS proteins with a heavy/light ratio greater than 1.2, the systems biology process may have resulted in a final score of less than 1; however, because all these proteins are GVT specific, they all are included in the final GVHD-free GVT signature. By using the integrated relevant interaction network as the base layout and differentially expressed proteins in the candidate protein list as response variables, we constructed a GeneTerrain (www.terrainatlas.com) visualization as shown in Figure 2A. An overview of the GeneTerrain object representation is shown in the Data Supplement. Briefly, the terrain base is a nodes-weighted graph that incorporates PPI networks. The z-axis adds data obtained from the proteomics experiments to generate a 2-dimensional terrain heat map. The first node is centered around CD8-α (CD8A) and includes CD58, IL1A, Ly75, FAS, and GPNMB and a secondary node centered around RAF1 and containing GUK1, EPHB4, and CXCL12. For example, CD8A is expressed on the surface of most cytotoxic CD8+ T lymphocytes and is a coreceptor of the T-cell receptor. The relative weight of CD4+ and CD8+ T cells in GVHD and GVT responses remains undetermined. Our data with the CD8A-centered node strongly suggest that GVT is majorly mediated by CD8 T cells and GVHD less so, which provides an opportunity to separate GVT and GVHD by manipulating these two T-cell subsets differently. A second cluster comprises MAP3K6, DBNL, and FERMT3, which are all implicated in immune response pathway activity. Penalty scores are designed to enrich for the most relevant proteins without altering the primary data. As a proof of concept, we performed the experiment without applying these proteins to show that their final scores have not been significantly changed by the use of penalty scores (Table 1; Data Supplement).

Plasma Proteomics of GVHD-Free GVT in a Validation Set

Our second proteomics experiment (IPAS-02) used a validation set that was designed to compare a pool of samples from five GVT-positive, GVHD-negative (labeled with the heavy isotope) patients with a pool of samples from the same patients taken before DLI when the patients were in relapse and thus GVT negative (labeled with the light isotope; Fig 1). A total of 733 proteins were identified and quantified using the SwissProt database (Data Supplement). Of these proteins, 125 were upregulated by more than 1.2-fold in GVT-positive, GVHD-negative post-DLI samples (and either not identified in GVT-positive, GVHD-negative pre-DLI samples or identified with a ratio of 1.2 or less). Similar results were obtained using the UniProt human proteome database (Data Supplement). A total of 602 proteins were identified, including 274 proteins that were upregulated by more than 1.2-fold in GVT-positive, GVHD-negative samples.

We then applied the same systems biology workflow as in the training set and obtained 30 proteins associated with

TABLE 1. Sixty-One Proteins in the Signature of GVT Without GVHD of IPAS-01, IPAS-02, and Combined Analyses With and Without Penalty Scores

Protein	Rp	IPAS-01, Average	IPAS-02, Average	IPAS-01, Maximum	IPAS-02, Maximum	Penalty Score	IPAS-01, Final Score	IPAS-02, Final Score	Combined, Final Score	Combined, Log$_{10}$ Final Score	Combined No Penalty, Log$_{10}$ Final Score
RPL23	85.29		64.27		78.55	1.00	0.00	6,699.15	6,699.15	3.83	3.83
DHX37	2.41	929.02		929.02		1.00	2,241.87	0.00	2,241.87	3.35	3.35
CENPM	15.57	62.18	34.80	62.18	34.80	1.00	968.38	542.00	968.38	2.99	2.99
MAP3K6	4.81	11.59	102.87	11.59	102.87	1.00	55.79	495.16	495.16	2.69	2.69
KDM5A	5.02	62.43		62.43		1.00	313.67	0.00	313.67	2.50	2.50
IL1A	9.76	31.70		31.70		1.00	309.42	0.00	309.42	2.49	2.49
FHOD1	3.00	88.44		88.44		1.20	221.18	0.00	221.18	2.34	2.42
CSE1L	3.96	43.56		43.56		1.00	172.34	0.00	172.34	2.24	2.24
ILF2	12.98		12.50		12.50	1.00	0.00	162.28	162.28	2.21	2.21
TRPC4AP	6.03	15.79		24.79		1.00	149.53	0.00	149.53	2.17	2.17
RAF1	51.64	2.85		2.85		1.00	147.00	0.00	147.00	2.17	2.17
GUK1	8.27	14.25		14.25		1.08	109.13	0.00	109.13	2.04	2.07
NUP160	18.85		5.88		5.88	1.08	0.00	102.62	102.62	2.01	2.04
VEZF1	0.83	81.91		81.91		1.00	67.98	0.00	67.98	1.83	1.83
VGLL2	1.46	38.42		38.42		1.00	56.11	0.00	56.11	1.75	1.75
EPHB4	8.87	5.97		5.97		1.00	52.98	0.00	52.98	1.72	1.72
CD8A	26.18	1.65		1.95		1.00	51.06	0.00	51.06	1.71	1.71
PIN1	27.16	1.41	1.92	1.41	1.92	1.10	34.81	47.41	47.41	1.68	1.72
ARID1A	19.20		1.20		1.75	1.00	0.00	33.59	33.59	1.53	1.53
TRIM39	0.64	51.56		51.56		1.00	33.00	0.00	33.00	1.52	1.52
WAPAL	17.43		1.74		1.74	1.00	0.00	30.31	30.31	1.48	1.48
GIMAP8	0.47	63.31		63.31		1.00	29.97	0.00	29.97	1.48	1.48
FAS	24.16	1.39	1.20	1.39	1.20	1.20	27.91	24.16	27.91	1.45	1.53
LEKR1	0.23		101.89		101.89	1.00	0.00	23.18	23.18	1.37	1.37
CXCL12	24.21	1.52		1.74		2.10	20.06	0.00	20.06	1.30	1.62
DBNL	6.84	2.27	1.68	2.31	1.79	1.28	12.34	9.56	12.34	1.09	1.20
STX7	10.03		1.62		1.70	1.50	0.00	11.34	11.34	1.05	1.23
PTPRD	5.68	1.93		1.93		1.10	9.96	0.00	9.96	1.00	1.04
RAB3IL1	2.66	3.65		3.65		1.00	9.69	0.00	9.69	0.99	0.99
IL13RA1	4.04	0.57	1.85	0.57	2.13	1.00	2.32	8.60	8.60	0.93	0.93
POGK	0.42	20.56		20.56		1.00	8.55	0.00	8.55	0.93	0.93
OIP5	3.05	1.73		1.73		1.00	5.28	0.00	5.28	0.72	0.72
GPNMB	2.53	1.19	1.65	1.19	2.04	1.00	3.01	5.16	5.16	0.71	0.71
BOD1	0.52		9.42		9.42	1.00	0.00	4.90	4.90	0.69	0.69
PVRL3	2.84	1.39	1.13	1.51	1.64	1.00	4.29	4.66	4.66	0.67	0.67
CRIPT	1.50		2.94		2.94	1.00	0.00	4.42	4.42	0.64	0.64
EFCAB6	2.79		1.54		1.54	1.00	0.00	4.30	4.30	0.63	0.63
UBR2	2.25	0.99		1.85		1.00	4.18	0.00	4.18	0.62	0.62
SRGN	2.44	1.56	1.51	1.56	1.64	1.00	3.80	4.00	4.00	0.60	0.60
TRIM37	1.58	16.05		16.05		6.52	3.89	0.00	3.89	0.59	1.40
IFI44	0.65		5.46		5.46	1.00	0.00	3.56	3.56	0.55	0.55
SEMA7A	1.72	0.71	2.33	0.71	2.33	1.20	1.02	3.35	3.35	0.52	0.60

(Continued on following page)

TABLE 1. Sixty-One Proteins in the Signature of GVT Without GVHD of IPAS-01, IPAS-02, and Combined Analyses With and Without Penalty Scores (Continued)

Protein	Rp	IPAS-01, Average	IPAS-02, Average	IPAS-01, Maximum	IPAS-02, Maximum	Penalty Score	IPAS-01, Final Score	IPAS-02, Final Score	Combined, Final Score	Combined, Log$_{10}$ Final Score	Combined No Penalty, Log$_{10}$ Final Score
CD58	1.83	1.81	0.66	1.81	0.66	1.00	3.30	1.21	3.30	0.52	0.52
LYPLA1	3.30	1.37		1.37		1.50	3.01	0.00	3.01	0.48	0.66
FERMT3	1.57	1.47	1.75	1.62	1.89	1.00	2.54	2.96	2.96	0.47	0.47
CD160	0.95	2.75		2.75		1.00	2.60	0.00	2.60	0.42	0.42
SDCCAG3	1.19	1.49		1.49		1.00	1.77	0.00	1.77	0.25	0.25
ART3	0.46		2.57		3.57	1.00	0.00	1.65	1.65	0.22	0.22
LY75	1.17	1.45		1.45		1.10	1.55	0.00	1.55	0.19	0.23
CREG1	0.64	1.50		1.50		1.00	0.96	0.00	0.96	−0.02	−0.02
CRTAM	0.56	1.46		1.62		1.00	0.90	0.00	0.90	−0.05	−0.05
SSC5D	0.60	1.31		1.31		1.00	0.79	0.00	0.79	−0.10	−0.10
FAM65B	0.39	1.95		1.95		1.00	0.76	0.00	0.76	−0.12	−0.12
PRR12	0.46		1.39		1.39	1.00	0.00	0.65	0.65	−0.19	−0.19
NBEAL2	0.44	1.33		1.33		1.00	0.59	0.00	0.59	−0.23	−0.23
GOLM1	0.38	0.94	1.27	0.94	1.27	1.00	0.36	0.49	0.49	−0.31	−0.31
LILRB1	0.32	1.22	1.48	1.22	1.67	1.10	0.36	0.49	0.49	−0.31	−0.27
LECT2	1.62		1.11		1.31	5.18	0.00	0.41	0.41	−0.39	0.33
CLEC11A	0.25	1.42	1.37	1.42	1.43	1.00	0.36	0.36	0.36	−0.44	−0.44
NTM	0.17	1.27		1.57		1.00	0.27	0.00	0.27	−0.57	−0.57
NBL1	0.17		1.68		2.64	5.36	0.00	0.08	0.08	−1.09	−0.36

Abbreviations: GVHD, graft-versus-host disease; GVT, graft versus tumor; IPAS, intact-protein analysis system; Rp, relevance score.

a GVHD-free GVT signature from the pre- and post-DLI comparison in the validation set (Table 1). The GeneTerrain visualization is shown in Figure 2B. A major cluster is constituted by centromere protein M (CENPM), NUP160, WAPAL, DHX37, STX7, and IL enhancer binding factor 2 (ILF2), which are all proteins implicated in the activation of cytotoxic T cells consistent with a GVT response (Table 2). For example, CENPM (also called PANE1) is a known miHAg expressed on B-lymphoid cells that is highly relevant to GVT-mediated post-DLI.[12] The protein encoded by ILF2 is a transcription factor needed for T-cell expression of IL-2.[13] The second cluster is similar to the one found in IPAS-01 centered on MAP3K6 and DBNL.

Final 61-Protein GVHD-Free GVT Signature After DLI

The next step was to combine the proteins found in the training and validation sets. Combined analysis of the two IPAS experiments yielded 61 proteins: 49 with a final combined score of greater than 1 and 12 with a final combined score of less than 1, all specific to GVT without GVHD (Table 1). The GeneTerrain visualization is shown in Figure 2C. Four main clusters are seen. These include one similar to that seen in IPAS-01 centered on CD8A, one similar to that seen in IPAS-02 centered on CENPM-WAPAL

and also containing DHX37, and one found in both experiments and centered on MAP3K6 and DBNL. A new fourth cluster appeared as a result of the enrichment in proteins specific for GVT without GVHD by the combination of the two sets. This additional cluster is centered around CSE1L, which is a RAS-related nuclear protein with a potential role in RAS/RAF/MAPK signaling in T cells. Some potential important novel markers, such as cytotoxic and regulatory T-cell molecule (CRTAM), which has been shown to determine the CD4+ cytotoxic T-lymphocyte lineage,[14] are not part of a cluster. A simple explanation is that CRTAM has been discovered recently, so the literature on CRTAM is only starting to emerge, which means that databases, such as STRING (Search Tool for the Retrieval of Interacting Genes/Proteins), have not yet integrated PPI networks for this protein.

Forty-Three of 61 Markers Are Found in a Single-Cell RNA Sequencing Analysis of PRAME- and Cytomegalovirus-Specific T Cells Post-HCT Relapse

Because the signature contained mostly intracellular proteins and additional samples from the discovery sets were not available to us, we next used single-cell profiling on T cells sorted and stimulated by the tumor antigen PRAME from a patient who relapsed after HCT compared with

FIG 2. Biomarker selection through GeneTerrain visualization tools shown in a two-dimensional (2D) panel. (A) Intact-protein analysis system (IPAS)-01: GeneTerrain visualization shown in 2D for the graft-versus-tumor (GVT) signature obtained by comparing GVT-positive, graft-versus-host disease (GVHD)–negative with GVT-positive, GVHD-positive post–donor lymphocyte injection (DLI) samples. The nodes between genes represent the protein-protein interaction pairs from the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) database. The protein-protein interaction confidence score in STRING contains functional protein associations derived from in-house predictions and homology transfers as well as from several externally maintained databases. Each interaction is assigned a score between 0 and 1, which is meant to be the probability that the interaction really exists given the available evidence. (B) IPAS-02: GeneTerrain visualization shown in 2D for the GVT signature from the pre- and post-DLI comparison. (C) Combined IPAS-01 and IPAS-02: GeneTerrain visualization shown in 2D for GVT without GVHD from both analyses.
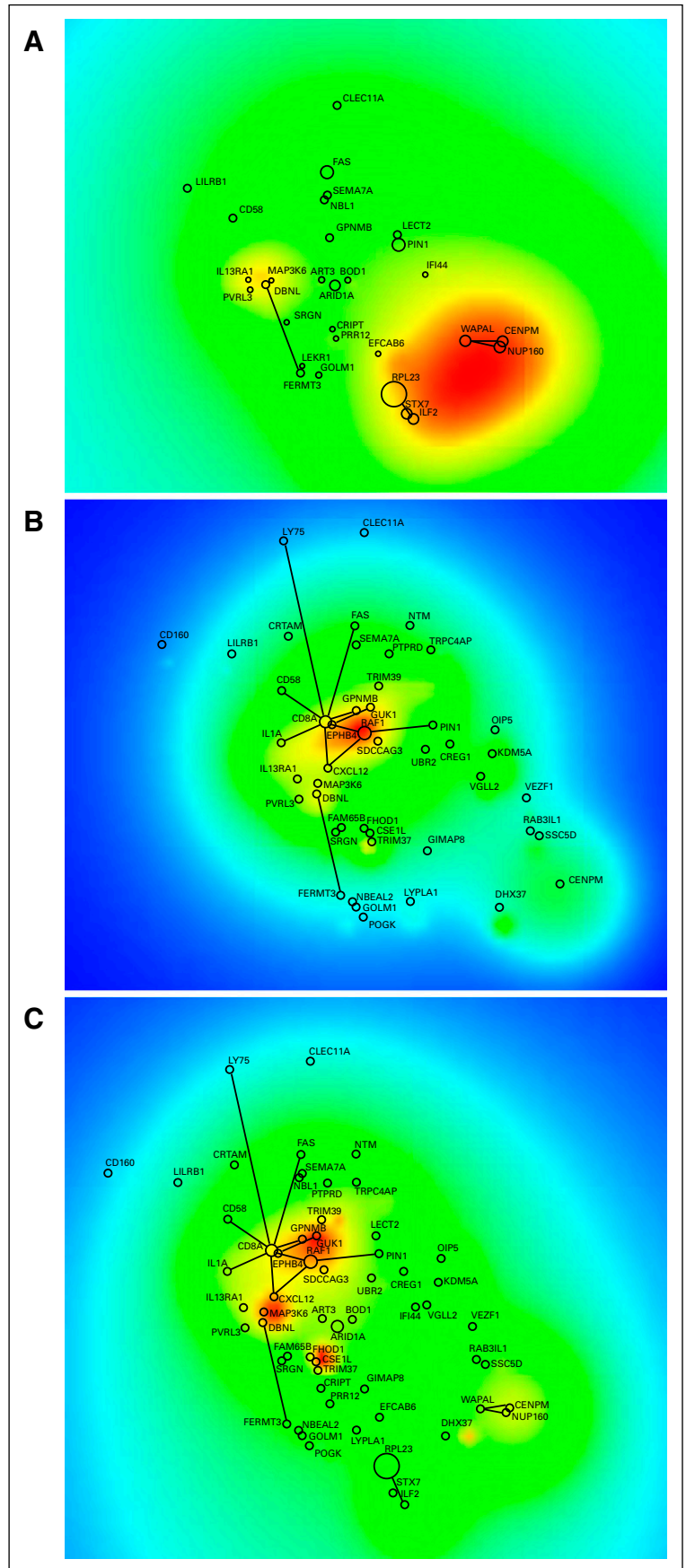
TABLE 2. GVT Biomarkers Mapped on CMV- and PRAME-Specific T Cells Through Single-Cell RNA Sequencing

Gene	T-Cell Classification	Cellular Location
RPL23	PRAME T-cell enriched	Intracellular
CENPM	PRAME T-cell enriched	Intracellular
ILF2	PRAME T-cell enriched	Intracellular
TRPC4AP	PRAME T-cell enriched	Membrane
GUK1	PRAME T-cell enriched	Intracellular
PIN1	PRAME T-cell enriched	Intracellular
WAPAL	PRAME T-cell enriched	Intracellular
CD58	PRAME T-cell enriched	Membrane
FERMT3	PRAME T-cell enriched	Membrane
CD160	PRAME T-cell enriched	Membrane, extracellular
CRTAM	PRAME T-cell enriched	Membrane
OIP5	Global T-cell activation	Intracellular
CRIPT	Global T-cell activation	Intracellular
IFI44	Global T-cell activation	Intracellular
LYPLA1	Global T-cell activation	Intracellular
NBL1	Global T-cell activation	Membrane
KDM5A	Global T-cell activation	Intracellular
CSE1L	Global T-cell activation	Intracellular
RAF1	Global T-cell activation	Intracellular
CD8A	Global T-cell activation	Membrane
TRIM39	Global T-cell activation	Intracellular
CXCL12	Global T-cell activation	Membrane, intracellular, extracellular
SRGN	Global T-cell activation	Intracellular
DHX37	Focal T-cell expression	Intracellular
MAP3K6	Focal T-cell expression	Intracellular
NUP160	Focal T-cell expression	Intracellular
VEZF1	Focal T-cell expression	Intracellular
ARID1A	Focal T-cell expression	Intracellular
FAS	Focal T-cell expression	Intracellular
DBNL	Focal T-cell expression	Membrane, intracellular, extracellular
STX7	Focal T-cell expression	Intracellular
POGK	Focal T-cell expression	Intracellular
BOD1	Focal T-cell expression	Intracellular
PVRL3	Focal T-cell expression	Membrane
UBR2	Focal T-cell expression	Intracellular
TRIM37	Focal T-cell expression	Intracellular
SDCCAG3	Focal T-cell expression	Intracellular
LY75	Focal T-cell expression	Membrane, extracellular
CREG1	Focal T-cell expression	Extracellular

(Continued in next column)

TABLE 2. GVT Biomarkers Mapped on CMV- and PRAME-Specific T Cells Through Single-Cell RNA Sequencing (Continued)

Gene	T-Cell Classification	Cellular Location
FAM65B	Focal T-cell expression	Membrane, intracellular
PRR12	Focal T-cell expression	Membrane, intracellular
NBEAL2	Focal T-cell expression	Membrane, intracellular
FHOD1	Focal T-cell expression	Intracellular
NTM	No or scant expression in T cells	Membrane
IL1A	No or scant expression in T cells	Extracellular
VGLL2	No or scant expression in T cells	Intracellular
EPHB4	No or scant expression in T cells	Membrane, intracellular, extracellular
GIMAP8	No or scant expression in T cells	Intracellular
LEKR1	No or scant expression in T cells	Not known
PTPRD	No or scant expression in T cells	Membrane
RAB3IL1	No or scant expression in T cells	Intracellular
IL13RA1	No or scant expression in T cells	Membrane
GPNMB	No or scant expression in T cells	Membrane
EFCAB6	No or scant expression in T cells	Intracellular
SEMA7A	No or scant expression in T cells	Membrane
ART3	No or scant expression in T cells	Membrane, extracellular
SSC5D	No or scant expression in T cells	Membrane, intracellular, extracellular
GOLM1	No or scant expression in T cells	Membrane, intracellular
LILRB1	No or scant expression in T cells	Membrane, intracellular, extracellular
LECT2	No or scant expression in T cells	Intracellular, extracellular
CLEC11A	No or scant expression in T cells	Intracellular, extracellular

Abbreviations: CMV, cytomegalovirus; GVT, graft versus tumor.

T cells sorted and stimulated by the viral antigen cytomegalovirus (CMV) pp65a as a surrogate measurement for an antitumoral response post-HCT. A mean reads number of 381,079 per cell and median gene number of 1,091 per cell were analyzed for an average of 1,436 cells per condition. Among the 61 genes identified in the proteomic GVHD-free GVT signature, 11 were more highly expressed in PRAME-specific T cells compared with CMV-specific T cells or nonreactive T cells. Thirty-two genes were enriched in both CMV- and PRAME-specific T cells, which represent a general activation marker of T cells. In contrast, 18 genes were not expressed in T-cell populations. The function and main GeneGO (Clarivate,

Philadelphia, PA) process for each protein are described in Table 2. The expression profile of four representative GVT markers in the single-cell RNA sequencing analysis of PRAME-specific T cells (RPL23, ILF2, CD58, and CRTAM) is shown in Figure 3. RPL23 is a component of the 60S ribosomal subunit and has been shown to link the oncogenic RAS signaling to p53-mediated tumor suppression.[15] T-cell responses to RPL23 also are increased in autoimmune diseases,[16] which suggests a role for the RAS/RAF/MAPK pathway in T cells that is currently underexplored. CD58 is the ligand of CD2 on T cells, and engagement of both lead to T-cell activation and adhesion. The CD58/CD2 axis is the primary costimulatory pathway for CD8+ T cells that lack CD28, which suggests an alternate activation mechanism during GVT for exhausted T cells.[17] CRTAM is upregulated in CD4+ and CD8+ T cells and encodes a type 1 transmembrane protein

with V and C1-like immunoglobulin domains.[18] It has been shown to negatively regulate ZEB1 (zinc finger E-box–binding homeobox 1) in T cells[19] and to determine the CD4+ cytotoxic T-lymphocyte lineage,[14] and it might determine the CD8+ cytotoxic T-lymphocyte lineage as well. The expression profiles of all 61 GVHD-free GVT markers are shown in the Data Supplement. These results suggest that tumor-specific T cells serve as a source of cells that express the GVT signature but require additional validation in future studies.

DISCUSSION

The current study has identified the first plasma proteomic signature for GVHD-free GVT after DLI using an in-depth tandem mass spectrometry–based analysis of plasma combined with a novel six-step systems biology approach. Previous studies have focused on cellular T-cell markers.[20]
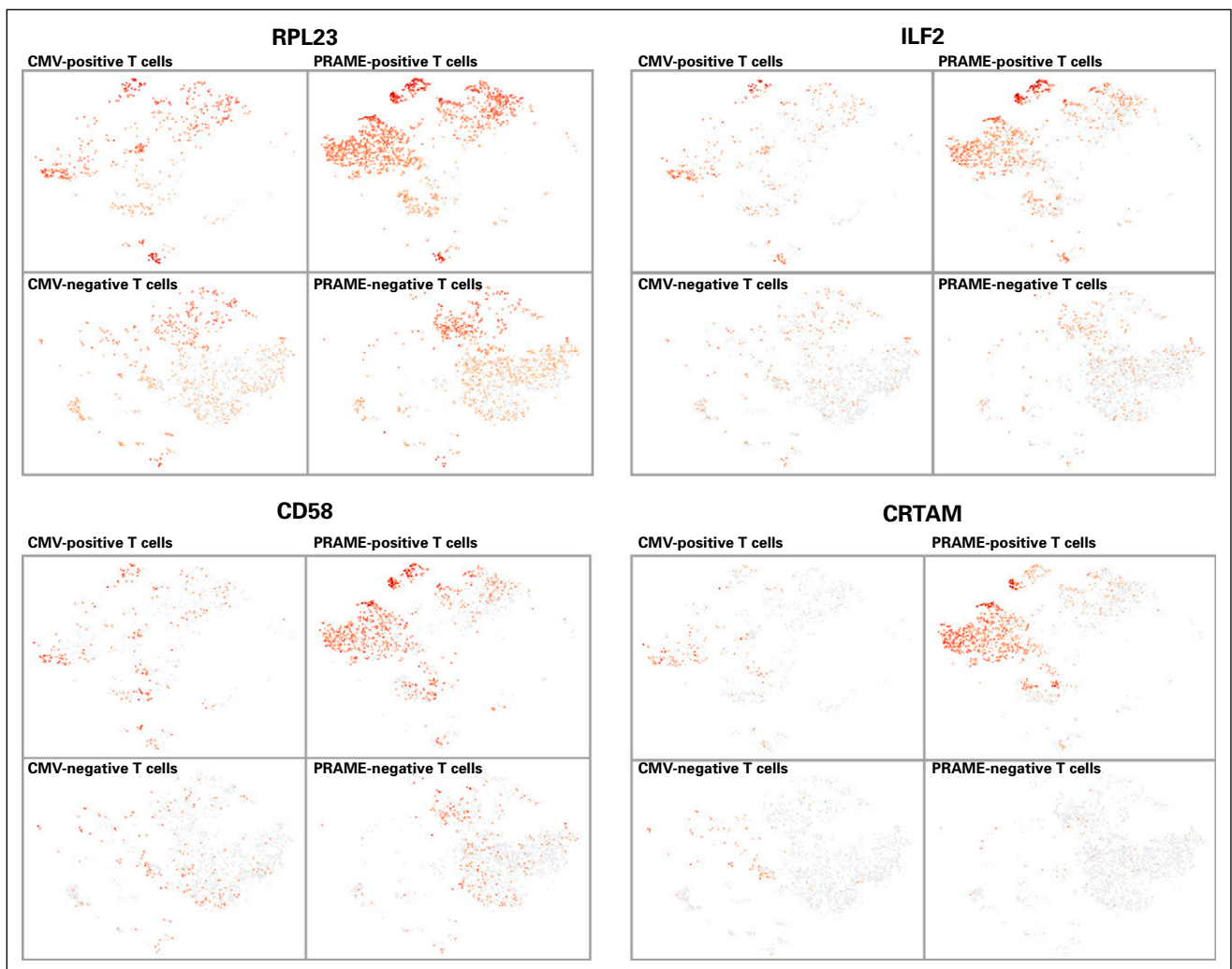

FIG 3. Expression profile of four representative graft-versus-tumor (GVT) markers in single-cell RNA sequencing analysis of PRAME- and cytomegalovirus (CMV) pp65–specific T cells. Expression profiles are shown for representative GVT markers RPL23, ILF2, CD58, and CRTAM in CMVpp65-positive or -negative T cells and PRAME-positive or -negative T cells.

This pipeline allowed us to identify a final plasma signature of 61 proteins from initially thousands of proteins. One systems biology novelty of our approach was the performance of a one-layer extension on PPI data using interactions one node away from the gene in the original list that are called outer genes. However, we found that this strategy did not significantly change the final signature of 61 proteins (data not shown). A strength of our approach is that we filtered out nonrelevant proteins using a penalty score, which led to a more-specific list of candidate proteins and avoided contaminants. This study favored a large-scale proteomics approach as opposed to a hypothesis-driven candidate approach.[21] We have shown that for GVHD markers, this method is efficient in discovering new candidate markers.[10,11] Compared with our previous studies, we experimentally removed the GVHD proteins by assigning them the light isotope. Of note, this approach showed that the proteins identified and their ratio have not been influenced much by the implementation of penalty scores. In an independent experiment with single-cell profiling of T cells from a patient with relapse after HCT that were activated in vitro, 43 of the 61 proteomic genes also were found in activated T cells, which suggests that the proteins identified are biologically relevant in different antitumoral responses.

The biology of the GVT markers not yet described is as follows. The function of TRPC4AP (transient receptor potential cation channel subfamily C member 4–associated protein) has been shown to be involved in the ubiquitination of E3 ligase skp2[22] and the activation of c-Jun NH(2) terminal kinase and transcription factor AP-1.[23] Guanylate kinase 1 (GUK1) is an enzyme that catalyzes the transfer of a phosphate group from ATP to guanosine monophosphate (GMP) to form guanosine diphosphate (GDP) and is believed to be a good target for cancer chemotherapy.[24] Its expression on tumor-specific T cells was not previously reported. PIN1 (peptidylprolyl *cis/trans* isomerase, NIMA-interacting 1) catalyzes the *cis/trans* isomerization of peptidyl-prolyl peptide bonds and thus catalytically regulates the postphosphorylation conformation of its substrates and is involved in the regulation of T-cell biology. In particular, its implication has been shown in systemic lupus erythematosus and T-cell acute lymphoblastic leukemia progression.[25-27] WAPL cohesin release factor has been shown to restrict chromatin loop extension.[28] Of note, it was part of a microRNA-mRNA network in allogeneic T-cell responses.[29] Fermitin family member 3 (FERMT3) is a member of the kindlins that mediates PPI involved in integrin activation. Mutations in this gene cause the autosomal recessive leukocyte adhesion deficiency syndrome-3.[30] Its role in T cells has not been studied. CD160 is another surface protein tightly expressed on peripheral cytotoxic CD8 T lymphocytes and natural killer cells,[31] and soluble CD160 enhances CD8+ T cells, which results in increased interferon-γ, IL-2, and tumor necrosis

factor-α secretion as well as cytolysis against tumor cells in vitro and in vivo.[32]

Not surprisingly, serglycin (SRGN), which serves as a mediator of granule-mediated apoptosis through the macromolecular complex of granzymes and perforin and determines the secretory granule repertoire of cytotoxic T lymphocytes,[33] was also upregulated in this study. The Fas cell surface death receptor (FAS) that plays a critical role in the activation of the death-inducing signaling complex with Fas-associated death domain protein and triggers a downstream caspase cascade that leads to apoptosis[34] also was upregulated.

The chemokine CXCL12 has been proposed to be able to distinguish immune cells that induce GVT going to the bone marrow from immune cells that induce GVHD.[35] It is one of the rare proteins for which the final score was modified more than twice by the penalty score but is still included overall as a GVT protein. Several of the tripartite motif (TRIM) members encode for miHAgs.[21] Although TRIM42, 22, and 37 are located on chromosomes 3, 11, and 17, respectively, TRIM39 is on chromosome 6 in the major histocompatability class I region and, in our study, showed global expression on activated T cells. RAF1 is the cellular homolog of viral raf proto-oncogene (v-raf) and is also a MAP3K that functions downstream of the Ras family of membrane-associated GTPases to which it binds directly. CSE1L is a RAS-related nuclear protein that binds strongly to nuclear localization signal-free importin-α, and this complex is then released in the cytoplasm by the combined action of RANBP1 and RANGAP1. The role of RAS/RAF/MAPK and phosphatidylinositol 3-kinase signaling in T cells remains understudied. One particular protein of non–T-cell origin, IL-1α, was found to be elevated in the plasma of patients with GVT response and may play a significant role in antitumor effects. Of note, IL-1α–engineered tumor cells rarely develop into tumors, and if they do, the tumors are quickly destroyed through a mechanism that involves CD8+ T cells and natural killer cells, and IL-1α enhances immunoediting in the tumor microenvironment.[36]

Although our approach identified several proteins for GVHD-free GVT, there are limitations in this study. First, plasma samples were pooled together. Second, only a small percentage of the whole circulating proteins are identified and quantified with tandem mass spectrometry. Third, the systems biology pipeline relies on knowledge from the published domain, which makes the application of this method to diseases that have not been well studied, such as GVT, difficult. Finally, because most of the proteins identified were intracellular, we could not apply immunoassays as we have done previously. As a result of the lack of samples available from additional DLI patients, we sought to look at the 61 genes from the proteomic signature in T cells from a patient who experienced relapse after HCT

and that were stimulated by the tumor antigen PRAME as a surrogate measurement for an antitumoral response post-HCT. Forty-three of these 61 genes also were found with this technique. Although challenged by the paucity of samples available after DLI, validation of this proteomic signature in larger cohorts is warranted.

AFFILIATIONS

[1]Indiana University School of Informatics and Computing, Indianapolis, IN

[2]Indiana University School of Medicine, Indianapolis, IN

[3]University of Alabama at Birmingham School of Medicine, Birmingham, AL

[4]Fred Hutchinson Cancer Research Center, Seattle, WA

[5]University of Michigan, Ann Arbor, MI

[6]MD Anderson Cancer Center, Houston, TX

[7]National Heart, Lung, and Blood Institute, Bethesda, MD

CORRESPONDING AUTHOR

Sophie Paczesny, MD, PhD, Indiana University School of Medicine, 1044 W Walnut St, Room R4-425, Indianapolis, IN 46202; e-mail: sophpacz@iu.edu.

EQUAL CONTRIBUTION

X.L., Z.Y., and Y.C. contributed equally to the study. S.I., J.Y.C., H.W., and S.P. contributed equally to the study.

AUTHOR CONTRIBUTIONS

Conception and design: Xiaowen Liu, Jake Y. Chen, Huanmei Wu, Sophie Paczesny

Financial support: Sophie Paczesny

Provision of study material or patients: Sung W. Choi, Sophie Paczesny

Collection and assembly of data: Xiaowen Liu, Zongliang Yue, Sawa Ito, Sophie Paczesny

Data analysis and interpretation: Xiaowen Liu, Zongliang Yue, Yimou Cao, Lauren Taylor, Qing Zhong, Sung W. Choi, Samir Hanash, Sawa Ito, Sophie Paczesny

Manuscript writing: All authors

Final approval of manuscript: All authors

Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/po/author-center.

Samir Hanash

Patents, Royalties, Other Intellectual Property: Patents submitted for lung and pancreatic cancer diagnostic markers (Inst)

Jake Y. Chen

Travel, Accommodations, Expenses: WuXi AppTec

Sophie Paczesny

Patents, Royalties, Other Intellectual Property: Inventor on a patent entitled "Methods of detection of graft-versus-host disease" (US 20130115232A1, WO 2013066369A3)

No other potential conflicts of interest were reported.

REFERENCES

1. Ringdén O, Labopin M, Gorin NC, et al: Is there a graft-versus-leukaemia effect in the absence of graft-versus-host disease in patients undergoing bone marrow transplantation for acute leukaemia? Br J Haematol 111:1130-1137, 2000

2. Bar M, Sandmaier BM, Inamoto Y, et al: Donor lymphocyte infusion for relapsed hematological malignancies after allogeneic hematopoietic cell transplantation: Prognostic relevance of the initial CD3+ T cell dose. Biol Blood Marrow Transplant 19:949-957, 2013

3. Mathew NR, Baumgartner F, Braun L, et al: Sorafenib promotes graft-versus-leukemia activity in mice and humans through IL-15 production in FLT3-ITD-mutant leukemia cells. Nat Med 24:282-291, 2018 [Erratum: Nat Med 24:526, 2018]

4. Warren EH, Greenberg PD, Riddell SR: Cytotoxic T-lymphocyte-defined human minor histocompatibility antigens with a restricted tissue distribution. Blood 91:2197-2207, 1998

5. Bleakley M, Otterud BE, Richardt JL, et al: Leukemia-associated minor histocompatibility antigen discovery using T-cell clones isolated by in vitro stimulation of naive CD8+ T cells. Blood 115:4923-4933, 2010

6. Dossa RG, Cunningham T, Sommermeyer D, et al: Development of T-cell immunotherapy for hematopoietic stem cell transplantation recipients at risk of leukemia relapse. Blood 131:108-120, 2018

7. Invernizzi R, Travaglino E, Lunghi M, et al: Survivin expression in acute leukemias and myelodysplastic syndromes. Leuk Lymphoma 45:2229-2237, 2004

8. Koestner W, Hapke M, Herbst J, et al: PD-L1 blockade effectively restores strong graft-versus-leukemia effects without graft-versus-host disease after delayed adoptive transfer of T-cell receptor gene-engineered allogeneic CD8+ T cells. Blood 117:1030-1041, 2011

9. Kim YM, Sachs T, Asavaroengchai W, et al: Graft-versus-host disease can be separated from graft-versus-lymphoma effects by control of lymphocyte trafficking with FTY720. J Clin Invest 111:659-669, 2003

10. 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 369:529-539, 2013

11. Yu J, Storer BE, Kushekhar K, et al: Biomarker panel for chronic graft-versus-host disease. J Clin Oncol 34:2583-2590, 2016

12. Brickner AG, Evans AM, Mito JK, et al: The PANE1 gene encodes a novel human minor histocompatibility antigen that is selectively expressed in B-lymphoid cells and B-CLL. Blood 107:3779-3786, 2006

13. Kao PN, Chen L, Brock G, et al: Cloning and expression of cyclosporin A- and FK506-sensitive nuclear factor of activated T-cells: NF45 and NF90. J Biol Chem 269:20691-20699, 1994

14. Takeuchi A, Badr MS, Miyauchi K, et al: CRTAM determines the CD4+ cytotoxic T lymphocyte lineage. J Exp Med 213:123-138, 2016

15. Meng X, Tackmann NR, Liu S, et al: RPL23 links oncogenic RAS signaling to p53-mediated tumor suppression. Cancer Res 76:5030-5039, 2016

16. Ito Y, Hashimoto M, Hirota K, et al: Detection of T cell responses to a ubiquitous cellular protein in autoimmune disease. Science 346:363-368, 2014

17. Leitner J, Herndler-Brandstetter D, Zlabinger GJ, et al: CD58/CD2 is the primary costimulatory pathway in human CD28-CD8+ T cells. J Immunol 195:477-487, 2015

18. Yeh JH, Sidhu SS, Chan AC: Regulation of a late phase of T cell polarity and effector functions by Crtam. Cell 132:846-859, 2008

19. Rojas-Marquez C, Valle-Rios R, Lopez-Bayghen E, et al: CRTAM is negatively regulated by ZEB1 in T cells. Mol Immunol 66:290-298, 2015

20. Bachireddy P, Hainz U, Rooney M, et al: Reversal of in situ T-cell exhaustion during effective human antileukemia responses to donor lymphocyte infusion. Blood 123:1412-1421, 2014

21. Warren EH, Deeg HJ: Dissecting graft-versus-leukemia from graft-versus-host-disease using novel strategies. Tissue Antigens 81:183-193, 2013

22. Jamal A, Swarnalatha M, Sultana S, et al: The G1 phase E3 ubiquitin ligase TRUSS that gets deregulated in human cancers is a novel substrate of the S-phase E3 ubiquitin ligase Skp2. Cell Cycle 14:2688-2700, 2015

23. Soond SM, Terry JL, Riches DW: TRUSS, a tumor necrosis factor receptor-1-interacting protein, activates c-Jun NH(2)-terminal kinase and transcription factor AP-1. FEBS Lett 580:4591-4596, 2006

24. Fitzgibbon J, Katsanis N, Wells D, et al: Human guanylate kinase (GUK1): cDNA sequence, expression and chromosomal localisation. FEBS Lett 385:185-188, 1996

25. Chae U, Park SJ, Kim B, et al: Critical role of XBP1 in cancer signalling is regulated by PIN1. Biochem J 473:2603-2610, 2016

26. Wei S, Yoshida N, Finn G, et al: Pin1-targeted therapy for systemic lupus erythematosus. Arthritis Rheumatol 68:2503-2513, 2016

27. Franciosa G, Diluvio G, Gaudio FD, et al: Prolyl-isomerase Pin1 controls Notch3 protein expression and regulates T-ALL progression. Oncogene 35:4741-4751, 2016

28. Haarhuis JHI, van der Weide RH, Blomen VA, et al: The cohesin release factor WAPL restricts chromatin loop extension. Cell 169:693-707.e14, 2017

29. Sun Y, Tawara I, Zhao M, et al: Allogeneic T cell responses are regulated by a specific miRNA-mRNA network. J Clin Invest 123:4739-4754, 2013

30. Ruppert R, Moser M, Sperandio M, et al: Kindlin-3-mediated integrin adhesion is dispensable for quiescent but essential for activated hematopoietic stem cells. J Exp Med 212:1415-1432, 2015

31. del Rio ML, Lucas CL, Buhler L, et al: HVEM/LIGHT/BTLA/CD160 cosignaling pathways as targets for immune regulation. J Leukoc Biol 87:223-235, 2010

32. Xiao H: Soluble CD160 enhances antitumor immunity against murine H22 hepatocarcinoma in vivo. Presented at the 14th Asia Pacific Oncologists Annual Meeting Melbourne, Australia, November 20-22, 2017

33. Sutton VR, Brennan AJ, Ellis S, et al: Serglycin determines secretory granule repertoire and regulates natural killer cell and cytotoxic T lymphocyte cytotoxicity. FEBS J 283:947-961, 2016

34. Peter ME, Hadji A, Murmann AE, et al: The role of CD95 and CD95 ligand in cancer. Cell Death Differ 22:885-886, 2015

35. Kotsiou E, Davies JK: New ways to separate graft-versus-host disease and graft-versus-tumour effects after allogeneic haematopoietic stem cell transplantation. Br J Haematol 160:133-145, 2013

36. Voronov E, Dotan S, Krelin Y, et al: Unique versus redundant functions of IL-1α and IL-1β in the tumor microenvironment. Front Immunol 4:177, 2013