frontiers in
IMMUNOLOGY

HYPOTHESIS AND THEORY ARTICLE
published: 06 November 2014
doi: 10.3389/fimmu.2014.00529

In silico derivation of HLA-specific alloreactivity potential from whole exome sequencing of stem-cell transplant donors and recipients: understanding the quantitative immunobiology of allogeneic transplantation

Max Jameson-Lee[1], Vishal Koparde[2†], Phil Griffith[1†], Allison F. Scalora[1†], Juliana K. Sampson[2], Haniya Khalid[1], Nihar U. Sheth[2], Michael Batalo[1], Myrna G. Serrano[2], Catherine H. Roberts[1], Michael L. Hess[3], Gregory A. Buck[2], Michael C. Neale[4], Masoud H. Manjili[5] and Amir Ahmed Toor[1]*

[1] *Stem Cell Transplant Program, Massey Cancer Center, Virginia Commonwealth University, Richmond, VA, USA*
[2] *The Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond, VA, USA*
[3] *Department of Internal Medicine, Virginia Commonwealth University, Richmond, VA, USA*
[4] *Department of Psychiatry and Statistical Genomics, Virginia Commonwealth University, Richmond, VA, USA*
[5] *Department of Microbiology and Immunology, Virginia Commonwealth University, Richmond, VA, USA*

Edited by:
Seiamak Bahram, Université de Strasbourg, France

Reviewed by:
Stanislaw Stepkowski, University of Toledo College of Medicine, USA
Myra Coppage, University of Rochester Medical Center, USA

****Correspondence:***
Amir Ahmed Toor, Stem Cell Transplant Program, Massey Cancer Center, Virginia Commonwealth University Health Systems, 1300 Marshall Street, Richmond, VA 23298-0157, USA
e-mail: atoor@vcu.edu

[†]Vishal Koparde, Phil Griffith and Allison F. Scalora have contributed equally to this work.

Donor T-cell mediated graft versus host (GVH) effects may result from the aggregate alloreactivity to minor histocompatibility antigens (mHA) presented by the human leukocyte antigen (HLA) molecules in each donor–recipient pair undergoing stem-cell transplantation (SCT). Whole exome sequencing has previously demonstrated a large number of non-synonymous single nucleotide polymorphisms (SNP) present in HLA-matched recipients of SCT donors (GVH direction). The nucleotide sequence flanking each of these SNPs was obtained and the amino acid sequence determined. All the possible nonameric peptides incorporating the variant amino acid resulting from these SNPs were interrogated *in silico* for their likelihood to be presented by the HLA class I molecules using the Immune Epitope Database stabilized matrix method (SMM) and NetMHCpan algorithms. The SMM algorithm predicted that a median of 18,396 peptides weakly bound HLA class I molecules in individual SCT recipients, and 2,254 peptides displayed strong binding. A similar library of presented peptides was identified when the data were interrogated using the NetMHCpan algorithm. The bioinformatic algorithm presented here demonstrates that there may be a high level of mHA variation in HLA-matched individuals, constituting a HLA-specific alloreactivity potential.

Keywords: alloreactivity potential, stem-cell transplant, whole exome sequencing, HLA, minor histocompatibility antigen

INTRODUCTION

Graft versus host disease (GVHD) is a major impediment in achieving optimal outcomes in patients undergoing allogeneic stem-cell transplantation (SCT) from human leukocyte antigen (HLA) identical related and unrelated donors (URD) (1–3). Further, it remains unclear why with only relatively minor variation in GVHD prophylaxis, some patients with HLA-matched donors develop severe GVHD, whilst others with HLA-mismatched donors may not experience any (4–6). In HLA-matched donor-recipient pairs (DRP), a major contributor to GVHD occurrence are the peptides encoded by loci outside the major histocompatibility (MHC) locus on chromosome 6. These peptides, functionally defined as minor histocompatibility antigens (mHA), are presented by specific HLA molecules and are responsible for both the clinically beneficial graft versus tumor responses, and the deleterious GVHD (7–10). As of 2012, around 49 mHA recognized by CD4+ or CD8+ T lymphocytes have been described (11). Further complicating this problem is the HLA specificity of various mHA, and the heterogeneity observed in the HLA distribution in various populations across the world (12, 13). Therefore, in order to understand the biology and role of mHA in generating GVHD, it is critical to quantify the extent of genetic variation between individuals.

Exploring genetic variation outside the MHC locus is also important to understand why, with relatively simple adjustments to the treatment protocols patients successfully engraft when transplanted with HLA-mismatched donors. This is true for both URD umbilical cord blood transplant, and related haploidentical SCT (6). Moreover, completely HLA-mismatched solid organ transplants result in successful engraftment, albeit with low-level life-long immunosuppression. Furthermore, organs, such as kidney and heart tissues, are prone to rejection when transplanted; yet, these organs are seldom targeted in GVHD, even in its chronic form, which affects nearly all organ systems. This makes it imperative to understand the role of mHA in generating alloreactivity, and the extent to which the magnitude of genetic variation outside the MHC locus contributes to allograft complications, such as GVHD or graft rejection.

To examine these quantitative relationships, whole exome sequencing of SCT donor and recipients genomes was performed to measure the antigenic variability existing between them (14). A large number of single nucleotide polymorphisms (SNP) were identified between donors and recipients. These differences were classified as, either possessing, a GVH vector, polymorphisms present at loci in the recipient and absent in the donor, or, a HVG vector, present in the donor and absent in the recipient. The large number of SNPs in the exome, termed *alloreactivity potential*, suggests that in all individuals undergoing SCT, there is a very high probability of there being peptides, which may function as mHA. However, given the observed frequency of GVHD, seemingly, not all of these SNPs would lead to immunogenic peptides being generated, to yield clinically relevant mHA responses. This may be because, for HLA class I molecules on an antigen-presenting cell to present a peptide to an effector T lymphocyte, first, the endogenous protein must be cleaved by the proteasome, then the resulting peptides must bind HLA class I molecules to be presented. This would initiate either an immune response or result in tolerance, depending on the cellular and cytokine milieu at the time of antigen presentation (15).

It is possible to determine the genetic variation between SCT recipients and donors, and to then bioinformatically determine the amino acid sequence of peptides resulting from SNPs encountered in their exomes. Further, bioinformatic techniques have been developed to determine which peptide antigens may be presented by specific HLA molecules. The Immune Epitope Database (IEDB; http://www.iedb.org) has characterized hundreds of thousands of peptides that can bind several hundred MHC complexes. From this large dataset, researchers have developed tools to predict peptide-HLA binding probabilities (16). Initially, matrix-based methods such as stabilized matrix method (SMM) (17) were developed to determine binding affinities. More recently, neural network-based algorithms such as NetMHC can use binding information from neighboring residues to predict dissociation constants between HLA molecules and putative mHA (18). Finally, "pan-specific" algorithms have developed that are able to predict peptide-binding HLA alleles with limited experimental binding data (19).

In this paper, the putative mHA in HLA-matched DRP and the *in silico* determined HLA class I binding affinity of these peptides is explored utilizing a bioinformatic approach based on exome sequencing of donors and recipients of SCT. The algorithm developed, lays a framework for future analysis of large SCT patient cohorts, and defines a personalized *HLA-specific alloreactivity potential*. The alloreactivity potential concept is analogous to the idea of potential energy in physics, i.e., the stored energy in a system. Thus, HLA-specific alloreactivity potential would give an estimate of the likelihood that GVHD or graft rejection may develop in a HLA-matched DRP in the absence of immunosuppression. Our work demonstrates that the number of potentially immunogenic peptides varies considerably across HLA-matched related (MR) and URD, constituting a large alloreactivity potential.

METHODS

WHOLE EXOME SEQUENCING

Patients with recurrent hematological malignancies enrolled in a Virginia Commonwealth University Institutional Review Board approved protocol (Clinicaltrial.gov identifier: NCT00709592) were included in this study. To identify all the potentially immunogenic differences that exist in a SCT DRP, whole exome sequencing was performed on previously cryopreserved DNA from the donors and recipients enrolled in this study as previously described (14). Of the nine DRP examined, four were from HLA-A, B, C, and DRB1 MRD, and 5 from URD. Histocompatibility testing was performed using high-resolution typing for both HLA class I (**Table 1**) and HLA class II loci (*not shown*). The whole exome sequence of individual donors and recipients was compared both within pairs, and to a reference genome to identify all the SNPs, which were subsequently characterized as either synonymous or non-synonymous. Next, all the non-synonymous SNP (nsSNP) present in the recipient, but absent in the donor were identified, and designated as possessing a graft versus host (GVH) vector ($nsSNP_{GVH}$).

Table 1 | HLA typing of the donor–recipient pairs.

D-RPair	HLA-A	HLA-A	HLA-B	HLA-B	HLA-C	HLA-C
2	02:01	24:02	15:16	27:05	02:02	17:01
3	03:01	11:01	07:02	55:01	03:03	07:02
4	23:01	30:02	15:03	44:03	02:10	07:18
5	01:01	03:01	570101	07:02	07:02	07:01
7	01:01	02:01	44:02	55:01	03:03	05:01
8	01:01	24:02	07:02	55:01	03:03	07:02
10	01:01	03:01	080101	40:01	03:04	07:01
16	01:01	26:01	13:02	27:05	02:02	06:02
23	03:01	24:02	07:02	57:01	06:02	07:02

Patients 2, 4, 16, and 23 underwent MRD and the others URD SCT. Patient 2 had a single locus HLA-B antigen mismatch; patients 3, 7, and 10 had a male donor/female recipient combination and others were gender matched.

DERIVING HLA-SPECIFIC ALLOREACTIVITY POTENTIAL

To derive the amino acid sequence of the oligopeptides, i.e., potential mHA, resulting from these nsSNPs and their binding affinity to the relevant HLA in each DRP, a bioinformatics pipeline was developed. This pipeline has the following components: (1) determine $nsSNP_{GVH}$ between the exomes of transplant donors and recipients; (2) generate putative immunogenic peptides *in silico* from these genomic differences; and (3) analyze the binding affinity of these polymorphic peptides to the HLA in that individual (**Figure 1**). This third step estimates the likelihood of these peptides to be presented by the six patient-specific HLA class I molecules to determine *candidate* mHA. A complete description of this bioinformatic pipeline follows.

CREATION OF PEPTIDE LIBRARIES

All the $nsSNP_{GVH}$ for each DRP were exported as variant call files (VCF) to the ANNOVAR software package (20). Next, using the DB SNP130 database and hg18 genome coordinates of the $nsSNP_{GVH}$, amino acid sequences of the putative peptides were generated using the "seq_padding" option of the "annotate_variation" function in ANNOVAR. Endogenous peptides are presented by HLA class I molecules, and the average length of peptides binding HLA

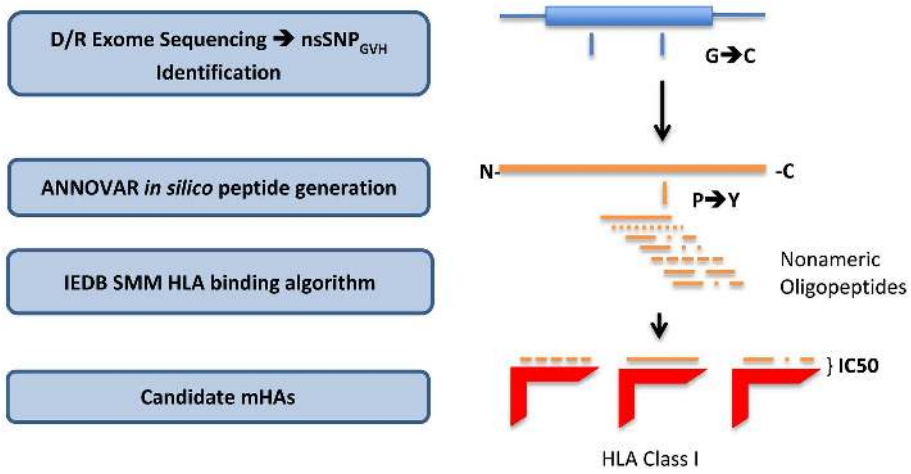

FIGURE 1 | Bioinformatics workflow for calculating HLA-specific alloreactivity potential in individual DRP. Starting with donor and recipient whole exome sequence data, non-synonymous SNP with a GVH vector (nsSNP$_{GVH}$) were identified, and peptide fragments generated using the ANNOVAR software package. These peptides, together with HLA data (**Table 1**) were then analyzed with IEDB SMM and NetMHCpan algorithms separately. Individual DRP binding data were then analyzed and candidate mHAs cataloged.

class I is 9 amino acids. Therefore, for each polymorphism, ANNOVAR returned 8 amino acids on either side of the nsSNP$_{GVH}$-encoded amino acid, resulting in a 17-mer peptide. This effectively generated nine nonamers from each nsSNP$_{GVH}$-encoded polymorphism; thus, the resulting peptides would have the polymorphic amino acid at positions 1 to 9, from the C- to the N-terminal position (**Figure 1**).

IN SILICO VARIANT PEPTIDE-HLA BINDING AFFINITY DETERMINATION

The 17-mer peptides generated by ANNOVAR resulting from the nsSNP$_{GVH}$ were analyzed by the IEDB-MHC I-peptide binding prediction tools version 2.9.1, downloaded from (http://tools.immuneepitope.org/analyze/html_mhcibinding20090901B/download_mhc_I_binding.html). Nine oligopeptides were created for each 17-mer peptide using a 9-mer sliding window. The binding affinity of each of these 9-mers to the patient-specific HLA-A, HLA-B, and HLA-C (**Table 1**) were determined by running each 9-mer independently through the IEDB-MHC I prediction software. The output of this iterative process included variables, such as, the gene name and coordinates, the polymorphic peptide sequence, and the calculated IC50 value via the SMM algorithm (a partial example of output in Table S1 in Supplementary Material). IC50 values in nano-Molar (nM) represent the concentration of the test peptide, which will displace 50% of a standard peptide from the HLA molecule in question. The lower the IC50 for a peptide, the stronger the binding affinity of that peptide for the HLA in question. The cutoff in our analysis to classify a putative peptide as being *presented* by HLA, is an IC50 of $<$500 nM (intermediate affinity binding; http://tools.immuneepitope.org/mhci/help/). Those peptides that bound to HLA with an IC50 of $<$50 nM were designated *strongly presented* (high affinity binding).

To validate the findings from the SMM algorithm, the ANNOVAR generated 17-mer peptide libraries were next interrogated using the NetMHCpan software (http://www.cbs.dtu.dk/services/NetMHCpan/). To accomplish this, two software programs were developed to analyze the peptide data and query NetMHCpan remotely. The first program sequentially sent packets of 30 protein sequences to NetMHCpan. The protein sequences were sent in order by patient and HLA, and a sliding 9-mer window was selected to interrogate HLA binding, similar to SMM IEDB algorithm. NetMHC then returned *html* results, which were then stored on the local server. The second program examined the returned *html* results and organized it in a comma-separated-value (.csv) file, which could then be opened in Microsoft Excel for further analysis.

Results from the SMM IEDB algorithm and NetMHCpan were compared in each DRP by HLA loci and polymorphic peptides. Specifically, HLA locus and polymorphic peptide were combined to make a single variable within each patient dataset, allowing for the removal of duplicate peptides and identification of unique polymorphic peptides found by both or one methods. Presented and strongly presented polymorphic peptides were compared between the two methods, and then combined to get a comprehensive list of unique polymorphic peptide-HLA complexes for each patient.

DERIVING HLA-SPECIFIC ALLOREACTIVITY POTENTIAL

Given the large number of peptides strongly binding HLA identified in each DRP, area under the curve for the IC50 of the strongly binding peptides was determined to summarize the data. The peptide-HLA IC50s were plotted in an ascending order (descending order of affinity). First the non-linear distribution function of the peptides up to an IC50 of 100 nM was computed (a polynomial function of the second order). To obtain the area under the curve depicting the peptide-HLA complexes and their corresponding dissociation constants, the definite integral of the curve was determined. The definite integral by definition is the area of the x–y plane bounded by the curve Eq. (1),

$$\int_a^b f(x)\,dx \tag{1}$$

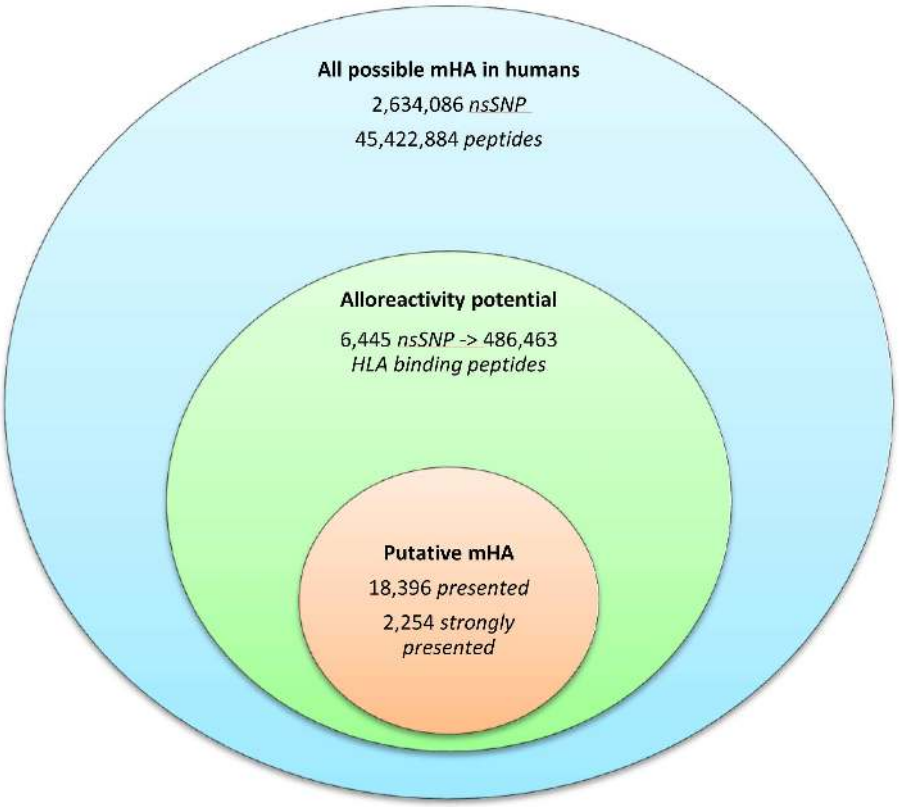

FIGURE 2 | The burden of minor histoincompatibility in human SCT. **(A)** All possible mHA in human beings: data generated from NCBI dbSNP database (22). **(B)** Alloreactivity potential: the current patient cohort had an average of 6,445 nsSNPs/DRP, which when converted into peptide fragments averaged 486,463 possible mHA/DRP. **(C)** Putative mHA: each DRP had its $nsSNP_{GVH}$-encoded peptides filtered by predicted binding to six HLA class I alleles specific to that DRP. Average number of peptides with binding affinity labeled *presented* (<500 nM), and *strongly presented* (<50 nM) is shown.

where $f(x)$ denotes the function of the curve and a and b are the bounds on the x–axis, i.e., the lowest value of the IC50 recorded and the cutoff chosen.

TISSUE EXPRESSION OF POLYMORPHIC PEPTIDES

Relative gene (and protein) expression level is a critical factor contributing to HLA class I presentation of a peptide derived from the gene (21). To investigate the tissue-specific expression of genes incorporating *presented* peptides, software from the European Bioinformatics Institute, Illumina Body Map, (http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-513/) was used to correlate *presented* peptides from the peptide library with relative gene expression in different tissues represented in this software.

RESULTS

CREATION OF POLYMORPHIC PEPTIDES

Whole exomes of 9 SCT DRP were sequenced, identifying an average of 6,445 nsSNP between donors and recipients. To determine the nsSNP that would be associated with possible mHA, peptide sequences were generated that incorporated the polymorphic amino acid at each position 1 to 9, in non-americ peptides using the ANNOVAR software. Theoretically, this could yield nine different peptides for each SNP (**Figure 1**). However, a nsSNP near either the 3′ or 5′ end of a sequence of a gene (N or C terminus of a protein) would lead to fewer peptides. The ANNOVAR output yielded on average 486,463 potential peptides encoded by nsSNPs and presented by the six HLA molecules in these patients (range: 1,043,514-366,426 peptides/DRP). This output was generally greater than the calculated possibilities since it also included peptides resulting from splice variants of the various proteins bearing SNP encoded amino acids. In all, these peptides constituted the total pool of variant peptides, which may be immunogenic in a DRP (**Figure 2**).

HLA-SPECIFIC ALLOREACTIVITY POTENTIAL

The 9-mer peptides bearing the polymorphic amino acid, with a GVH vector were then analyzed for their binding affinities to the individual HLA class I in each patient to determine the variant peptides potentially presented to the donor T-cells. The IEDB-SMM HLA class I binding prediction algorithm was utilized to calculate the binding affinity of the peptide output from ANNOVAR, and to rank putative mHA for their ability to be presented by individual HLA. After filtering for splice variants and duplicate peptide representation in the dataset, a median of 18,396 (range: 1,926–72,294) peptides were identified that bound HLA-A, -B, and -C with an intermediate affinity (IC50 < 500 nM) in the nine DRP,

and were designated as *presented*. Further, a median 2,254 (177–21,548) peptides were predicted to bind MHC class I with a high affinity (IC50 < 50 nM) and were designated as *strongly presented* (**Figure 2**). When separated by the donor type (MRD, $n = 4$, versus URD, $n = 5$), the HLA-matched unrelated DRPs had a significantly higher number of both *presented* and *strongly presented* peptides as determined by IEDB SMM algorithm ($P = 0.016$; Mann–Whitney U test) (**Figure 3**). The difference in the number of presented peptides between unrelated and related donors corroborated the large alloreactivity potential identified earlier in these donor types by whole exome sequencing (14).

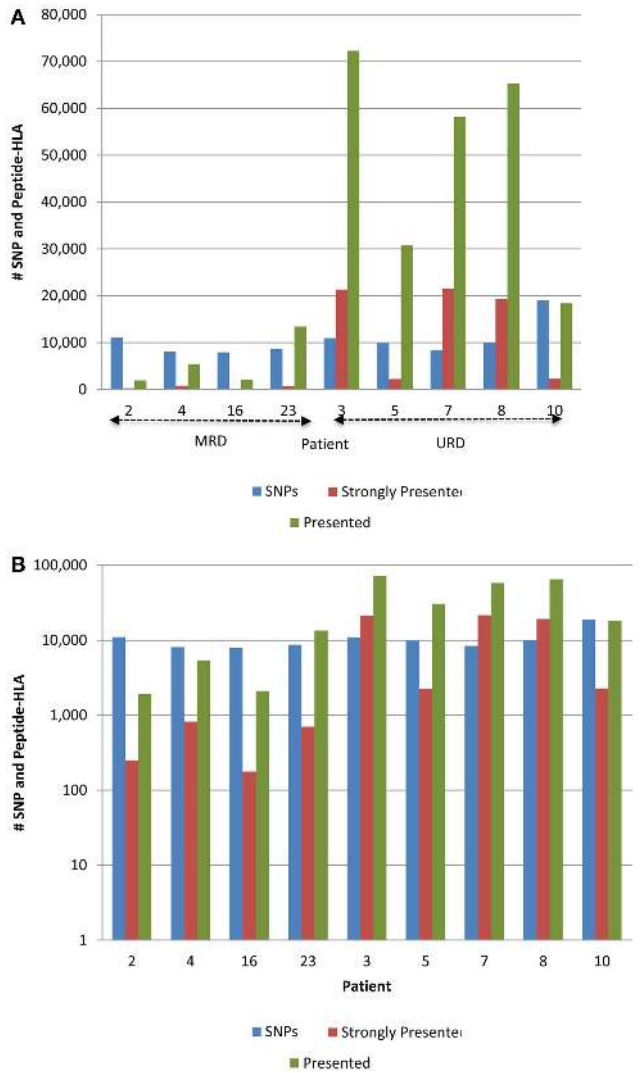

FIGURE 3 | Whole exome sequence variation and resulting HLA-binding oligopeptides in MRD and URD. **(A)** Number of nsSNP, and the resulting *presented* (IC50 < 500 nM) and *strongly presented* (IC50 < 50 nM) peptides (GVH vector) presented by the HLA in each patient. **(B)** Same data as in **(A)**, presented with the *y*-axis changed to log-scale to better illustrate the SNP to HLA-binding peptide ratio between MRD and URD. Significant difference observed in the distribution of SMM-IEDB predicted presented and strongly presented peptides between MRD and URD. Patients 2, 4, 16, 23 – MRD; patients 3, 5, 7, 8, 10 – URD SCT recipients.

To summarize the mass of information regarding the numerous HLA-binding peptides and their binding affinities, the peptides were ranked according to their binding affinity, that is, the IC50 values, and the distribution of their binding affinities was determined (**Figure 4**). For the analysis reported here, this operation was performed without filtering duplicate peptide-HLA complexes resulting from splice variants. Area under the curve (AUC; nM•Peptide) for each DRP was then computed for peptides with an IC50 up to 100 nM. Once again, marked differences were observed in the calculated AUC between MRD and URD (**Table 2**). This summarized measure hypothetically represents a *HLA-specific alloreactivity potential* for each unique DRP, and may be considered as an example of the cumulative mHA differences observed between the HLA-matched donors and recipients.

In a further analysis, when the reciprocal of the IC50 for each peptide (a more direct numerical reflection of the binding affinity) was plotted for each peptide, a Power distribution

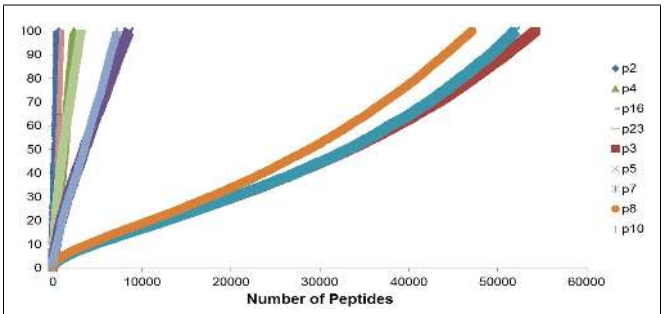

FIGURE 4 | Peptide-HLA complexes with IC50 values up to 100 nM plotted in descending order of binding affinity. Depicting difference in the number of peptide-HLA complexes (*x*-axis) and their IC50 values (*y*-axis), for each DRP. Lower IC50 values correspond to greater binding affinity between putative peptide and relevant HLA. IC50 distribution is non-linear and described as a polynomial function of the second order, forming a continuum. Marked difference observed between MRD and URD (see **Table 2**) for the AUC calculated from these plots.

Table 2 | HLA-specific alloreactivity potential.

Patient	AUC (nM.Peptide)
2	$0.0361*10^6$
4	$0.1191*10^6$
16	$0.0417*10^6$
23	$0.1906*10^6$
3	$2.5802*10^6$
5	$0.4751*10^6$
7	$2.2249*10^6$
8	$1.9886*10^6$
10	$0.3754*10^6$

*All the peptides with an SMM-IC50 of < 100 nM were plotted in order of ascending IC50, and the area under the curve for the resulting graph for each patient was determined (Formula 1). This value represents a summary measure of the number of peptides with a high binding affinity and their binding affinities in each DRP and is described by the dimensionless unit, nM.Peptide. See **Figure 4** for the individual plots. Unrelated DRP are shaded gray.*

Table 3 | Number of *presented* and *strongly presented* peptides predicted by the IEDB SMM and NetMHCpan algorithms.

Donor–recipient pair	Number of $nsSNP_{GVH}$	SMM *presented* peptide-HLA	SMM *strongly presented* peptide-HLA	NetMHC *presented* peptide-HLA	NetMHC *strongly presented* peptide-HLA	Shared *presented* peptide-HLA
2	4,446	1,926	250	3,883	1,376	1,332
4	4,448	5,412	825	3,962	885	2,441
16	3,290	2,111	177	1,071	427	417
23	3,657	13,456	705	787	118	534
3	7,227	72,294	21,339	7,242	2,509	4,881
5	6,572	30,730	2,254	2,759	538	1,865
7	6,725	58,209	21,548	5,231	2,178	2,931
8	6,573	65,298	19,275	4,831	2,000	2,445
10	9,203	18,396	2,283	5,002	989	2,065

The number of unique peptide-HLA complexes identified in silico for each donor-recipient pair. Last column represents number of unique peptides predicted to bind the relevant HLA by both algorithms. Unrelated DRP are shaded gray. Presented (intermediate affinity HLA binding) and strongly presented (strong affinity HLA binding) peptide-HLA complexes have IC50 of <500 and <50 nM, respectively.

was observed, analogous to T-cell clonal frequency distribution previously reported (Figure S1 in Supplementary Material) (23).

VERIFYING HLA BINDING AFFINITY OF THE VARIANT PEPTIDE LIBRARY IN UNIQUE DRP

To confirm the IEDB-SMM algorithm findings, a second peptide-HLA binding affinity prediction tool, NetMHCpan was used to interrogate the variant peptide libraries from the unique DRP and its output compared with the IEDB SMM. The NetMHCpan yielded a median of 3,962 peptides categorized as *presented* and 989 peptides as *strongly presented* in the nine DRP studied (MRD versus URD, $P = 0.063$ and 0.11, respectively, Mann–Whitney U test) (**Table 3**). The IEDB-SMM and NetMHCpan datasets were then combined and unique peptide-HLA complexes predicted to be presented by both algorithms determined (*shared* peptides). The median number of *shared* unique peptides presented/DRP was 2,065 (range: 417–4,881) (**Table 3**). A representative data table depicting peptide sequences and respective IC50 values for binding to a single HLA locus, in a patient, predicted by both algorithms is given in Table S1 in Supplementary Material. Plotting the IC50 of unique *presented* peptide-HLA complexes derived utilizing both algorithms, demonstrated not only a very large number of complexes identified by both algorithms, but also that a large proportion of these complexes were categorized as *strongly presented* (**Figure 5**). Furthermore, a weak, but significant correlation was identified between the IC50 predictions for both the algorithms in the shared peptide-HLA complex datasets ($N = 9$, median Pearson's correlation coefficient $R = 0.62$, $P < 0.01$). Additionally, when the distribution of peptides presented on the three class I HLA loci was examined, no discernable preference for particular HLA loci was observed in terms of likelihood of peptide presentation (Figure S2A,B in Supplementary Material), except for a possible HLA-C dominance in URD recipients in the SMM algorithm.

TISSUE DISTRIBUTION OF PEPTIDES

For a peptide to be relevant in terms of its contribution to GVHD risk, in addition to its potential for presentation on the relevant HLA in a specific DRP, the relevant protein needs to be expressed in the tissues. When the putative mHA (presented peptides, IC50 < 500 nM) were cataloged, according to the tissue-specific expression of the genes they were derived from, most organ systems had genes with potential mHA (**Figure 6**). Further, although several antigens are expressed in organs, such as, colon, liver, and lungs, frequent target organs for GVHD; a large number of genes bearing potentially antigenic peptides are also expressed in other organ systems such as the kidney and adipose tissue seldom targeted by GVHD (Table S2 in Supplementary Material).

DISCUSSION

Allogeneic SCT represents a unique model system to study donor T-cell responses to neo-antigens encountered in the recipient. However, clinical transplantation is characterized by a vast repertoire of variant antigens, which in theory would result in a complex expansion of the T-cell repertoire (24, 25). The findings reported here provide a direct estimate of the antigenic variation, which may be encountered by the donor cytotoxic T-cell (CTL) populations following SCT. Starting from nsSNPs in the exomes of donors and recipients, the reported analysis determined the resulting variant nonameric peptides and gave an *in silico* estimate of the binding affinity (reflected by the IC50) of these peptides to the relevant HLA in the transplant recipients. The existence of this very large library of immunogenic peptides in HLA-matched DRP, immediately raises the question as to why only some and not all the patients develop GVHD.

If all the peptides in this large library of potential mHA were presented to non-tolerant T-cells, then GVHD would potentially develop in all SCT patients, particularly with URD, where the magnitude of immunogenic peptides is considerably larger than MRD. Supporting this notion is the observation that development of extensive chronic GVHD in patients is relatively common when conventional immunosuppressive regimens are used. Further, our findings offer a possible explanation for why most patients develop GVHD, despite having HLA identical donors, and do so more frequently when the donors are unrelated (26, 27). Alternatively, the large magnitude of mHA between HLA-matched donors also gives an insight into why patients undergoing HLA-mismatched transplants such as haploidentical or mismatched URD transplants have

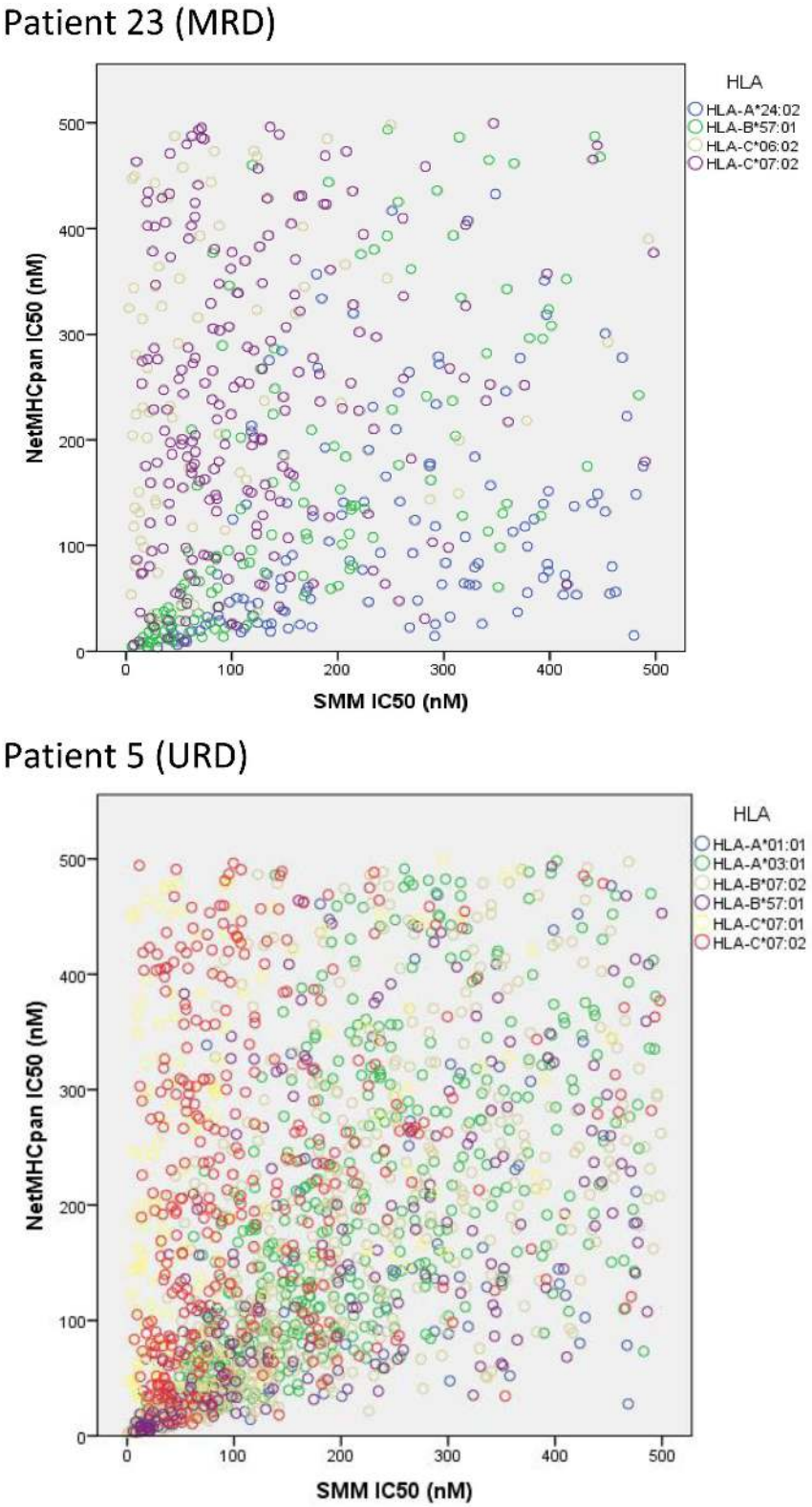

FIGURE 5 | Unique peptide-HLA complexes (GVH vector) with IC50 < 500 nM predicted by both SMM and NetMHCpan. Scatter plots depict the IC50 for unique polymorphic peptide-HLA complexes predicted by the two different algorithms studied. Each circle corresponds to a unique peptide-HLA complex, with color depicting specific HLA. A large number of patient-HLA-specific strong-binding peptides identified by both programs, using SNP data derived from exome sequencing. Only *shared* peptide-HLA complexes predicted to have an IC50 < 500 nM by both algorithms included.

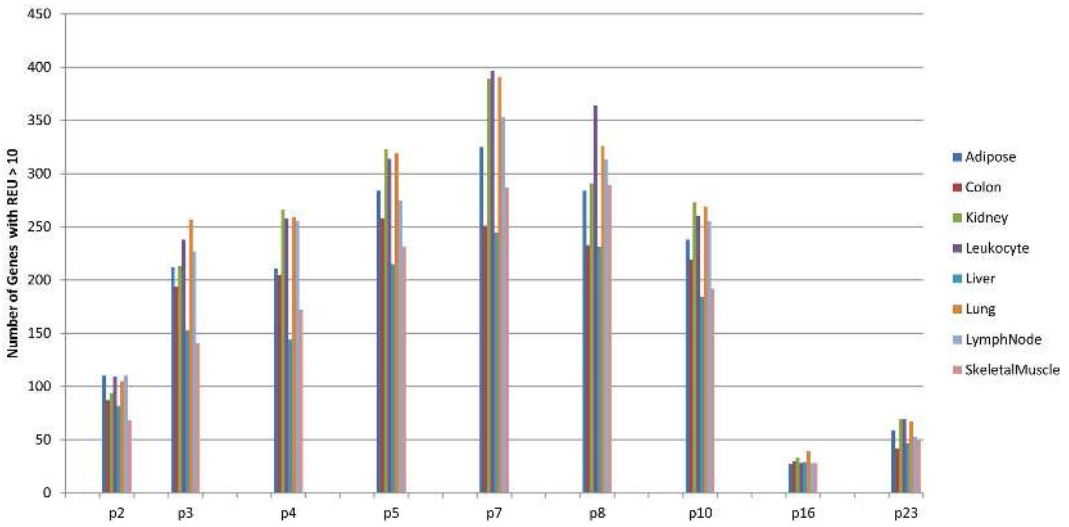

FIGURE 6 | Tissue distribution of presented mHA with gene expression. Number of genes coding for mHA (IC50 < 500 nM by SMM algorithm) and expressed at a relative expression unit (REU) value > 10. European Bioinformatics Institute Illumina Body Map was used to correlate presented peptides with relative gene expression in 16 tissues. Several hundred genes per organ expressed have $nsSNP_{GVH}$, which may generate HLA binding peptides (SMM IEDB dataset).

clinical outcomes, which are not dramatically different from those of HLA-MR donors, that is, if appropriate GVHD prophylaxis is used in the first few weeks of the transplant (28, 29). This paradox may be understood, if one considers the mHA as the *targets* for GVHD and HLA as the *mediators* of this phenomenon. Thus, if the number of *targets* is relatively similar in HLA-matched and haploidentical-related donor, and in the HLA-matched and -mismatched URD transplant recipients; the difference introduced by HLA mismatching is overcome by adjustments in the GVHD prophylaxis regimens. One may postulate that even though thousands of immunogenic peptides are present, the conditions at the time of transplantation determine eventual outcome following transplant, that is, whether tolerance will develop or GVHD ensue following the initial interaction between recipient mHA-HLA complexes and donor T-cells. As an example, when the proteasome inhibitor bortezomib is added to the conditioning regimen, by inhibiting peptide generation and consequently diminishing antigen presentation to donor T-cells in the very first weeks of the transplant, it reduces the risk of GVHD in URD SCT (6). If the model outlined above is correct, then the enormous magnitude of immunogenic peptides constituting the HLA-specific alloreactivity potential will constitute an antigenic "pressure" upon the non-tolerant donor T-cells when first encountered, influencing the evolving T-cell repertoire following SCT. This antigenic pressure may be mitigated by agents, which influence either antigen presentation (e.g., bortezomib) or the T-cell response (e.g., anti-thymocyte globulin, calcinuerin inhibitors, mycophenolate mofetil, post-transplant cyclophosphamide). An observation from this dataset that supports this hypothesis is that the frequency distribution of the binding affinities of the peptides to the HLA molecules follows the Power law (Figure S1 in Supplementary Material). This frequency distribution is similar to the T-cell clonal frequency distribution observed when T-cell clonality is measured using high-throughput T-cell receptor β sequencing (23). This suggests that the T-cell repertoire and clonal frequency emerging after SCT may be proportional to the antigenic peptide-HLA binding affinities. Thus, peptides strongly bound to the HLA will elicit a strong T-cell clonal response, if they engage a T-cell receptor and appropriate co-stimulation is provided. And since the peptide antigen binding affinities form a continuum, rather than discrete clusters of high and low affinity, the T-cell repertoire frequency similarly forms a continuum, described by the Power law. Another conclusion to be considered from the non-discrete distribution of peptide-HLA binding affinity is that other non-recipient derived antigens, such as pathogen-associated peptides may also lie on this continuum. This may result in *cross-reactivity* between autologous antigens and pathogen-associated peptides (30). A manifestation of this in the transplant setting is the triggering of GVHD or graft rejection events by viral infections, such as cytomegalovirus or human herpes virus 6 virus infections (31, 32).

Can these findings be used to develop a clinically relevant model for allogeneic SCT? One possible explanation of the variant outcomes following SCT is that post-transplant emergent T-cell clones either develop tolerance to the many antigens encountered or fail to do so depending on the milieu encountered in the host. Early interventions, such as administration of anti-thymocyte globulin, (33) bortezomib, or post-transplant cyclophosphamide have a large impact on late post-transplant outcomes. Similar tolerance induction is observed following cellular interventions such as regulatory T-cell infusion and conditioning, which up regulates NK-T-cells at the time of SCT (34). This suggest that if a large antigenic pressure from the HLA-specific alloreactivity potential exists in all patients, then tissue injury and cytokine milieu at the time of SCT may be influential in determining the development of GVHD. Thus, if there is tissue injury following SCT, even if it is *sub-clinical*, multiple antigens are presented, then in the absence of

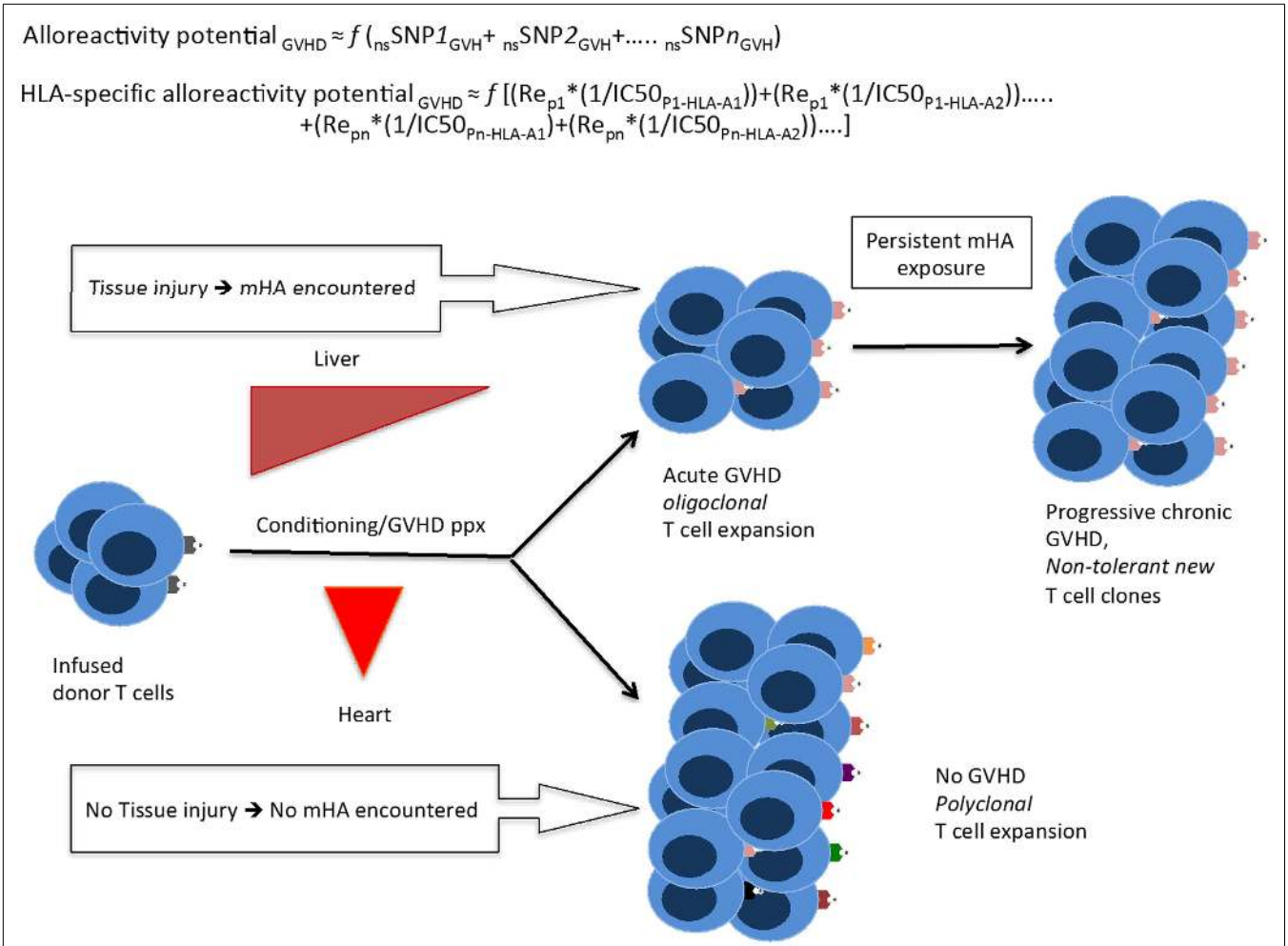

FIGURE 7 | A quantitative model for the development of GVHD. Whole exome sequencing identifies all the nsSNP with a GVH vector, yielding a putative *alloreactivity potential*, which may be a function (*f*) of the cumulative influence of these polymorphisms. This is represented as a series, listing the sequence of polymorphic exome loci. Substituting individual $nsSNP_{GVH}$ in the equation by peptide-HLA *binding affinity* (reciprocal of IC50)*relative expression level of the gene bearing the $nsSNP_{GVH}$ (for each HLA molecule) yields the HLA-specific alloreactivity potential, in this Re is the relative expression of protein with $nsSNP_{GVH}$ and resulting peptides. In this series, the expression, $Re_{p1}*(1/IC50_{P1\text{-}HLA\text{-}A1})$ for each specific peptide-HLA complex, hypothetically represents the T-cell clone-specific AP. Multiple peptides constituting this series then drive a proportional oligoclonal T-cell expansion in GVHD, as many different mHA are presented by the HLA in an individual, the final distribution conforming to the Power law. Since T-cell clonal expansion in response to presented antigens may be influenced by factors such as tissue injury, cytokine milieu, and immunosuppression intensity; the GVHD likelihood, and its phenotype may in turn be determined not only by the ubiquitous mHA but also by the tissue volume and its state (inflammation/injury), and most importantly time at which organ injury/inflammation occurs relative to T-cell infusion.

adequate immunosuppression, the T-cell repertoire that develops results in the development of GVHD. On the other hand, if tissue injury is minimized and there is adequate immunosuppression, when the initial T-cell antigen-presenting cell interactions take place, peripheral (or central) tolerance would emerge. Following that, depending on the presence or absence of thymic tissue, T-cell clones developing from infused stem cells may perpetuate this process based on the prevailing T-cell population and target-tissue antigen presentation, perhaps influenced by the state of tissue injury (**Figure 7**). In such a model, inflammation provoked by the acute GVHD initiated by infused donor-derived T-cells reacting to recipient antigens is perpetuated in the form of "auto-reactivity" by the T-cells, developing from infused stem cells in the absence of normal thymic processing. This concept may not be novel in itself; however, our model provides a biologically plausible explanation reconciling mHA differences observed in HLA-matched DRP.

Correlating the variant peptides with tissue protein expression levels, in our dataset, the immunogenic peptides appear to be uniformly distributed in the major organ systems of the body. This raises the following question: why do solid organ transplant recipients develop rejection, but GVHD does not commonly affect most such organs, such as the kidney and heart? The data presented in this paper suggest a possible answer to this question if the above quantitative model of immunobiology of transplantation is considered. Hypothetically, in the days following SCT, when the infused donor T-cells encounter widespread variant

immunogenic recipient antigens in *inflamed* tissues with a large tissue interface for T-cell antigen-presenting cell interaction, i.e., skin, GI mucosa, liver, and lungs, there is a corresponding polyclonal T-cell allo-immune response, which may result in GVHD affecting the targeted organs. In contrast, the relatively smaller tissue interface in the absence of direct injury, in organs such as the heart and kidney, do not trigger an immunogenic response in the face of an ongoing, *competing* oligoclonal T-cell response elicited by the larger organ systems with injury. When solid organ transplantation is performed, tissue injury even if sub-clinical, in the transplanted organ resulting from the transplant procedure serves as the injury stimulus triggering graft rejection. Based on these data, a theoretical model has been proposed to investigate the notion of alloreactivity potential and its relationship with GVHD onset and propagation over time as in a "chaotic dynamical system" (35).

A potential therapeutic application of this analysis would be the ability to "titrate" the intensity of immunosuppressive therapy in the peri-transplant period based on the magnitude of the HLA-specific alloreactivity potential. This study supports the need for intensive immunosuppression in patients undergoing URD allogeneic SCT, making this algorithm a useful analysis for treatment planning (36). For example, if a patient has a high number of predicted mHA and these are over-represented in lung tissue, therapies can be specifically tailored for that patient and symptoms of lung GVHD treated more promptly. Alternatively, large-scale protein expression studies by Ponten et al. concluded that most proteins are expressed in most tissues, although in varying quantities (37). This raises the question of which parameter plays a larger role in peptide presentation by MHC class I HLA: the absolute molar amount of protein expressed in a tissue, or the binding affinity for a particular peptide; in theory, it may be a combination of the two (**Figure 7**).

As with any *in silico* work, this work can only be considered preliminary and the peptide-HLA class I combinations predicted in our work, will need experimental verification. Acknowledging this limitation, it should be noted that the accuracy of these algorithms has been reviewed and they have been found to be useful predictors of HLA presentation. A similar large number of peptides binding HLA in EBV-transformed B cell lines have been identified when directly characterizing the "ligandome" presented by these cells (38). Further, in a vaccinia virus challenge mouse model, the NetMHC algorithm was able to predict epitopes responsible for 95% of the CTL response with an IC50 threshold of <500 nM (39). Similarly, Armistead et al. found that with an IC50 threshold of <500 nM, all peptides predicted by SMM-IEDB algorithm bound HLA-A 0201 in their assays (40). To put our data in context, a database from all known nsSNPs that had been deposited in NCBI's dbSNP database is presented in **Figure 2** and is labeled as all possible mHA in human beings (22, 41). In light of these findings, it is not at all surprising that we find a large library of immunogenic mHA in each DRP, and there may exist a similar alloreactivity potential mediated by HLA class II.

In conclusion, the findings reported here demonstrate that whole exome sequencing, followed by *in silico* peptide generation and HLA binding affinity determination reveal a large and previously unmeasured *HLA-specific* alloreactivity potential. This potential is predictably larger in patients undergoing URD SCT and mirrors previously described T-cell clonal frequency distribution. We posit that these methodologies may be used to develop mathematical models to better understand the immunopathology of SCT from both HLA-matched and mismatched donors and may in the future allow more precise titration of the immunosuppression intensity in individual transplant recipients.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Massey Cancer Center Pilot Project Grant, JUP 11-0.9, and Virginia's Commonwealth Health Research Board Award #236-11-13 for funding. The authors also gratefully acknowledge Dr. Jamie Teer (Moffitt Cancer Center, Tampa, FL) for his helpful suggestions, critical in determining peptide sequence from exome sequence variation, and Ms. Cheryl Jacocks-Terrell, MS for her assistance with preparing the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Journal/10.3389/fimmu.2014.00529/abstract

REFERENCES

1. Arora M, Weisdorf DJ, Spellman SR, Haagenson MD, Klein JP, Hurley CK. HLA-identical sibling compared with 8/8 matched and mismatched unrelated donor bone marrow transplant for chronic phase chronic myeloid leukemia. *J Clin Oncol* (2009) **27**(10):1644–52. doi:10.1200/JCO.2008.18.7740
2. Toor AA, Sabo RT, Chung HM, Roberts C, Manjili RH, Song S, et al. Favorable outcomes in patients with high donor-derived T cell count after in vivo T cell-depleted reduced-intensity allogeneic stem cell transplantation. *Biol Blood Marrow Transplant* (2012) **18**(5):794–794. doi:10.1016/j.bbmt.2011.10.011
3. Weisdorf DJ, Nelson G, Lee SJ, Haagenson M, Spellman S, Antin JH, et al. Sibling versus unrelated donor allogeneic hematopoietic cell transplantation for chronic myelogenous leukemia: refined HLA matching reveals more graft-versus-host disease but not less relapse. *Biol Blood Marrow Transplant* (2009) **15**(11):1475–8. doi:10.1016/j.bbmt.2009.06.016
4. Valcarcel D, Sierra J, Wang T, Kan F, Gupta V, Hale GA, et al. One-antigen mismatched related versus HLA-matched unrelated donor hematopoietic stem cell transplantation in adults with acute leukemia: center for international blood and marrow transplant research results in the era of molecular HLA typing. *Biol Blood Marrow Transplant* (2011) **17**(5):640–8. doi:10.1016/j.bbmt.2010.07.022
5. Brunstein CG, Fuchs EJ, Carter SL, Karanes C, Costa LJ, Wu J, et al. Alternative donor transplantation after reduced intensity conditioning: results of parallel phase 2 trials using partially HLA-mismatched related bone marrow or unrelated double umbilical cord blood grafts. *Blood* (2011) **118**(2):282–8. doi:10.1182/blood-2011-03-344853
6. Koreth J, Stevenson KE, Kim HT, McDonough SM, Bindra B, Armand P, et al. Bortezomib-based graft-versus-host disease prophylaxis in HLA-mismatched unrelated donor transplantation. *J Clin Oncol* (2012) **30**(26):3202–8. doi:10.1200/JCO.2012.42.0984
7. Shlomchik WD. Graft-versus-host disease. *Nat Rev Immunol* (2007) **7**(5):340–52. doi:10.1038/nri2000
8. Warren EH, Deeg HJ. Dissecting graft-versus-leukemia from graft-versus-host disease using novel strategies. *Tissue Antigens* (2013) **81**(4):183–93. doi:10.1111/tan.12090
9. Mullally A, Ritz J. Beyond HLA: the significance of genomic variation for allogeneic hematopoietic stem cell transplantation. *Blood* (2007) **109**(4):1355–62. doi:10.1182/blood-2006-06-030858
10. Spierings E, Kim Y, Hendriks M, Borst E, Sergeant R, Canossi A, et al. Multicenter analyses demonstrate significant clinical effects of minor histocompatibility antigens on GvHD and GvL after HLA-matched related and unrelated hematopoietic stem cell transplantation. *Biol Blood Marrow Transplant* (2013) **19**(8):1244–53. doi:10.1016/j.bbmt.2013.06.001

11. Warren EH, Zhang XC, Li S, Fan W, Storer BE, Chien JW, et al. Effect of MHC and non-MHC donor/recipient genetic disparity on the outcome of allogeneic HCT. *Blood* (2012) **120**(14):2796–2796. doi:10.1182/blood-2012-04-347286
12. Spierings E, Hendriks M, Absi L, Canossi A, Chhaya S, Crowley J, et al. Phenotype frequencies of autosomal minor histocompatibility antigens display significant differences among populations. *PLoS Genet* (2007) **3**(6):e103. doi:10.1371/journal.pgen.0030103
13. Spellman S, Warden MB, Haagenson M, Piet BC, Goulmy E, Warren EH, et al. Effects of mismatching for minor histocompatibility antigens on clinical outcomes in HLA-matched, unrelated hematopoietic stem cell transplants. *Biol Blood Marrow Transplant* (2009) **15**(7):856–63. doi:10.1016/j.bbmt.2009.03.018
14. Sampson JK, Sheth NU, Koparde VN, Scalora AF, Serrano MG, Lee V, et al. Whole exome sequencing to estimate alloreactivity potential between donors and recipients in stem cell transplantation. *Br J Haematol* (2014) **166**(4):566–70. doi:10.1111/bjh.12898
15. Lundegaard C, Lund O, Buus S, Nielsen M. Major histocompatibility complex class I binding predictions as a tool in epitope discovery. *Immunology* (2010) **130**(3):309–18. doi:10.1111/j.1365-2567.2010.03300.x
16. Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, Nielsin M. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. *Nucleic Acids Res* (2008) **36**:W509–12. doi:10.1093/nar/gkn202
17. Peters B, Sette A. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinformatics* (2005) **6**:132. doi:10.1186/1471-2105-6-132
18. Nielsen M, Lundegaard C, Blicher T, Lamberth K, Harndahl M, Justesen S, et al. NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and-B locus protein of known sequence. *PLoS One* (2007) **2**(8):e796. doi:10.1371/journal.pone.0000796
19. Zhang H, Lundegaard C, Nielsen M. Pan-specific MHC class I predictors: a benchmark of HLA class I pan-specific prediction methods. *Bioinformatics* (2009) **25**(1):83–9. doi:10.1093/bioinformatics/btn579
20. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* (2010) **38**(16):e164. doi:10.1093/nar/gkq603
21. Hoof I, van Baarle D, Hildebrand WH, Kesmir C. Proteome sampling by the HLA class I antigen processing pathway. *PLoS Comput Biol* (2012) **8**(5):e1002517. doi:10.1371/journal.pcbi.1002517
22. Nijveen H, Kester MG, Hassan C, Viars A, de Ru AH, de Jager M, et al. HSPVdb-the human short peptide variation database for improved mass spectrometry-based detection of polymorphic HLA-ligands. *Immunogenetics* (2011) **63**(3):143–53. doi:10.1007/s00251-010-0497-1
23. Meier J, Roberts C, Avent K, Hazlett A, Berrie J, Payne K, et al. Fractal organization of the human T cell repertoire in health and after stem cell transplantation. *Biol Blood Marrow Transplant* (2013) **19**(3):366–77. doi:10.1016/j.bbmt.2012.12.004
24. Ofran Y, Kim T, Brusic V, Blake L, Mandrell M, Wu CJ, et al. Diverse patterns of T-cell response against multiple newly identified human Y chromosome-encoded minor histocompatibility epitopes. *Clin Cancer Res* (2010) **16**(5):1642–51. doi:10.1158/1078-0432.CCR-09-2701
25. Berrie JL, Kmieciak M, Sabo RT, Roberts CH, Idowu MO, Mallory K, et al. Distinct oligoclonal T cells are associated with graft versus host disease after stem-cell transplantation. *Transplantation* (2012) **93**(9):949–57. doi:10.1097/TP.0b013e3182497561
26. Flomenberg N, Baxter-Lowe LA, Confer D, Fernandez-Vina M, Filipovich A, Horowitz M, et al. Impact of HLA class I and class II high-resolution matching on outcomes of unrelated donor bone marrow transplantation: HLA-C mismatching is associated with a strong adverse effect on transplantation outcome. *Blood* (2004) **104**(7):1923–30. doi:10.1182/blood-2004-03-0803
27. Saber W, Opie S, Rizzo JD, Zhang M, Horowitz MM, Schriber J. Outcomes after matched unrelated donor versus identical sibling hematopoietic cell transplantation in adults with acute myelogenous leukemia. *Blood* (2012) **119**(17):3908–16. doi:10.1182/blood-2011-09-381699
28. Bashey A, Zhang X, Sizemore CA, Manion K, Brown S, Holland HK, et al. T-cell-replete HLA-haploidentical hematopoietic transplantation for hematologic malignancies using post-transplantation cyclophosphamide results in outcomes equivalent to those of contemporaneous HLA-matched related and unrelated donor transplantation. *J Clin Oncol* (2013) **31**(10):1310–6. doi:10.1200/JCO.2012.44.3523
29. Fuchs EJ. Haploidentical transplantation for hematologic malignancies: where do we stand? *Hematology Am Soc Hematol Educ Program* (2012) **2012**:230–6. doi:10.1182/asheducation-2012.1.230
30. D'Orsogna LJ, Roelen DL, Doxiadis II, Claas FH. TCR cross-reactivity and allorecognition: new insights into the immunogenetics of allorecognition. *Immunogenetics* (2012) **64**(2):77–85. doi:10.1007/s00251-011-0590-0
31. Cainelli F, Vento S. Infections and solid organ transplant rejection: a cause-and-effect relationship? *Lancet Infect Dis* (2002) **2**(9):539–49. doi:10.1016/S1473-3099(02)00370-5
32. Wang L, Dong L, Zhang M, Lu D. Correlations of human herpesvirus 6B and CMV infection with acute GVHD in recipients of allogeneic haematopoietic stem cell transplantation. *Bone Marrow Transplant* (2008) **42**(10):673–7. doi:10.1038/bmt.2008.238
33. Portier DA, Sabo RT, Roberts CH, Fletcher DS, Meier J, Clark WB, et al. Antithymocyte globulin for conditioning in matched unrelated donor hematopoietic cell transplantation provides comparable outcomes to matched related donor recipients. *Bone Marrow Transplant* (2012) **47**(12):1513–9. doi:10.1038/bmt.2012.81
34. Kohrt HE, Turnbull BB, Heydari K, Shizuru JA, Laport GG, Miklos DB, et al. TLI and ATG conditioning with low risk of graft-versus-host disease retains antitumor reactions after allogeneic hematopoietic cell transplantation from related and unrelated donors. *Blood* (2009) **114**(5):1099–1099. doi:10.1182/blood-2009-03-211441
35. Toor AA, JamesonLee M, Kobulnicky JD, Meier J, Roberts CH, Scalora A, et al. Stem cell transplant as a dynamical system: are clinical outcomes deterministic? *Front Immunol* (2014).
36. Feldhahn M, Doennes P, Schubert B, Schilbach K, Rammensee H, Kohlbacher O. miHA-match: computational detection of tissue-specific minor histocompatibility antigens. *J Immunol Methods* (2012) **386**(1–2):94–100. doi:10.1016/j.jim.2012.09.004
37. Ponten F, Gry M, Fagerberg L, Lundberg E, Asplund A, Berglund L, et al. A global view of protein expression in human cells, tissues, and organs. *Mol Syst Biol* (2009) **5**:337. doi:10.1038/msb.2009.93
38. Hassan C, Kester MG, de Ru AH, Hombrink P, Drijfhout JW, Nijveen H, et al. The human leukocyte antigen-presented ligandome of B lymphocytes. *Mol Cell Proteomics* (2013) **12**(7):1829–43. doi:10.1074/mcp.M112.024810
39. Moutaftsi M, Peter B, Pasquetto V, Tscharke DC, Sidney J, Bu H, et al. A consensus epitope prediction approach identifies the breadth of murine TCD8+-cell responses to vaccinia virus. *Nat Biotechnol* (2006) **24**(7):817–9. doi:10.1038/nbt1215
40. Armistead PM, Liang S, Li H, Lu S, Van Bergen CA, Alatrash G. Common minor histocompatibility antigen discovery based upon patient clinical outcomes and genomic data. *PLoS One* (2011) **6**(8):e23217. doi:10.1371/journal.pone.0023217
41. Sherry S, Ward M, Kholodov M, Baker J, Phan L, Smigielski E, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* (2001) **29**(1):308–11. doi:10.1093/nar/29.1.308

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 02 August 2014; accepted: 07 October 2014; published online: 06 November 2014.

Citation: Jameson-Lee M, Koparde V, Griffith P, Scalora AF, Sampson JK, Khalid H, Sheth NU, Batalo M, Serrano MG, Roberts CH, Hess ML, Buck GA, Neale MC, Manjili MH and Toor AA (2014) In silico derivation of HLA-specific alloreactivity potential from whole exome sequencing of stem-cell transplant donors and recipients: understanding the quantitative immunobiology of allogeneic transplantation. Front. Immunol. 5:529. doi: 10.3389/fimmu.2014.00529

This article was submitted to Alloimmunity and Transplantation, a section of the journal Frontiers in Immunology.