Discovery and Validation of Graft-versus-Host Disease Biomarkers

Sophie Paczesny, MD, PhD

Blood and Marrow Transplant Program, University of Michigan, Ann Arbor, MI; and Bone Marrow and Stem Cell Transplantation Program, Indiana University Melvin and Bren Simon Cancer Center, Riley Hospital for Children, and Wells Center for Pediatric Research, Indianapolis, IN

Correspondence: Sophie Paczesny, MD, PhD, Bone Marrow and Stem Cell Transplantation Program, Indiana University Melvin and Bren Simon Cancer Center and Wells Center for Pediatric Research, 1044 W. Walnut Street, Room 425, Indianapolis, IN, Ph: 317-278-5487, Fax: 317-274-0138, E-mail: sophpacz@iupui.edu

Running Title: Biomarkers of Graft-versus-Host Disease
Abstract

Allogeneic hematopoietic stem cell transplantation (allo-HSCT) is the most effective tumor immunotherapy available. Although allo-HSCT provides beneficial graft-versus-tumor effects, acute graft-versus-host disease (aGVHD) is the primary source of morbidity and mortality after HSCT. Diagnosis of aGVHD is typically based on clinical symptoms in one or more of the main target organs (skin, liver, gastrointestinal tract) and confirmed by biopsy. However, currently available diagnostic and staging tools often fail to identify patients at higher risk of GVHD progression, unresponsiveness to therapy, or death. In addition, there are shortcomings in the prediction of GVHD before clinical signs develop, indicating the urgent need for noninvasive and reliable laboratory tests. Through the continuing evolution of proteomics technologies seen in recent years, plasma biomarkers have been identified and validated as promising diagnostic tools for GVHD and prognostic tools for non-relapse mortality. These biomarkers may facilitate timely and selective therapeutic intervention but should be more widely validated and incorporated into a new grading system for risk stratification of patients and better-customized treatment. This review identifies biomarkers for detecting GVHD, summarizes current information on aGVHD biomarkers, proposes future prospects for the blinded evaluation of these biomarkers, and discusses the need for biomarkers of chronic GVHD.
Introduction

Allo-HSCT is a widely used therapy for a variety of malignant and nonmalignant hematologic diseases. In malignant disease, the donor immune system recognizes residual tumor cells as foreign and eradicates them by immunological means, known as the graft-versus-tumor (GVT) effect. Unfortunately, donor immune cells may also attack normal host tissue, particularly the skin, liver, and gastrointestinal (GI) tract, resulting in acute graft-versus-host disease (aGVHD). GVHD remains one of the major barriers to a more widespread and successful application of allo-HSCT. Known risk factors that increase the incidence and severity of aGVHD include the use of HLA-mismatched unrelated donors rather than HLA-matched sibling donors.¹

A major barrier to GVHD research and treatment is that diagnosis and prognosis rely almost entirely on the presence of clinical symptoms and should be confirmed by biopsy of the involved target organs. No currently used laboratory tests can predict the risk of developing GVHD, responsiveness to treatment, or patient survival. The importance of biomarkers in HSCT settings is crucial, because the ability to identify patients at high risk for GVHD early in their transplantation and treatment course has important therapeutic consequences including more stringent monitoring and/or preventive care. The ability to identify patients who will not respond to traditional treatment and are at particularly high risk for subsequent morbidity and mortality could enable tailored treatment plans such as additional immunosuppressive treatments that might be more effective if introduced early.
This review provides an update on the discovery and validation of the most clinically relevant biomarkers of aGVHD and insights for specific recommendations on their use in clinical trials.

**GVHD biomarker qualification**

A good diagnostic GVHD test will be able to distinguish patients with GVHD from those without GVHD (e.g., GI GVHD vs. infectious colitis, skin GVHD vs. drug rash). Furthermore, the biomarker should (i) specifically and sensitively reflect the disease state, (ii) be noninvasive (e.g., blood or urine test), and (iii) be rapid, simple, accurate, inexpensive, and standardized. Ideally the same test should also be prognostic for outcomes such as response to treatment, survival, and non-relapse mortality (NRM) to allow for early risk stratification before initiation of treatment. A biomarker that predicts aGVHD before clinical signs appear will have tremendous impact and allow for preemptive interventions. The ideal biomarker for GVHD should differentiate between GVHD and GVT effects.

**Single versus multiple biomarkers**

When several biomarkers exist, usually no single biomarker is sufficiently sensitive or specific on its own for either diagnostic or predictive testing. Thus, the simultaneous use of several markers may increase specificity or predictive/diagnostic performance. Presumably, a combination of tissue-specific and systemic biomarkers will be more informative for aGVHD diagnosis. However, if a biomarker is not highly correlated to other biomarkers or clinical predictors, one or two biomarkers could be
sufficient. Proportional odds logistic regression models are used to create a composite panel that will generate a receiver operating characteristic (ROC) curve. The ROC curve is a plot of the true-positive rate (sensitivity = 1 – false negative error rate) versus the false-positive rate (1 – specificity), and is associated with rules that classify an individual as “positive” if is marker value is above a threshold c for all possible thresholds. 3-6

Types of GVHD biomarkers

MicroRNAs. MicroRNAs (miRNAs) are 21- to 25-nucleotide transcripts that repress gene function through interaction with target mRNAs. MiRNAs control gene activity at multiple levels, specifically transcription, translation, and protein degradation. Circulating miRNAs have recently been reported as promising biomarkers in diverse diseases7 including GVHD.8

Cellular biomarkers. The functions and numbers of several different immune cell populations are altered in GVHD. Some of these immune cell subsets, particularly regulatory T cells, seem promising as biomarkers of aGVHD9-11 and cGVHD.12 Dendritic cells, monocytes, and gamma-delta T cells are also promising as markers after HSCT.13,14 In cGVHD, B cells and their modulators (e.g., B cell-activating factor) are important biomarkers.15

Proteomic biomarkers. The term “proteomics” indicates PROTEins expressed by a genOME and describes the systematic analysis of protein profiles of a biological sample. Unlike the genome, the proteome varies with time and is defined as the proteins present in a single sample at a certain time point. Thus, proteins represent
ideal biomarkers in the post-transplant setting and have been widely studied, as
detailed in the following sections.

**Proteomics strategies in the search for new GVHD biomarkers**

*The study of blood by proteomics*

Ideal clinical tests are based on noninvasive collection, which enables repetitive
sample collection from the same patient in a short amount of time. Enormous effort has
been made to develop standardized methods for clinical sample collection for proteomic
studies.\(^{16,17}\)

The levels of individual blood proteins represent a summation of multiple, disparate
events that occur in every organ system. Plasma and sera contain proteins shed by the
affected tissue as well as proteins that reflect secondary systemic changes. However,
plasma and sera are highly complex mixtures containing high levels of many different
proteins with a wide dynamic range, spanning 12 orders of magnitude from albumin to
the lowest abundance protein. Low-abundance proteins such as cytokines and their
receptors\(^ {18,19}\) are often the most clinically relevant. To detect low-abundance proteins,
depletion of predominant proteins and subsequent fractionation of the proteome are
required.

Urine samples represent an alternative to plasma/sera samples for biomarker
discovery. Urine has three main advantages compared with plasma/sera: (i) it can be
obtained in large quantities; (ii) the protein mixture is far less complex and the variation
in protein abundance is low\(^ {20}\); and (iii) it is more stable.\(^ {21}\) However, urine yields better
information about diseases in organs that are directly involved in its production and excretion, such as the kidneys.

_Tissue proteomics_

Although blood biomarkers are ideal for use in a clinical setting, one goal of research into the fundamental biology of GVHD is to identify markers that are target-tissue specific. Thus, the ideal sample for discovery of biologically relevant GVHD proteins may be the target-tissue itself. However, finding tissue-specific markers has thus far proven difficult because of the cellular heterogeneity of tissues, and the limited material available in biopsies for tissue proteomics. To date, there is no method capable of amplifying the amount of proteins requiring, at best, pooling of several biopsies.\textsuperscript{22-24}

_Proteomics approaches for biomarker discovery_

Advances in engineering have increased data throughput, enabling the study of complete sets of molecules ("-omics") with exponentially increased speed, accuracy, and cost-effectiveness. Thus, analysis of the entire spectrum of molecular and cellular organization is now possible, providing insight into the mechanism of disease, with fewer \textit{a priori} assumptions. However, arriving at proteins from genes (~20,000) requires two more levels of complexity: the transcriptome (~100,000 RNA transcripts) and proteome (~1,000,000 proteins). Here, this review focuses on proteomics for the molecular diagnosis of GVHD post-HSCT, since proteins are more informative than other cellular metabolites regarding the ongoing pathophysiology of a disease.\textsuperscript{25} Both mass spectrometry (MS)-based and non-MS-based proteomic approaches have been used to screen potential GVHD biomarkers.
**Array-based approaches.** Immunoassays are analytical tests that harness the unique properties of antibodies and are therefore extremely sensitive. The unique characteristics of antibodies are derived from two important properties: (i) their exceptional binding specificity that enables the measurement of picomolar ($10^{-12}$) protein amounts in blood samples; and (ii) the binding strength between an antibody and its target that makes tests accurate and precise, even at low concentrations. To screen for aGVHD biomarkers, microarrays dotted with hundreds of antibodies measure hundreds of proteins in complex biological matrices. The advantages of immunoassays are that they are: (i) suited for the characterization of complex protein mixtures such as human plasma, (ii) quantitative, (iii) able to detect low-abundance proteins such as cytokines, and (iv) high-throughput. The disadvantages are limitations to the number of antibodies included in each array (i.e., introduction of bias) and high cross-reactivity between antibodies and nontarget proteins.

**The advent of MS for protein analysis.** Most non-antibody proteomic strategies are based on MS, which is a powerful tool for characterizing and assessing qualitative and quantitative changes in complex protein mixtures. Two types of MS techniques have been used in clinical proteomics: (i) pattern profiling and (ii) detailed protein characterization. Pattern profiling compares polypeptide spectra obtained by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) MS to show which patients suffer from a particular disease without identifying individual profile components. A variant of this is surface-enhanced laser desorption/ionization (SELDI-TOF) MS. Because factors influencing the final oligopeptide profiles of body fluid samples are so complex, MS profiling has not yet met the standards required for clinical
practice. This technique has been used to screen aGVHD biomarker candidates in both serum\textsuperscript{30} and saliva.\textsuperscript{31}

Other approaches rely on protein separation employing gel-based techniques followed by MS\textsuperscript{32-35}. However, gel-free separation methods such as liquid chromatography (LC)\textsuperscript{36-39} and capillary electrophoresis\textsuperscript{40} provide better separation because they overcome several limitations of gel separation, such as lengthy analysis time; poor separation of proteins with low molecular weight, or extreme isoelectric point. MS is the final step in the analytical procedure and reliably identifies proteins and determines their isoforms and posttranslational modifications. MS also allows unambiguous quantification, particularly when tandem MS (MS/MS) is used\textsuperscript{36,37,41,42}. New instrumentation such as the ultra-high resolution linear ion trap Orbitrap mass spectrometer also facilitates top-down LC-MS/MS\textsuperscript{43}.

These approaches are not suitable for validation purposes because of the time required for analysis, but they remain the most efficient methods for biomarker discovery in clinical research.

\textit{Informatics to mine disease proteomes and prioritize candidate biomarkers.} Massive amounts of complex and heterogeneous proteomic data are generated from a single experiment. Many proteins are identified, but prioritization is uncertain with no mechanism for attrition and few available high-quality reagents to validate these discoveries. Proteomics workflows that have led to biomarker discovery include the one detailed here\textsuperscript{38,39}. First, the acquired spectra are matched to a sequence database to identify proteins, with a false discovery rate <5\%.\textsuperscript{44,45} False
discovery rates are now recognized as markers of significance in -omics studies.\textsuperscript{46,47} To sequentially refine the list of candidate proteins, proteins can be selected based on relationship to the tissue of origin using the Human Protein Atlas website, which aims to annotate human proteins using antibodies to systematically analyze the cellular distribution of proteins in normal and pathologic tissues.\textsuperscript{48} Another refinement can be obtained by using pathway analysis tools. Physician scientists are usually well versed in the complex pathophysiology of their disease of interest and might be more successful at mining their dataset based on information from published literature than pathway analysis, which depends on the validity of the annotations of the different molecules and their biological processes. Then, candidate proteins would require further validation by other biotechnological techniques, ideally those that allow for high-throughput as detailed in the next paragraph. The steps of these workflows are summarized in Figure 1.

\textbf{Proteomics approach for high-throughput validation of GVHD biomarkers}

Although proteomics holds great promise for biomarker development, gaps still remain between biomarker discovery and biomarker validation. Most noteworthy is the paucity of affinity-capture reagents, such as high-quality antibodies, which leads to a bias in the prioritization of candidate markers. Furthermore, the number of samples required for validation increases as the biomarker advances through each test phase, augmenting the need for high-throughput assays. The most applicable approach for
quantitating individual proteins for validation is the sandwich enzyme-linked immunosorbent assay (ELISA), which is highly specific and uses two antibodies specific for the candidate protein. This procedure is also relatively simple and highly reproducible among performers and laboratories, limiting inter- and intra-assay variability. Finally, because of the urgent need for GVHD blood tests, testing proteins primarily with available antibodies is an excellent strategy. The main disadvantage of validation by ELISA is the large volume of plasma required. Thus, multiplexing technologies are preferred if cross-reactivity does not occur. \(^{27}\) Recently, selected reaction monitoring-mass spectrometry (SRM-MS) that can measure the absolute quantification of multiple native peptides, and by extension, their parent proteins, has emerged as a potentially useful technique to screen and quantify targeted multiplexed proteins in patient plasma samples with high sensitivity, absolute specificity, and high throughput. \(^{49-51}\)

**Identified and validated GVHD biomarkers**

GVHD is not only a systemic immunological disorder; it also affects specific organ systems, including the skin and GI tract. Noteworthy systemic and target-specific aGVHD biomarkers have been reviewed along with their diagnostic and prognostic values and whether they were identified using aGVHD pathology or a proteomics discovery approach (Table 1).

**Biomarkers identified using the pathology of aGVHD**

**Systemic biomarkers of aGVHD.** Because of the cytokine storm that occurs early after donor graft infusion, cytokines and their receptors have been tested as potential aGVHD
biomarkers.\textsuperscript{52} Soluble IL-2 receptor α chain (sIL-2Rα) concentrations were increased in aGVHD patients in many studies.\textsuperscript{53-57} In two studies, IL-18 concentrations were closely correlated with IL-2Rα concentrations, and IL-2Rα concentrations with GVHD severity.\textsuperscript{56} However, some studies found that sIL-2Rα concentrations were also increased in patients with other transplantation-related complications (TRCs), such as veno-occlusive disease and sepsis.\textsuperscript{53} Similarly, TNFα and its receptors, particularly TNFR1 concentrations were higher in aGVHD patients than in patients without GVHD.\textsuperscript{58-62} The same precautions used to evaluate sIL-2Rα levels in other TRCs should be considered when evaluating TNFα and TNFR1. The roles of TNFα/TNFR1 and IL-2/IL-2Rα in aGVHD pathogenesis are supported by evidence suggesting that antibodies directed against TNFR1 or IL-2Rα are effective therapies for steroid-refractory aGVHD.\textsuperscript{63} C-reactive protein is a nonspecific inflammatory protein that is elevated in GVHD patients along with IL-6, the main cytokine that induces C-reactive protein release.\textsuperscript{61,64-66} Uguccioni et al. found that IL-8 concentrations correlated with aGVHD.\textsuperscript{67} However, Schots et al. showed that IL-8 was released in all types of TRCs, rather than specifically in cases of aGVHD.\textsuperscript{61} Similarly, increases in IL-8 and other cytokines (e.g., IL-6, IL-10, and IL-18) were not confirmed in a study of patients who underwent a reduced intensity conditioning regimen; however, this same study found that IL-12 was elevated in association with aGVHD development.\textsuperscript{68} Chemokines and chemokine receptors that were also implicated in immune cell migration from lymphoid organs to target organs, were also elevated in patients with aGVHD.\textsuperscript{69,70}

Hepatocyte growth factor (HGF) is a multifunctional cytokine secreted by mesenchymal cells that acts primarily on cells of epithelial origin. Okamoto et al.
observed higher serum HGF concentrations in patients who developed severe aGVHD than in patients who did not develop the disease. HGF appears to belong to a different category of biomarkers that represent a physiologic response to GVHD damage. In this respect, HGF seems similar to cytokeratin-18 (KRT18) fragments, markers of epithelial apoptosis that are associated with intestinal and hepatic GVHD damage. Indoleamine 2,3-dioxygenase, the rate-limiting enzyme in the kynurenine pathway of tryptophan degradation, functions in a potent immunoregulatory loop and is involved in the pathology of GVHD. Recently, LC-MS/MS was used to measure major tryptophan metabolites in serial urine specimens from 51 patients. Surviving patients had significantly lower metabolite levels on days 28, 42, and 90 after HSCT than patients dying of GVHD.

**Target-specific biomarkers of aGVHD.** As mentioned above, KRT18 is the only target-specific marker of GVHD that is based on GVHD pathology. In addition, few biomarker studies have evaluated their prognostic value for outcomes such as NRM, primarily because of small sample sizes. One exception is a recent study that measured KRT18 along with markers of endothelial dysfunction and found that patients with steroid-refractory GVHD were not exposed to an overwhelming T-cell attack but to a progressive microangiopathy that led to organ failure. Few studies have evaluated the predictive value of biomarkers. Studies measuring the kinetics of cytokine concentrations during the first month post-HSCT showed that monitoring IL-2Rα and TNFR1 concentrations during the engraftment period allows for early detection of aGVHD. The mean serum IL-2Rα and TNFR1 concentrations appear to increase during the first 2 weeks following transplantation in patients with aGVHD.
August et al. reported that sIL-2R, TNFR1, and CD8 have high predictive values for aGVHD occurrence. A recent study by Rezvani et al. found that a decrease of 0.5 g/dl in serum albumin from pretransplantation levels predicted the development of grade III/IV aGVHD and overall survival 6 months after initiation of aGVHD treatment in a cohort of 401 patients with aGVHD grades II–IV after reduced intensity HSCT. Because measuring serum albumin concentration is easy and inexpensive, the authors suggest incorporating albumin into the set of other validated biomarkers to improve the prediction of aGVHD severity and mortality.

**Biomarkers identified using proteomics discovery**

**Systemic biomarkers of aGVHD.** Four studies have identified the proteomic pattern of aGVHD using MS-based approaches on small sample sizes. One group identified a pattern that was validated in a larger study. Imanguli et al. used complementary proteomic techniques to characterize the salivary proteome of 41 patients who underwent allo-HSCT. Hori et al. used the SELDI technique to screen plasma proteins specific for aGVHD in a mouse model. One peak appeared to distinguish GVHD plasma from non-GVHD plasma and was identified as CCL8. Using SELDI-TOF MS, Srinivasan et al. identified an aGVHD-specific peptide pattern in training samples that was validated in an independent set of aGVHD samples obtained on the day aGVHD symptoms appeared. Kaiser et al. used capillary electrophoresis-MS to identify peptide patterns in urine samples as early indicators of aGVHD development. Two prominent GVHD-indicative polypeptides were identified as a 1.85-kDa peptide from leukotriene A4 hydrolase and a 1.83-kDa peptide from albumin. This
aGVHD-specific peptide set pattern was used to screen 63 samples collected from 33 patients after allo-HSCT. A subsequent blind evaluation of 599 samples from 141 patients enabled the prediction of aGVHD before the appearance of clinical symptoms with a sensitivity of 83% and specificity of 76%. Using MS/MS, 3 of the 31 peptides that contributed to the aGVHD pattern were identified as fragments from collagen α-1 chain I (downregulated) and α-1 chain III (upregulated).

Paczesny et al. identified and validated a panel of proteins using an antibody microarray. To validate this aGVHD biomarker panel, they randomly divided the samples into a training set (282 patients) and a validation set (142 patients), which is currently one of the most reliable approaches for validation. This approach identified and validated a four-protein biomarker panel for GVHD diagnosis (IL-2Rα, TNFR1, IL-8, and HGF) with high specificity. Because of the large sample size of this study, Paczesny et al. provided the first evidence that these biomarkers are associated with GVHD clinical outcomes and prognosis.

**Target-specific biomarkers of aGVHD.** Clinical symptoms of the skin (e.g., maculopapular rash) and GI tract (e.g., nausea, diarrhea) caused by GVHD can be difficult to distinguish from other causes (e.g., infection, drugs). Thus, biomarkers specific for GVHD and target organs may improve diagnosis. Using intact-protein analysis system, plasma pooled from 10 patients with skin GVHD labeled with the heavy isotope was compared with plasma from 10 patients without GVHD labeled with the light isotope for the analysis of thousands of spectra and hundreds of proteins. Elafin emerged as the lead biomarker candidate for detecting skin GVHD at the time of clinical diagnosis. Paczesny et al. demonstrated that plasma elafin concentrations in
samples from 492 patients had significant diagnostic and prognostic value, including the prediction of long-term survival. These data also provide proof-of-principle demonstration that biomarkers of disease-related tissue-specific changes can be detected in plasma. Using the same proteomics strategy, Ferrara et al. discovered 74 proteins with increased ratios (i.e., heavy to light) in patients with GI GVHD, and five were of GI origin. Regenerating islet-derived 3-alpha (REG3\(\alpha\)) was the lead candidate and was validated as a biomarker of lower GI GVHD using ELISA in two independent sets including 1,014 patients from three centers. In a follow-up study, REG3\(\alpha\) was compared with KRT18 and HGF, which were previously identified as GI GVHD markers, and showed higher diagnostic precision for lower GI GVHD than the other two biomarkers.

There are also limitations in predicting response to GVHD therapy. Recently, Luft et al. showed that markers of endothelial dysfunction are elevated in steroid-refractory GVHD. Levine et al. measured six previously validated diagnostic biomarkers from samples prospectively obtained at treatment initiation and on days 14 and 28 in a multicenter, randomized, four-arm phase II clinical trial for newly diagnosed aGVHD. At each time point, the panel of six biomarkers predicted important clinical outcomes (post-therapy nonresponse at day 28 and mortality at day 180 from onset).

**Decision making strategies.**

Because GVHD target-specific markers are primarily indicative of GVHD damaged tissue, they represent indicators of the disease when clinical signs are present. Therefore, GVHD target-specific markers complementary to biomarkers and
infectious agents’ analyses are useful in distinguishing skin rashes and diarrhea from other causes at time of diagnosis. If only one target, such as diarrhea, is involved, the most specific GI GVHD biomarker described so far, REG3α, should be tested. If both diarrhea and rashes are present at the same time, which is suggestive of a more systemic disease, REG3α and elafin should be tested along with systemic biomarkers.

GVHD should be analyzed as a single disease process early in the transplant course because mechanisms involved in alloreactivity start as soon as donor T cells are activated and in contact with host antigen-presenting cells. We expect that biomarkers will be identified that correlate with subclinical disease. A good example of such a marker is sIL-2Rα that was shown in a prospective study to be elevated at days 7 and 15 post-HCST and to predict severe aGVHD with 57% sensitivity and 82% specificity.

Biomarkers of GVHD have also recently been used to follow the response to GVHD treatment. However, the key to developing useful surrogate endpoints is to identify biomarkers that reflect fundamental aspects of the treatment’s effects on disease pathogenesis. A “perfect” surrogate endpoint, as described by Prentice, can be presented as T induces S that induces O, where T is the treatment, S is the surrogate endpoint, and O is a clinical outcome. In this case, the biomarker S measures all the effects of T on O. A more likely situation arises when T has a direct effect on O that is not mediated through S.
Design considerations for future interventional trials with biomarkers

Given the progress in GVHD biomarker identification and validation, clinical trial design will begin incorporating biomarkers. The number of specimens that should be tested depends on the objective of the study and the extent of biomarker variability in the study. The following factors contribute to variability: the prevalence of the disease; the subtypes of disease among the study samples (e.g., skin GVHD or GI GVHD); the capacity of the biomarkers to discriminate among the different disease subtypes; the number of biomarkers under study; and the statistical algorithm used to select promising biomarkers. Thus, as suggested by Pepe et al., there are no simple methods for recommending samples sizes. They proposed that computer simulations guide the choice of sample sizes with the guidance of investigators on biologically plausible models to generate data.

Study using diagnostic biomarkers in place of invasive biopsies. Target-specific diagnostic biomarkers that can differentiate skin GVHD from other rashes and GI GVHD from other forms of enteritis will replace invasive biopsies. Observational trials in which samples and biopsies taken at GVHD onset are prospectively compared should be the first step. The goal is to replace invasive tests (biopsies) that have specificity and sensitivity greater than 85% with a blood test that should achieve similar specificity and sensitivity. The outcome evaluated is the presence or absence of GVHD which has an incidence in most centers of approximately 50%. If we define similar sensitivity and specificity as being within 15 percentage points of 85% (70% or more),
50 patients per arm or a total of 100 patients would be needed to conclude that a blood test measuring biomarkers can replace biopsies.

**Study using biomarkers that predict unresponsiveness to GVHD therapy.**

Another potential clinical application of GVHD biomarkers is to stratify patients based on risk at the time of GVHD onset before initiation of therapy. GI GVHD is considered high risk for mortality, but without further risk stratification, the standard of care for all patients with GI GVHD is prompt initiation of systemic steroid treatment, with second-line agents reserved for patients who fail initial therapy. Unfortunately, most patients who require second-line therapy die, highlighting the need to refine risk beyond the current grading system. The University of Michigan BMT group recently developed a risk stratification algorithm for patients with new onset GI GVHD that incorporates clinical stage, histologic grade, and plasma levels of the newly discovered GI GVHD biomarker, REG3α.\(^{38}\)

In addition, early identification of patients at high risk for steroid unresponsiveness may allow alternative testing or additional therapies before the development of refractory disease. Martin et al. recently highlighted the urgent need to develop well-designed prospective phase II treatment trials with a well-established benchmark of success, good statistical plan, and well-defined primary endpoint for GVHD treatment studies, which still need to be standardized between centers.\(^{82}\) Biomarkers might facilitate decision making if their incorporation can be validated in a prospective randomized trial. To estimate the effectiveness of measuring a biomarker, the trial must be designed such that participants are randomized to different groups: one
in which a biomarker is measured and provided to the clinician and another in which it is not (Figure 2). The next step is to include the intervention and control strategies. A schema for treating newly diagnosed GVHD using biomarkers is shown in Figure 3. Comparison of box B and box D shows if additional treatment at onset of GVHD improves response rates and lowers non-relapse mortality in high risk patients identified by biomarkers. Comparison of box A and box C shows if GVHD treatment is less toxic (faster taper and shorter steroids exposition may decrease infection rates) in low risk patients identified by biomarkers. As previously mentioned, the “best” treatment for treating GVHD is still under debate. The expectation is that these markers will also have a specificity and sensitivity similar to diagnostic markers. The outcomes measured in this study are response 28 days post-treatment and non-relapse mortality 180 days post-treatment. Approximately 50 to 70 patients may need to be included in each arm for a total of 200 to 280 patients in the study.

Study using biomarkers that predict risk for future occurrence of GVHD to pre-emptively treat GVHD. To determine whether their validated biomarkers can predict GVHD before the appearance of clinical symptoms, Paczesny et al. evaluated the four most informative biomarkers (i.e., IL2R-α, TNFR1, elafin, and REG3α) in day 7 and 14 post-HSCT samples from 513 patients who underwent unrelated HSCT and had not yet developed GVHD. Measurement of this biomarker panel pre-HSCT predicted grade II-IV GVHD with a specificity of 75% and sensitivity of 57%. Indicators of success must include not only a lower incidence of GVHD but also reductions in infectious complications and relapse. Ultimately, a randomized trial is needed to assess
the effectiveness of GVHD pre-emption. Assuming an early biomarker (e.g., at day 14 post–HSCT) exists with reasonable specificity and sensitivity, the study design should be defined narrowly, such that a single treatment strategy is clinically reasonable in the absence of the biomarker result. For GVHD, a reasonable treatment strategy would be standard steroid treatment (2 mg/kg/day) or low-dose steroid treatment (1 mg/kg/day), which may produce similar outcomes with less toxicity.85 A schema for a pre-emptive trial to decrease GVHD incidence using biomarkers is shown in Figure 4. Comparison of box B and box D shows if pre-emptive treatment lowers acute GVHD in high risk patients identified by biomarkers. The expectation is that these markers will have a lower specificity and sensitivity than diagnostic markers. It is also likely that at least one or two clinical parameters, most likely age and conditioning intensity, will need to be included. The outcomes measured are incidence of GVHD and non-relapse mortality 180 days post-HSCT. Taking an optimistic approach and assuming a simplified algorithm (no more than 4 parameters) with specificity and sensitivity greater than 80% has been identified, approximately 80 to 120 patients may need to be included in each arm for a total of 320 to 480 patients for the study.

**Future Directions**

Future directions include a blinded evaluation of these biomarkers with samples collected in a multicenter prospective study that will reduce center effects and facilitate the successful design of subsequent trials. Ideally, a single formula will be developed to predict a patient’s risk for aGVHD, allowing for personalized medicine.
Furthermore, biomarkers may represent novel therapeutic targets that could be inhibited by future aGVHD-specific drugs. Because these drugs would target the appropriate effector T cells, they should increase efficacy and lower toxicity.

GVHD was traditionally categorized as acute (i.e., arising before day 100 post-HSCT) or chronic (i.e., occurring after that time). This definition was updated by the new National Institutes of Health classification, which includes late-onset aGVHD (after day 100) and an overlap syndrome with features of both aGVHD and cGVHD. So far, development of post-HSCT biomarkers has focused on aGVHD biomarkers. However, future efforts in biomarker discovery and validation will be particularly valuable for cGVHD.

Conclusions

Proteomics is a revolutionary field that includes detection technologies for proteins, which are the molecules that best represent the real-time pathophysiology of alloreactivity. In a short time, proteomics has led to the identification of novel mechanisms of allo-HSCT that likely would not have been discovered by traditional hypothesis-driven research. A promising approach is to use protein biomarkers in risk stratification to better use current treatment modalities. Furthermore, the findings presented in this review demonstrate the potential of biomarkers for exploring targeted therapeutics. The principal barrier to be circumvented is the validation of biomarker concentrations in different allo-HSCT settings such as conditioning intensity or donor sources (particularly cord blood, T cell-depleted graft). Achieving this aim will require a
much larger validation study, ideally a multicenter prospective trial. Once an algorithm for each setting is established, personalized medicine will be possible.

Acknowledgments

Sophie Paczesny (S.P.) is supported by grants from the National Institutes of Health (RC1HL101102 and P01-CA039542). S.P. is an investigator of the Eric Hartwell Research fund and the Amy Strelzer Manasevit Research Program.

The author would like to thank Drs James Ferrara and Samir Hanash for their guidance and mentorship in discovery and validation of graft-versus-host disease biomarkers. The author would also like to thank Drs Thomas Braun and John Levine for their counseling and valuable comments in statistical design.

Conflict of interest statement

The author declares no competing financial interests.
References


64. Schots R, Van Riet I, Othman TB, et al. An early increase in serum levels of C-reactive protein is an independent risk factor for the occurrence of major complications and 100-day transplant-related mortality after allogeneic bone marrow transplantation. *Bone Marrow Transplant*. 2002;30(7):441-446.


Table 1A: Candidate Acute GVHD Biomarkers with diagnostic and prognostic significance identified based on GVHD pathology

<table>
<thead>
<tr>
<th>Protein</th>
<th>Study</th>
<th>Number of patients in the study</th>
<th>Target organ</th>
<th>DiagnosisEndpoint</th>
<th>AssociationDirection</th>
<th>Prognostic Endpoint</th>
<th>PredictiveEndpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albumin</td>
<td>Rezvani2011</td>
<td>401</td>
<td>Systemic/GI tract/renal</td>
<td>ND</td>
<td>↓</td>
<td>2 year NRM, 6 months OS in RIC patients</td>
<td>aGVHD 3-4 in RIC patients</td>
</tr>
<tr>
<td>CRP</td>
<td>Schots2002</td>
<td>96</td>
<td>Systemic</td>
<td>All TRC</td>
<td>↑</td>
<td>Day 100 NRM</td>
<td>TRC</td>
</tr>
<tr>
<td></td>
<td>Pihush2006</td>
<td>350</td>
<td>Systemic</td>
<td>All TRC</td>
<td>↑</td>
<td>ND</td>
<td>TRC</td>
</tr>
<tr>
<td>CXCL10</td>
<td>Piper2007</td>
<td>34</td>
<td>Systemic</td>
<td>aGVHD 0-4</td>
<td>↑</td>
<td>ND</td>
<td>aGVHD 0-4 by day 100</td>
</tr>
<tr>
<td>HGF</td>
<td>Okamoto2001</td>
<td>38</td>
<td>Systemic</td>
<td>aGVHD 0-4</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Harris2012</td>
<td>826+ 128*</td>
<td>GI tract</td>
<td>aGVHD 0-4</td>
<td>↑</td>
<td>6 months NRM</td>
<td>ND</td>
</tr>
<tr>
<td>IL-2Rα</td>
<td>Miyamoto1996</td>
<td>30</td>
<td>Systemic</td>
<td>aGVHD 0-4</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Foley1998</td>
<td>36</td>
<td>Systemic</td>
<td>aGVHD 0-4, sepsis, VOD</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Shaiegan2006</td>
<td>39</td>
<td>Systemic</td>
<td>aGVHD 2-3</td>
<td>↔</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>August2011</td>
<td>62</td>
<td>Systemic</td>
<td>ND</td>
<td>↑</td>
<td>ND</td>
<td>Prospective study aGVHD 3-4 by day 60</td>
</tr>
<tr>
<td>IL-6</td>
<td>Malone2007</td>
<td>147</td>
<td>Systemic</td>
<td>aGVHD 0-4 by day 100</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Schots2003</td>
<td>84</td>
<td>Systemic</td>
<td>All TRC</td>
<td>↑</td>
<td>1 year NRM</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Pihush2006</td>
<td>350</td>
<td>Systemic</td>
<td>All TRC</td>
<td>↑</td>
<td>ND</td>
<td>TRC</td>
</tr>
<tr>
<td>IL-8</td>
<td>Uguccioni1993</td>
<td>30</td>
<td>Systemic</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Schots2003</td>
<td>84</td>
<td>Systemic</td>
<td>↑</td>
<td>1 year NRM</td>
<td>TRC</td>
<td></td>
</tr>
<tr>
<td>IL-10</td>
<td>Liem1998</td>
<td>46</td>
<td>Systemic</td>
<td>aGVHD 2-4, cGVHD, infection,</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Protein</td>
<td>Study Year</td>
<td>Study Type</td>
<td>Condition</td>
<td>Marker</td>
<td>Change</td>
<td>Control</td>
<td>Notes</td>
</tr>
<tr>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>----------------------------</td>
<td>-------------------------</td>
<td>--------</td>
<td>---------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>IL-12</td>
<td>Mohty 2005</td>
<td>Systemic</td>
<td>aGVHD 2-4 in RIC patients</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>IL-18</td>
<td>Fujimori 2000</td>
<td>Systemic</td>
<td>aGVHD 0-4</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Shaiegan 2006</td>
<td>Systemic</td>
<td>aGVHD 2-3</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>KRT18</td>
<td>Luft 2007</td>
<td>GI tract</td>
<td>GI GVHD vs. non GVHD-enteritis</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
<td>Response to GVHD treatment</td>
</tr>
<tr>
<td></td>
<td>Luft 2011</td>
<td>Systemic</td>
<td>Refractory GVHD</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Harris 2012</td>
<td>GI tract</td>
<td>GI GVHD</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
<td>6 months NRM, response to GVHD treatment</td>
</tr>
<tr>
<td>REG3α</td>
<td>Harris 2012</td>
<td>GI tract</td>
<td>GI GVHD</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>Syndecan-1</td>
<td>Seidel 2003</td>
<td>Systemic</td>
<td>aGVHD 0-4</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>TNFα</td>
<td>Holler 1990</td>
<td>Systemic</td>
<td>All TRC</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>Abdallah 1997</td>
<td>Systemic</td>
<td>All TRC, aGVHD 0-1 vs. 2-4</td>
<td>↔</td>
<td>ND</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schots 2003</td>
<td>Systemic</td>
<td>All TRC</td>
<td>↑</td>
<td>1 year NRM</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>TNFR1</td>
<td>Choi 2008</td>
<td>Systemic</td>
<td>ND</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
<td>aGVHD 2-4, 1 year NRM</td>
</tr>
<tr>
<td></td>
<td>Kitko 2008</td>
<td>Systemic</td>
<td>ND</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
<td>aGVHD 2-4, 1 year NRM</td>
</tr>
<tr>
<td></td>
<td>August 2011</td>
<td>Systemic</td>
<td>ND</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
<td>Prospective study aGVHD 3-4 by day 60</td>
</tr>
<tr>
<td>Tryptophan metabolites</td>
<td>Landfried 2011</td>
<td>Systemic</td>
<td>ND</td>
<td>↑</td>
<td>6 months NRM</td>
<td>ND</td>
<td></td>
</tr>
</tbody>
</table>
Table 1B: Candidate Acute GVHD Biomarkers with diagnostic and prognostic significance identified with proteomics discovery

<table>
<thead>
<tr>
<th>Protein</th>
<th>Study</th>
<th>No of patients in the study</th>
<th>Target organ</th>
<th>Diagnosis Endpoint</th>
<th>Association Direction</th>
<th>Prognostic Endpoint</th>
<th>Predictive Endpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peptides set</td>
<td>Kaiser 2004</td>
<td>35</td>
<td>Systemic</td>
<td>aGVHD 0-4, sepsis</td>
<td>Presence vs. Absence</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Srinivasan 2006</td>
<td>34</td>
<td>Systemic</td>
<td>aGVHD</td>
<td>Presence vs. Absence</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Imanguli 2007</td>
<td>41</td>
<td>Systemic</td>
<td>1 month post-HSCT vs. pre-HSCT</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Weissinger 2007</td>
<td>33+141*</td>
<td>Systemic</td>
<td>aGVHD 2-4</td>
<td>Presence vs. Absence</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Proteins set</td>
<td>Wang 2005</td>
<td>9</td>
<td>Systemic</td>
<td>aGVHD vs. pre-HSCT</td>
<td>Presence vs. Absence</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Albumin</td>
<td>Kaiser 2004</td>
<td>35</td>
<td>Systemic</td>
<td>aGVHD 0-4, sepsis</td>
<td>Presence vs. Absence</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Imanguli 2007</td>
<td>41</td>
<td>Systemic</td>
<td>1 month post-HSCT vs. pre-HSCT</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCL8</td>
<td>Hori 2008</td>
<td>14</td>
<td>Systemic</td>
<td>aGVHD 0-4</td>
<td>Presence vs. Absence</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Collagen α-1</td>
<td>Weissinger 2007</td>
<td>33+141*</td>
<td>Systemic</td>
<td>aGVHD 2-4</td>
<td>Presence vs. Absence</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Elafin Paczesny 2010</td>
<td>20+ 492*</td>
<td>Skin</td>
<td>aGVHD 0-4 vs. non-GVHD rashes</td>
<td>↑</td>
<td>NRM, OS</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>HGF Paczesny 2009</td>
<td>42+ 282*+ 142*</td>
<td>Systemic/G I tract</td>
<td>aGVHD 0-4 vs. no GVHD</td>
<td>↑</td>
<td>NRM, OS</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>IL-2Rα Paczesny 2009</td>
<td>42+ 282*+ 142*</td>
<td>Systemic</td>
<td>aGVHD 0-4 vs. no GVHD</td>
<td>↑</td>
<td>NRM, OS</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>IL-8 Paczesny 2009</td>
<td>42+ 282*+ 142*</td>
<td>Systemic</td>
<td>aGVHD 0-4 vs. no GVHD</td>
<td>↑</td>
<td>NRM, OS</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Lactoferrin Imanguli 2007</td>
<td>41</td>
<td>Systemic</td>
<td>1 month post-HSCT vs. pre-HSCT</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Protein</td>
<td>Study Year</td>
<td>Sample Size</td>
<td>Disease</td>
<td>Condition</td>
<td>Outcome 1</td>
<td>Outcome 2</td>
<td>Outcome 3</td>
</tr>
<tr>
<td>------------------</td>
<td>------------</td>
<td>-------------</td>
<td>---------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>Leukotriene A4 hydrolase</td>
<td>Kaiser 2004</td>
<td>35</td>
<td>Systemic</td>
<td>aGVHD 0-4, sepsis</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>REG3α</td>
<td>Ferrara 2011</td>
<td>20+ 871*+ 143*</td>
<td>GI tract</td>
<td>GI-GVHD vs. non-GVHD enteritis</td>
<td>↑</td>
<td>NRM, OS</td>
<td>ND</td>
</tr>
<tr>
<td>SLPI</td>
<td>Imanguli 2007</td>
<td>41</td>
<td>Systemic</td>
<td>1 month post-HSCT vs. pre-HSCT</td>
<td>↑</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>TNFR1</td>
<td>Paczesny 2009</td>
<td>42+ 282*+ 142*</td>
<td>Systemic</td>
<td>aGVHD 0-4 vs. no GVHD</td>
<td>↑</td>
<td>NRM, OS</td>
<td>ND</td>
</tr>
</tbody>
</table>

* Validation on independent sets of patients
aGVHD: Acute Graft Versus Host Disease
cGVHD: Chronic Graft Versus Host Disease
MAHA: Microangiopathic Hemolytic Anemia
ND: Not Done
NRM: Non-Relapse Mortality
OS: Overall Survival
RIC: Reduced Intensity Conditioning
TRC: Transplant-Related complications, major complications including aGVHD, MAHA, VOD, pneumonitis and Infection
VOD: Veno-Occlusive Disease
↑ Plasma protein increased
↓ Plasma protein decreased
↔ Plasma protein unchanged
Figure legends

Figure 1. Proteomics biomarkers pipeline.

Figure 2. Fundamental design for a randomized trial to evaluate biomarker utility.

Figure 3. Biomarker strategy design for newly diagnosed GVHD.

Figure 4. GVHD pre-emptive trial design.
Biomarker Discovery

10s Biospecimens
>10000 MS/MS spectra

10-100 proteins
100s Biospecimens

Informatics
Peptide Prophet > 75%
Protein prophet > 95%
False Discovery rate < 5%
Quantification: Ratio heavy to light isotopes > 1.5)

10-100 Candidate Biomarkers

Biomarker Validation

10-100 proteins
100s Biospecimens

Data mining
Tissue of origin
Pathway analysis
Available immunoassays or multiple reaction monitoring

~10 Verified Biomarkers

Biomarker Qualification

< 10 proteins
1000s Biospecimens

Immuoassays only

Panel of Qualified Biomarkers
Study Biospecimen

Randomize

Measure Biomarker

Intervention tailored to biomarker concentration

Measure Outcomes

Do not measure Biomarker

Standard Intervention

Measure Outcomes
Newly diagnosed GVHD

- Onset GVHD Biospecimen
  - Randomize
    - Measure Biomarker
      - Low Risk: Standard Intervention (steroids)
        - Outcomes A
      - High Risk: "Best treatment" (2 agents)
        - Outcomes B
    - Do not measure Biomarker
      - Randomize
        - Standard Intervention (steroids)
          - Outcomes C
        - "Best treatment" (2 agents)
          - Outcomes D
Day +14 post-HSCT Biospecimen

Preemptive Trial

Measure Biomarker for all

Treat none

Low Risk
Placebo

High Risk
Placebo

Treat if High Risk

Low Risk
Placebo

High Risk
Standard GVHD Intervention (steroids)

A GVHD Incidence, 6 month-NRM
B GVHD Incidence, 6 month-NRM
C GVHD Incidence, 6 month-NRM
D GVHD Incidence, 6 month-NRM
Discovery and validation of graft-versus-host disease biomarkers

Sophie Paczesny

Information about reproducing this article in parts or in its entirety may be found online at: http://www.bloodjournal.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at: http://www.bloodjournal.org/site/misc/rights.xhtml#reprints

Information about subscriptions and ASH membership may be found online at: http://www.bloodjournal.org/site/subscriptions/index.xhtml

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include digital object identifier (DOIs) and date of initial publication.