Published OnlineFirst November 9, 2010; DOI: 10.1158/0008-5472.CAN-10-1832

Cancer Research

High-Throughput Characterization of 10 New Minor Histocompatibility Antigens by Whole Genome Association Scanning

Cornelis A.M. Van Bergen[1], Caroline E. Rutten[1], Edith D. Van Der Meijden[1], Simone A.P. Van Luxemburg-Heijs[1], Ellie G.A. Lurvink[1], Jeanine J. Houwing-Duistermaat[2], Michel G.D. Kester[1], Arend Mulder[3], Roel Willemze[1], J.H. Frederik Falkenburg[1], and Marieke Griffioen[1]

Abstract

Patients with malignant diseases can be effectively treated with allogeneic hematopoietic stem cell transplantation (allo-SCT). Polymorphic peptides presented in HLA molecules, the so-called minor histocompatibility antigens (MiHA), play a crucial role in antitumor immunity as targets for alloreactive donor T cells. Identification of multiple MiHAs is essential to understand and manipulate the development of clinical responses after allo-SCT. In this study, $CD8^+$ T-cell clones were isolated from leukemia patients who entered complete remission after allo-SCT, and MiHA-specific T-cell clones were efficiently selected for analysis of recognition of a panel of EBV-transformed B cells positive for the HLA restriction elements of the selected T-cell clones. One million single nucleotide polymorphisms (SNP) were determined in the panel cell lines and investigated for matching with the T-cell recognition data by whole genome association scanning (WGAs). Significant association with 12 genomic regions was found, and detailed analysis of genes located within these genomic regions revealed SNP disparities encoding polymorphic peptides in 10 cases. Differential recognition of patient-type, but not donor-type, peptides validated the identification of these MiHAs. Using tetramers, distinct populations of MiHA-specific $CD8^+$ T cells were detected, demonstrating that our WGAs strategy allows high-throughput discovery of relevant targets in antitumor immunity after allo-SCT. *Cancer Res; 70(22); 9073–83. ©2010 AACR.*

Introduction

Allogeneic hematopoietic stem cell transplantation (allo-SCT) is a curative treatment for malignant diseases (1). In case of relapse, persistent disease, or incomplete donor chimerism after transplantation, donor lymphocytes can be administered to initiate or enhance graft-versus-tumor (GVT) reactivity (2). Following HLA-matched allo-SCT, minor histocompatibility antigens (MiHA) are the target structures for donor-derived T cells. Presented in the context of HLA molecules, MiHAs are peptides derived from polymorphic proteins that differ between individuals due to genomic single nucleotide polymorphisms (SNP; refs. 3–6). T-cell responses directed against MiHAs have been monitored in detail in patients responding to donor lymphocyte infusion (DLI; refs. 7, 8). In these patients, detection of circulating MiHA-specific T cells coincided with rapid disappearance of malignant cells, indicating the *in vivo* efficacy of MiHA-specific T cells. Whether donor T cells mediate beneficial GVT effects or harmful graft-versus-host disease (GvHD) is likely to depend on the tissue distribution of the target antigen. Exclusive MiHA expression on malignant cells may result in a selective GVT effect, whereas T cells recognizing broadly expressed MiHAs may lead to detrimental GvHD (9).

Initially, biochemical identification of peptides eluted from HLA class I molecules recognized by MiHA-specific T cells was used for MiHA discovery (10–15). In addition, cDNA libraries were generated and screened for T-cell recognition to identify the MiHA-encoding gene transcript (8, 16–21). Finally, genetic linkage analysis has been applied to define the genomic region encoding the MiHA (22, 23). Irrespective of the MiHA identification method, ultimate validation of each MiHA was provided by the strict correlation between a specific SNP genotype and recognition by the T-cell clone.

SNPs between patient and donor may lead to MiHA expression by different mechanisms. The majority of MiHAs are encoded by "missense" SNPs that directly encode an amino acid polymorphism in the protein. In addition, SNPs

Authors' Affiliations: Departments of [1]Hematology, [2]Medical Statistics and Bioinformatics, and [3]Immunohaematology and Blood Transfusion, Leiden University Medical Center, Leiden, the Netherlands

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

C.A.M. Van Bergen and C.E. Rutten contributed equally to this work.

Corresponding Author: Cornelis A.M. Van Bergen, Department of Hematology, Leiden University Medical Center, P.O. Box 9600, 2300RC Leiden, the Netherlands. Phone: 31-71-5262274; Fax: 31-71-5266755; E-mail: c.a.m.van_bergen@lumc.nl.

doi: 10.1158/0008-5472.CAN-10-1832

Published OnlineFirst November 9, 2010; DOI: 10.1158/0008-5472.CAN-10-1832

"synonymous" in the normal open reading frame (ORF) have been shown to encode polymorphic amino acids in MiHA epitopes upon translation of alternative reading frames (ARF), as previously described for LB-ECGF-1H (8) and LB-ADIR-1F (15). MiHAs can also be generated by protein splicing, as illustrated for HwA-9, which is spliced together in a noncontiguous order by the proteasome (18). Furthermore, MiHAs can be encoded by gene transcripts that differ between patient and donor due to insertion or deletion of a single nucleotide, thereby inducing a shift in the translational reading frame, as shown by the LRH-1 MiHA (23). Finally, MiHAs may be encoded by gene transcripts that are homozygously deleted in the donor, as shown for MiHAs encoded by the *UGT2B17* gene (17), or by intronic SNPs altering gene expression by RNA splicing, as shown for ACC-6 (19).

Although various MiHA identification approaches have led to the successful discovery of 20 HLA class I–restricted MiHAs, their laboriousness and low efficiency hampered identification of the majority of MiHAs recognized by T-cell clones. Development of advanced array-based SNP genotyping, and the fact that MiHAs are encoded by polymorphic genomic regions containing specific SNPs, enabled application of whole genome association scanning (WGAs) as an alternative approach for MiHA discovery (24). Association analysis may identify associating SNPs that directly encode the amino acid polymorphism or may serve as markers for MiHA-encoding SNPs that are not measured by the array. These markers are in linkage disequilibrium with MiHA-encoding SNPs and therefore associate with T-cell recognition. The feasibility of WGAs as an approach to identify new MiHAs has been shown by the discovery of individual MiHAs in BCL2A1 (24), CD19 (25), SLC1A5 (26), UGT2B17 (26), SLC19A1 (27), P2RX7 (28), and DPH1 (28).

To understand the development of GVT reactivity and GvHD, characterization of large numbers of MiHAs is essential. High-throughput identification and expression analysis of MiHAs in patients with various clinical responses after allo-SCT may provide insight into the fundamental difference and may allow manipulation of the balance between GVT reactivity and GvHD. Here, we report the successful application of WGAs for high-throughput characterization of MiHAs in GVT reactivity. T-cell clones were generated from two patients who received an HLA-matched allograft and responded to DLI for recurrence of the malignant disease. A panel of 80 EBV-transformed B-cell lines (EBV-LCL) was tested for recognition by each T-cell clone. More than 1 million SNPs were measured in the EBV-LCL and subsequently searched for association with the EBV-LCL recognition pattern of each T-cell clone. Strong association with a genomic region was found for 12 of the 17 T-cell clones evaluated. For 10 of these 12 T-cell clones, novel MiHAs with balanced population frequencies presented in the common HLA-A*0201 or B*0702 molecules could be identified, illustrating the value and efficiency of WGAs for MiHA discovery. Using tetramer analysis, the kinetics of MiHA-specific T-cell responses in the patients were shown to coincide with the clinical responses that were observed after DLI.

Materials and Methods

Patients

Patient H (HLA-A*0201, B*0702, B*4402, Cw*0501, Cw*0702, DRB1*04, DRB1*15, DRB4*01, DRB5*01, DQB1*03, DQB1*06, DPB1*04) had myelodysplastic syndrome refractory anemia with excess of blasts type 2. She was transplanted with a T-cell–depleted peripheral blood stem cell graft from a male HLA-matched unrelated donor. Chimerism analysis showed a decrease in the percentage of donor cells to 86% during the first 6 months after transplantation; therefore, the patient was treated with DLI. After DLI, the patient developed GvHD limited to skin and mouth, which coincided with a rapid and sustained conversion to 100% donor chimerism. Patient Z (HLA-A*0101, A*0201, B*0702, B*4402, Cw*0501, Cw*0702, DRB1*11, DRB1*15, DQB1*03, DQB1*06, DPB1*04, DPB1*14) is a patient with chronic myelogenous leukemia (CML) transplanted with a T-cell–depleted peripheral blood stem cell graft from his HLA identical sibling. More than 1 year after transplantation, the patient was successfully treated with DLI for cytogenetic relapse of CML. The patient did not develop GvHD.

Cell collection, preparation, and EBV-LCL culture

Peripheral blood and bone marrow samples were collected from the patients, their donors, and third-party individuals after approval by the Leiden University Medical Center Institutional Review Board and informed consent according to the Declaration of Helsinki. Mononuclear cells were isolated by Ficoll gradient centrifugation and cryopreserved. Stable EBV-LCL were generated by *in vitro* transformation of thawed mononuclear cells of selected HLA-A*0201–positive and B*0702-positive individuals using EBV supernatant (29) followed by culture in Iscove's modified Dulbecco's medium (IMDM, Lonza) with 10% fetal bovine serum (FBS, Lonza).

Isolation, expansion, and selection of T-cell clones

Peripheral blood mononuclear cells (PBMC) obtained after DLI were thawed for isolation of T-cell clones. Monocytes were depleted using magnetic CD14 beads (Miltenyi Biotec) according to the manufacturer's instructions. CD14-negative cells were stained with allophycocyanin (APC)-conjugated anti-CD8 antibody (Invitrogen) and phycoerythrin (PE)-conjugated anti–HLA-DR antibody (Invitrogen). Propidium iodide (Sigma-Aldrich) was added to exclude dead cells, and viable HLA-DR$^+$ CD8$^+$ cells were single cell sorted by flow cytometry. Isolated T cells were collected in 96-well U-bottomed plates (Corning) in IMDM with 5% pooled human serum, 5% FBS, interleukin 2 (IL-2; 100 IU/mL), phytohemagglutinin (0.8 μg/mL, Murex Biotec Limited), and 5×10^4 irradiated allogeneic PBMC as feeder cells. Proliferating T-cell clones were restimulated every 10 to 14 days under identical stimulatory conditions at 2×10^5

Published OnlineFirst November 9, 2010; DOI: 10.1158/0008-5472.CAN-10-1832

T cells and 1×10^6 feeder cells per milliliter of culture medium. T-cell (5×10^3 per well) recognition of patient- and donor-derived EBV-LCL as well as third-party EBV-LCL (3×10^4 per well) were tested in 96-well U-bottomed plates in 100 to 150 μL IMDM with 5% pooled human serum, 5% FBS, and 20 IU/mL of IL-2. After overnight coincubation, release of IFN-γ was measured in culture supernatants by ELISA according to the manufacturer's instructions (Sanquin). For blocking studies, EBV-LCL were preincubated with saturating concentrations of antibodies against HLA-A*02 (BB7.2) and HLA-B*07 (VTM3A1) for 30 minutes at room temperature before addition of T cells.

T-cell receptor-β variable chain analysis

T-cell receptor-β variable chain (TCRBV) usage was investigated by flow cytometry using specific monoclonal antibodies as included in the TCRBV repertoire kit (Beckman Coulter). T-cell clones with identical TCRBVs were analyzed by DNA sequencing of the CDR3 region of the TCRBV after amplification by PCR using specific forward and reverse primers. TCRBV products were purified with the Qiaquick PCR purification kit (Qiagen) and sequenced using the dye terminator cycle sequencing kit (ABI-PRISM, Perkin-Elmer), according to the manufacturer's instructions.

Whole genome association scanning

A panel of 80 HLA-A*0201– and B*0702-expressing EBV-LCL was selected for WGAs. To allow simultaneous analysis of all EBV-LCL for recognition by T-cell clones, triplicate samples of EBV-LCL were dispensed at 6×10^4 cells in 50 μL per well in 96-well polypropylene U-bottomed plates (Greiner Bio-One). After addition of 50 μL medium containing 20% DMSO, the panel plates were sealed, frozen, and stored at −140°C. To test the recognition of the EBV-LCL panel by each individual T-cell clone, plates were thawed, washed, and incubated for 2 days, and 5×10^3 MiHA-specific T cells were added to each well. After 24 hours, supernatants were used for IFN-γ ELISA. Of each EBV-LCL, genomic DNA from 5×10^6 cells was isolated using Gentra Puregene Cell Kit (Qiagen). According to the manufacturer's instructions,

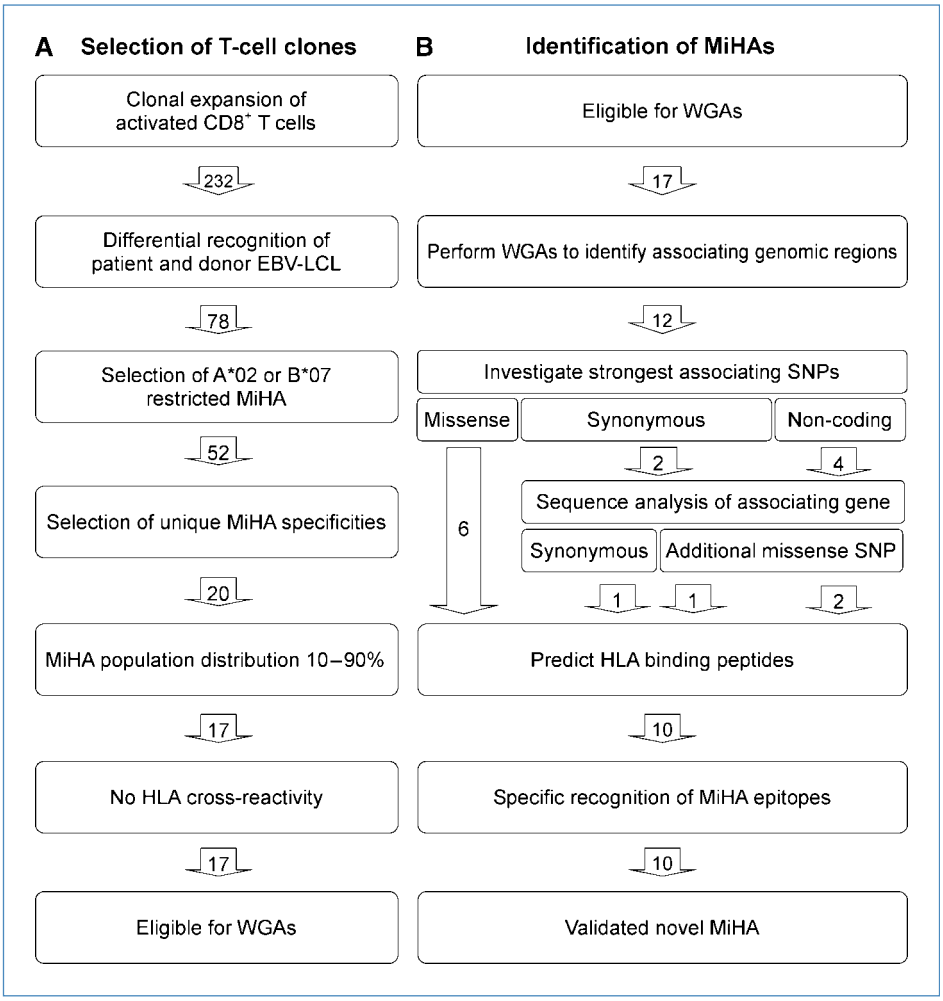

Figure 1. T-cell clone selection and WGAs strategy for MiHA identification. Strategy outline for selection of CD8+ T-cell clones (A) and MiHA identification by WGAs (B). The digits in the arrows depict the number of clones that was selected at each step.

Published OnlineFirst November 9, 2010; DOI: 10.1158/0008-5472.CAN-10-1832

PCR-free whole genome amplification was performed. The DNA samples were subsequently fragmented, purified, and hybridized to Illumina Human1M-duo arrays containing probes for 1.1 million SNPs (Illumina). After hybridization, the bead arrays were stained and fluorescence intensities were quantified on an Illumina Bead Array 500GX Scanning Station. Raw data were analyzed using Illumina Genome Studio software, and SNP genotype reports were generated allowing WGAs analysis using the Plink software (http://pngu.mgh.harvard.edu/~purcell/plink, version 1.03; ref. 30). For WGAs, categorization of the test results in two separate groups is required. Therefore, EBV-LCL were divided into MiHA^{pos} and MiHA^{neg} groups using five times the level of IFN-γ production in the absence of EBV-LCL as a threshold for recognition. WGAs was performed by combining T-cell recognition with SNP genotyping data. The level of matching between both patterns was calculated according to Fisher's exact test.

PCR amplification and gene sequencing

RNA from the patients and their donors was isolated from EBV-LCL using Trizol (Invitrogen) and transcribed into cDNA by reverse transcriptase (Invitrogen) using oligo-dT primers (Roche Diagnostics). Genes that were located in associating genomic regions were amplified by PCR using various forward and reverse primers that were selected to comprise entire gene transcript or specific SNP-containing regions (see Supplementary Table S1 online). PCR products were purified and analyzed by DNA sequencing, and the obtained patient and donor sequences were aligned to identify SNP disparities.

Peptide prediction and analysis

For each SNP that was identified by WGAs and gene sequencing, gene fragments comprising the polymorphic patient-type nucleotide flanked by 30 upstream and downstream nucleotides were translated into different reading frames. The online algorithm of NetMHC (31) was used to search for peptides with predicted binding to the appropriate HLA molecule. Peptide candidates were synthesized and dissolved in DMSO. Dilutions were prepared in IMDM and added to MiHA^{neg} donor EBV-LCL (3×10^4 per well) in 96-well U-bottomed plates for 2 hours at 37°C. T cells (5×10^3 per well) were added, and after 24 hours, the culture supernatant was used for IFN-γ ELISA.

Tetramer production and staining

Tetramers were constructed by folding peptides in appropriate biotinylated MHC-monomers followed by multimerization using streptavidin conjugated to PE or APC as previously described with minor modifications (32). Patient samples were thawed and stained with FITC-conjugated CD4 and CD14 antibodies to exclude nonrelevant cells in flow cytometric analysis. MiHA-specific T cells were visualized using PE- or APC-conjugated tetramers and PE-Cy7–labeled anti-CD8. On average, 5×10^5 viable cells were analyzed per staining using a Becton Dickinson FACS CantoII and DIVA software.

Results

Selection of MiHA-specific CD8$^+$ T-cell clones eligible for WGAs

CD8$^+$ T cells were isolated from PBMCs obtained 5 or 6 weeks, and 8 weeks after DLI from patient H and patient Z, respectively. Single activated CD8$^+$ T cells were isolated using flow cytometry based on expression of HLA-DR. Percentages of CD8$^+$ T cells expressing HLA-DR were 11% (week 5) and 19% (week 6) for patient H and 3% (week 8) for patient Z (data not shown). Clonal expansion was observed for 195 T-cell clones from patient H and 37 T-cell clones from patient Z. All 232 clones were tested for recognition of patient- and donor-derived EBV-LCL by IFN-γ ELISA. The strategy followed for selection of MiHA-specific T-cell clones eligible for WGAs is depicted in Fig. 1A. In total, 62 T-cell clones from patient H and 16 T-cell clones from patient Z differentially recognized patient, but not donor, EBV-LCL, indicating recognition of MiHAs. All 78 patient-specific T-cell clones were investigated for HLA restriction using specific HLA class I, HLA-A*02, and B*07 blocking antibodies. Restriction to HLA-A*02 or B*07 was observed for 43 T-cell clones from patient H and 9 T-cell clones from patient Z. To identify whether T-cell clones recognized the same MiHA, we tested a panel of 30 HLA-A*02– and B*07-expressing EBV-LCL. In addition, TCRBV diversity was determined by flow cytometric analysis using specific TCRBV monoclonal antibodies and by DNA sequencing of the TCRBV CDR3 regions (33). This led to identification of 20 unique MiHA recognition patterns, including 13 MiHAs recognized by single T-cell clones and 7 MiHAs recognized by groups of T-cell clones (Table 1). TCRBV diversity was observed in five of the seven groups, indicating the presence of polyclonal T-cell responses targeting these MiHAs. Application of WGAs requires a balanced distribution between MiHA^{pos} and MiHA^{neg} test samples. Recognition of 30 EBV-LCL was used to determine this parameter, resulting in exclusion of three clones recognizing rare (clones H11 and H15) or common (clone H9) MiHAs. Cross-reactivity to nonself HLA molecules was excluded by the lack of recognition of a panel of 25 HLA-A*02– and B*07-negative EBV-LCL.

Identification of associating genomic regions by WGAs

A panel of 80 HLA-A*0201– and B*0702-expressing EBV-LCL was generated to identify MiHA-encoding genes by WGAs. Genomic DNA was isolated from each individual EBV-LCL for SNP genotype analysis using Illumina Human 1M-duo bead chips. To estimate the quality of the array data, we extracted SNP genotypes for known MiHAs. For 16 of the 27 known MiHAs, probes were included on the Illumina Human 1M-duo bead array, and all SNP genotypes were accurately measured with call rates >95% and allele frequencies comparable with dbSNP data (see Supplementary Table S2 online). Next, recognition of all EBV-LCL by each WGAs-eligible T-cell clone was tested using IFN-γ ELISA. Figure 2 shows the recognition of all EBV-LCL for the selected T-cell clones. Clear clustering of MiHA^{pos} and

Published OnlineFirst November 9, 2010; DOI: 10.1158/0008-5472.CAN-10-1832

MiHA^{neg} EBV-LCL cells was observed for the majority of the T-cell clones. For clones H6 and H12, MiHA^{pos} EBV-LCL could not be clearly discriminated from MiHA^{neg} EBV-LCL due to the large number of EBV-LCL that were intermediately recognized by the T cells. WGAs was performed by comparing each T-cell recognition pattern with all SNP genotypes. The level of matching between both patterns was calculated according to Fisher's exact test. By selecting SNP genotypes that significantly associated with a T-cell recognition pattern with P values $<10^{-9}$, single genomic regions could be identified for 12 T-cell clones. The identified regions varied in length from 1 bp up to 100 kb, comprising a maximum number of three genes and containing 1 to 17 associating SNPs (see Supplementary Fig. S1 and Table S3 online). Lack of association with a genomic region was observed for five T-cell clones, including clones H6 and H12, which produced intermediate levels of IFN-γ upon coincubation with a number of EBV-LCL. Clones H13 and H14, which also failed to show association with a genomic region, recognized MiHAs with low population frequencies of 10% and 11%, respectively, suggesting that the power of WGAs was not sufficient to identify these MiHAs.

Identification of MiHA-encoding SNPs

The strongest associating SNPs within each region were further examined, and associating SNPs were found in coding and noncoding gene regions (Table 2). Associating missense SNPs, encoding amino acid polymorphisms in the normal ORF, were found in the genes *WNK1, SSR1, PRCP, PDCD11, EBI3*, and *APOBEC3B*. These SNPs may directly encode the MiHA. Clone H1 and its associating SNP rs12828016 located in exon 21 of the *WNK1* gene is shown as a representative example in Fig. 3A. Associating synonymous SNPs were found in the genes *ARHGDIB* and *GEMIN4*, and associating SNPs in noncoding regions were found in the genes *TTK, ERAP1, BCAT2*, and *ERGIC1*. Associating synonymous SNPs and SNPs in noncoding regions may serve as marker SNPs for missense SNPs, which are not measured by the bead array. To search for these disparate missense SNPs, all genes comprising associating synonymous SNPs and SNPs in noncoding regions as identified by WGAs were investigated by analyzing cDNA derived from both patient and donor. Sequence analysis of gene transcripts revealed the presence of disparate missense SNPs in the genes *ERAP1, BCAT2*, and *GEMIN4*. Figure 3B shows clone H8 and its associating SNP rs26654 located in intron 1 of the *ERAP1* gene as a representative example of a marker SNP, which is in linkage disequilibrium with SNP rs26654 located in exon 2 identified after gene sequencing. Sequence analysis revealed that *ARHGDIB* gene transcripts were completely identical between patient and donor, except for the single associating synonymous SNP rs4703 as identified by

Table 1. Isolation of CD8$^+$ T-cell clones recognizing 20 unique MiHAs

Patient	Clone type	T-cell clone characteristics*			
		HLA restriction	No. of clones	TCRBV diversity	MiHA frequency (%)
H	H 1	A*02	2	2	50
	H 2	A*02	1	1	41
	H 3	A*02	1	1	23
	H 4	A*02	14	3	50
	H 5	A*02	1	1	97
	H 6	A*02	1	1	52
	H 7	A*02	1	1	0
	H 8	B*07	11	1	67
	H 9	B*07	4	2	37
	H 10	B*07	1	1	60
	H 11	B*07	2	2	53
	H 12	B*07	1	1	39
	H 13	B*07	1	1	13
	H 14	B*07	1	1	10
	H 15	B*07	1	1	0
Z	Z 1	B*07	1	1	47
	Z 2	B*07	1	1	37
	Z 3	B*07	4	2	72
	Z 4	B*07	1	1	53
	Z 5	B*07	2	nt	34

Abbreviation: nt, not tested.

*T-cell clones were isolated from patient H and patient Z after DLI at 5 or 6 weeks, and at 8 weeks, respectively, based on HLA-DR expression.

Published OnlineFirst November 9, 2010; DOI: 10.1158/0008-5472.CAN-10-1832

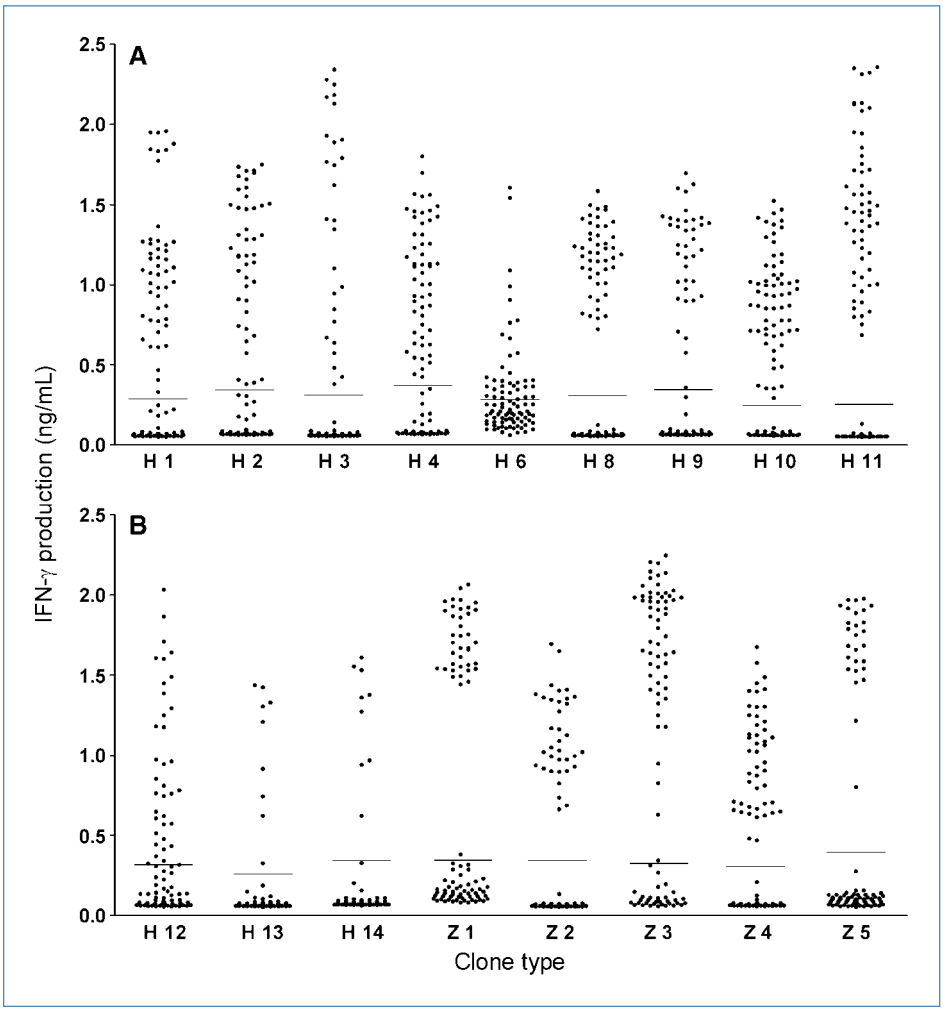

Figure 2. Recognition of a panel of 80 EBV-LCL by MiHA-specific T-cell clones. HLA-A*0201– and B*0702-restricted T-cell clones recognizing MiHAs with population frequencies between 10% and 90% were tested for recognition of a panel of 80 HLA-A*0201 and B*0702 EBV-LCL. Each dot represents the mean of triplicate IFN-γ production (ng/mL) as measured by ELISA after overnight coincubation of the T-cell clone (indicated on the *X* axis) with each EBV-LCL. Lines indicate the threshold of five times background IFN-γ production to segregate $MiHA^{pos}$ from $MiHA^{neg}$ EBV-LCL.

WGAs. These data suggest that this SNP directly encodes the MiHA in an ARF, as previously reported for other MiHAs (8, 15). The *TTK* and *ERGIC1* genes, identified by strong association of SNPs in noncoding regions, did not show any SNP disparity between patient and donor in coding regions. In conclusion, WGAs in combination with gene sequencing led to identification of disparate missense SNPs for 10 of the 12 genes (Table 2).

Prediction and recognition of MiHA epitopes

Gene sequences containing missense and synonymous SNPs identified by WGAs and DNA sequencing were translated in the normal ORF and in all ARFs, respectively, and searched for peptides with predicted binding to HLA-A*0201 and B*0702 using the online algorithm of NetMHC (31). All peptide variants with predicted binding to HLA-A*0201 and B*0702 are shown in Table 3. Peptides of 9 to 11 amino acids with predicted binding to HLA-A*0201 or B*0702 were synthesized, exogenously pulsed on $MiHA^{neg}$ donor EBV-LCL, and tested for T-cell recognition. For nine polymorphic peptides, strong recognition of a single predicted candidate peptide or variable levels of recognition for candidate peptides of various lengths was observed (Table 3). The SSR1-derived HLA-A*0201 binding candidate peptide, however, was not recognized. Longer 14-mer and 21-mer SSR1-derived peptides spanning the identified patient-type amino acid polymorphism were clearly recognized, albeit at high peptide concentrations of 50 μmol/L (data not shown), thereby confirming that the missense SNP encodes the MiHA epitope, but leaving the precise composition of the epitope unknown. For all identified MiHA epitopes as shown in Table 3, as well as for the longer SSR1 peptides, no recognition of donor-type peptides was observed. This validated the identification of 10 novel MiHAs encoded by SNPs as identified by our approach of WGAs (Fig. 1B).

Detection of MiHA-specific T-cell responses

To analyze the kinetics of the T-cell responses specific for the novel MiHAs as identified by WGAs, donor samples as well as sequentially obtained peripheral blood samples from the patients were stained with tetramers and analyzed by flow cytometry. Except for LB-BCAT2-1R, we could construct

Published OnlineFirst November 9, 2010; DOI: 10.1158/0008-5472.CAN-10-1832

stable tetramers that successfully stained the respective T-cell clones for all novel MiHAs. Tetramer analysis revealed the presence of T cells specific for the MiHAs in patient H, and for one MiHA in patient Z (Fig. 4). The frequencies of MiHA-specific T cells peaked at day 41 and 62 in patients H and Z, respectively, coinciding with conversion to full donor chimerism in patient H and disappearance of the BCR-ABL fusion product as a marker for disease in patient Z. No significant tetramer-positive events were detected using tetramers specific for LB-APOBEC3B-1K and for LB-EBI3-1I in patient Z.

Discussion

In this study, we developed an efficient strategy for T-cell selection and MiHA identification by WGAs. This WGAs-based strategy allowed detailed investigation of the specificity and diversity of antitumor T-cell responses in two patients successfully treated with HLA-matched allo-SCT and DLI. The strategy is based on efficient selection of MiHA-specific T-cell clones and a panel of 80 third-party EBV-LCL for simultaneous measurement of T-cell recognition and 1.1 million SNP genotypes. From the 17 selected T-cell specificities, 10 novel MiHAs were identified, illustrating the value of this WGAs-based strategy for broad characterization of MiHAs in clinical immune responses.

For 5 of the 17 selected T-cell specificities, we failed to show significant association with SNP genotypes by WGAs. Two T-cell clones recognized a limited number of EBV-LCL, suggesting that the power of the WGAs strategy may not be sufficient to identify MiHAs with low population frequencies. Two other T-cell clones for which we failed to show association with SNP genotypes by WGAs released intermediate levels of IFN-γ upon coincubation with a number of EBV-LCL, thereby hampering separation of EBV-LCL in $MiHA^{pos}$ and $MiHA^{neg}$ groups. Finally, for one T-cell clone, a balanced segregation of $MiHA^{pos}$ and $MiHA^{neg}$ EBV-LCL was observed; however, no significant association with any SNP genotype was detected. For this T-cell clone, association may be found by measuring more SNP genotypes, for example, by testing EBV-LCL from the International HapMap project, which have been genotyped for more than 4 million SNPs and recently shown to represent a valuable source for MiHA discovery (25–27).

Our WGAs-based strategy revealed a number of SNPs that associated with T-cell recognition with significant *P* values. We separated the missense SNPs from the synonymous SNPs and SNPs located in noncoding regions.

Table 2. Identification of MiHA-encoding SNPs

Clone type	Identified candidate SNPs			SNP function	SNP genotype		Gene	
	By association	*P*	By sequencing*		Patient	Donor	Name	Location
H 1	rs12828016	1.1×10^{-16}		Missense	AC	CC	*WNK1*	Exon 21
H 2	rs1886568	7.8×10^{-15}		—	GG	AA	*SSR1*	Intron 1
	rs10004	7.8×10^{-15}		Missense	GG	AA	*SSR1*	Exon 2
	rs3778333	7.8×10^{-15}		—	GG	AA	*SSR1*	Intron 2
H 3	rs2298668	1.1×10^{-12}		Missense	AC	AA	*PRCP*	Exon 4
H 4	rs608962	3.6×10^{-20}	Not found	—	GG	AA	*TTK*	Intron 20
H 8	rs26654	1.2×10^{-19}		—	AG	AA	*ERAP1*	Intron 1
			rs26653	Missense	CC	GG	*ERAP1*	Exon 2
H 9	rs736170	1.3×10^{-16}		—	AG	AA	*BCAT2*	Intron 1
	rs4801775	1.3×10^{-16}		—	AG	AA	*BCAT2*	Intron 3
	rs4802507	1.3×10^{-16}		—	AC	AA	*BCAT2*	Intron 3
			rs11548193	Missense	CG	CC	*BCAT2*	Exon 6
H 10	rs4703	5.5×10^{-16}		Synonymous	CG	GG	*ARHGDIB*	Exon 6
H 11	rs2986014	4.2×10^{-15}		Missense	AG	GG	*PDCD11*	Exon 24
	rs2281860	4.2×10^{-15}		—	AG	GG	*PDCD11*	Intron 30
	rs6580	4.2×10^{-15}		—	AG	AA	*CALHM2*	UTR 3′
Z 1	rs1045481	9.1×10^{-15}		Synonymous	AG	GG	*GEMIN4*	Exon 2
	rs1045491	9.1×10^{-15}		—	AG	GG	*GEMIN4*	UTR 3′
			rs4968104	Missense	AT	AA	*GEMIN4*	Exon 2
	rs1064245	9.1×10^{-15}		—	AC	AA	*FAM57A*	UTR 3′
Z 2	rs4740	3.4×10^{-14}		Missense	AG	GG	*EBI3*	Exon 5
	rs4905	3.4×10^{-14}		Synonymous	AG	AA	*EBI3*	Exon 5
Z 3	rs2076109	8.8×10^{-17}		Missense	AG	GG	*APOBEC3B*	Exon 3
Z4	rs564349	1.6×10^{-15}	Not found	—	GG	AA	*ERGIC1*	Intron 3

*Additional SNP disparities in coding regions identified after gene sequencing.

Published OnlineFirst November 9, 2010; DOI: 10.1158/0008-5472.CAN-10-1832

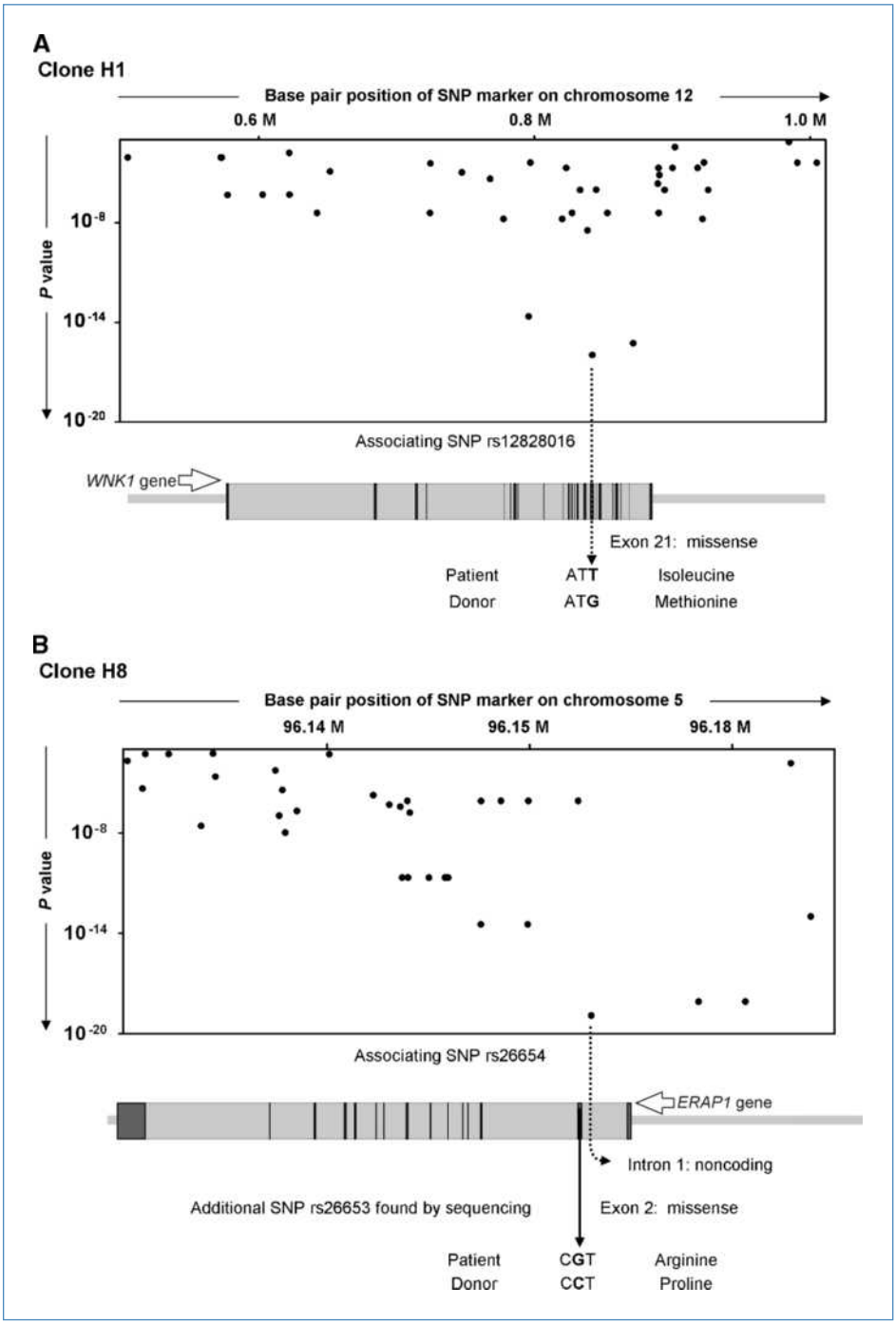

Figure 3. Identification of candidate genes and MiHA-encoding SNPs by WGAs. Association results and selection of candidate MiHA-encoding SNPs are shown for *WNK1* (A) and *ERAP1* (B) genes as representative examples. Significance of association for each individual array SNP as expressed by *P* value is depicted. Each dot represents the array SNP relative to its physical position on the genome. The image below each plot shows the gene aligned to the chromosomal position as indicated on the *X* axis of the plot. Exons and introns are represented by dark and light boxes, respectively. A, for clone H1, the strongest associating SNP is located in exon 21 of the *WNK1* gene and encodes an amino acid polymorphism. B, the strongest associating SNP for clone H8 is located in intron 1 of the *ERAP1* gene. Sequence analysis revealed a disparate missense SNP in exon 2 of the *ERAP1* gene.

Missense SNPs were directly analyzed for HLA binding prediction and identification of the MiHA peptide epitopes. All six missense SNPs were found to encode the MiHA epitope. Synonymous SNPs and SNPs in noncoding regions may serve as markers for MiHA-encoding SNPs that are not measured by the bead array. Therefore, genes comprising marker SNPs were first analyzed by gene sequencing to reveal all SNP disparities between patient and donor. For two T-cell clones, associating synonymous SNPs were found, of which one was shown to serve as a marker for a MiHA-encoding missense SNP and the other directly encoded the MiHA epitope translated in an ARF, as previously described for LB-ECGF-1H (8) and LB-ADIR-1F (15). Finally, for four T-cell clones, associating SNPs were located in noncoding regions. In two cases, intron SNPs were in linkage disequilibrium with MiHA-encoding missense SNPs, whereas no SNP disparities in coding gene regions were found for the other two T-cell clones.

Published OnlineFirst November 9, 2010; DOI: 10.1158/0008-5472.CAN-10-1832

For two T-cell clones showing strong association with intron SNPs, without SNP disparities in the coding regions of the associating genes of patient and donor, we explored the possibility that a monomorphic epitope is recognized, which may be under differential control of a SNP located in a regulatory region. The entire coding regions of the *TTK* and *ERGIC1* genes were cloned into an expression vector and introduced into MiHA-negative cells, but no T-cell recognition of the transfected genes was observed (data not shown). Therefore, we consider it likely that these MiHAs are encoded by SNPs located outside known protein coding regions of the candidate genes, for example, by an intron SNP of an alternative mRNA splice variant, as previously described for ACC-6 (19), and we speculate that these MiHAs may be most difficult to identify by WGAs.

As a final step in the discovery of MiHA, specific recognition of patient-type peptide was established. All identified MiHA epitopes were accurately predicted using the binding algorithm of NetMHC (31), as shown by T-cell recognition of donor EBV-LCL pulsed with the predicted patient-type, but not donor-type, peptide variants. Weak, but patient-specific, recognition was observed for peptides derived from APOBEC3B and *SSR1*. The *SSR1* epitope is a nonconventional peptide of 14 amino acids. Although recognition was weak, the T-cell clone differentially recognized patient-type, but not donor-type, 14-mer peptides demonstrating T-cell specificity for this polymorphism. Weak recognition of pulsed synthetic peptides may indicate posttranslational modification of endogenous epitopes, as previously described for the *SMCY* antigen (34). Another explanation may be that the T-cell clone expresses a T-cell receptor with a relatively low avidity, which is, however, sufficiently high to recognize MiHAs that are abundantly expressed at the cell surface of patient cells.

The MiHAs described in this study were identified as targets for T cells isolated from patients receiving DLI for incomplete donor chimerism or recurrent disease after allo-SCT. Specific tetramer staining of $CD8^+$ T cells was shown for the majority of the novel MiHAs, and peak frequencies coincided with the development of clinical

Table 3. Prediction and recognition of MiHA epitopes

Clone type	MiHA name	SNP	Amino acid polymorphism*	Restriction	Predicted HLA binding peptides*	Measured EC_{50} peptide recognition (nmol/L)
H 1	LB-WNK1-1I	rs12828016	Ile/Met	A*0201	RTLSPEIITV	4
					TLSPEIITV	>5,000
H 2	LB-SSR1-1S	rs10004	Ser/Leu	A*0201	SLAVAQDLT	>5,000
H 3	LB-PRCP-1D	rs2298668	Asp/Glu	A*0201	FMWDVAEDLKA	1
					FMWDVAEDL	17
H 8	LB-ERAP1-1R	rs26653	Arg/Pro	B*0702	HPRQEQIALLA	20
					HPRQEQIALL	nt
					HPRQEQIAL	nt
H 9	LB-BCAT2-1R	rs11548193	Arg/Thr	B*0702	QPRRALLFVIL	6
					QPRRALLFVI	195
					QPRRALLFV	>5,000
					GVSQPRRALL	>5,000
					GVSQPRRAL	nt
H 10	LB-ARHGDIB-1R	rs4703	Arg/Pro†	B*0702	LPRACWREA	4
					LPRACWREAR	45
					LPRACWREART	249
H 11	LB-PDCD11-1F	rs2986014	Phe/Leu	B*0702	GPDSSKTFLCL	28
					GPDSSKTFL	>5,000
Z 1	LB-GEMIN4-1V	rs4968104	Val/Glu	B*0702	FPALRFVEV	3
Z 2	LB-EBI3-1I	rs4740	Ile/Val	B*0702	RPRARYYIQVA	480
					RPRARYYIQV	13
					RPRARYYIQ	nt
					AVRPRARYYI	nt
Z 3	LB-APOBEC3B-1K	rs2076109	Lys/Glu	B*0702	KPQYHAEMCFL	4,447
					KPQYHAEMCF	285
					KPQYHAEMC	1,604

Abbreviation: nt, not tested.

*Patient-type amino acid residues are underlined.

†Amino acid polymorphism after translation in an ARF.

Published OnlineFirst November 9, 2010; DOI: 10.1158/0008-5472.CAN-10-1832

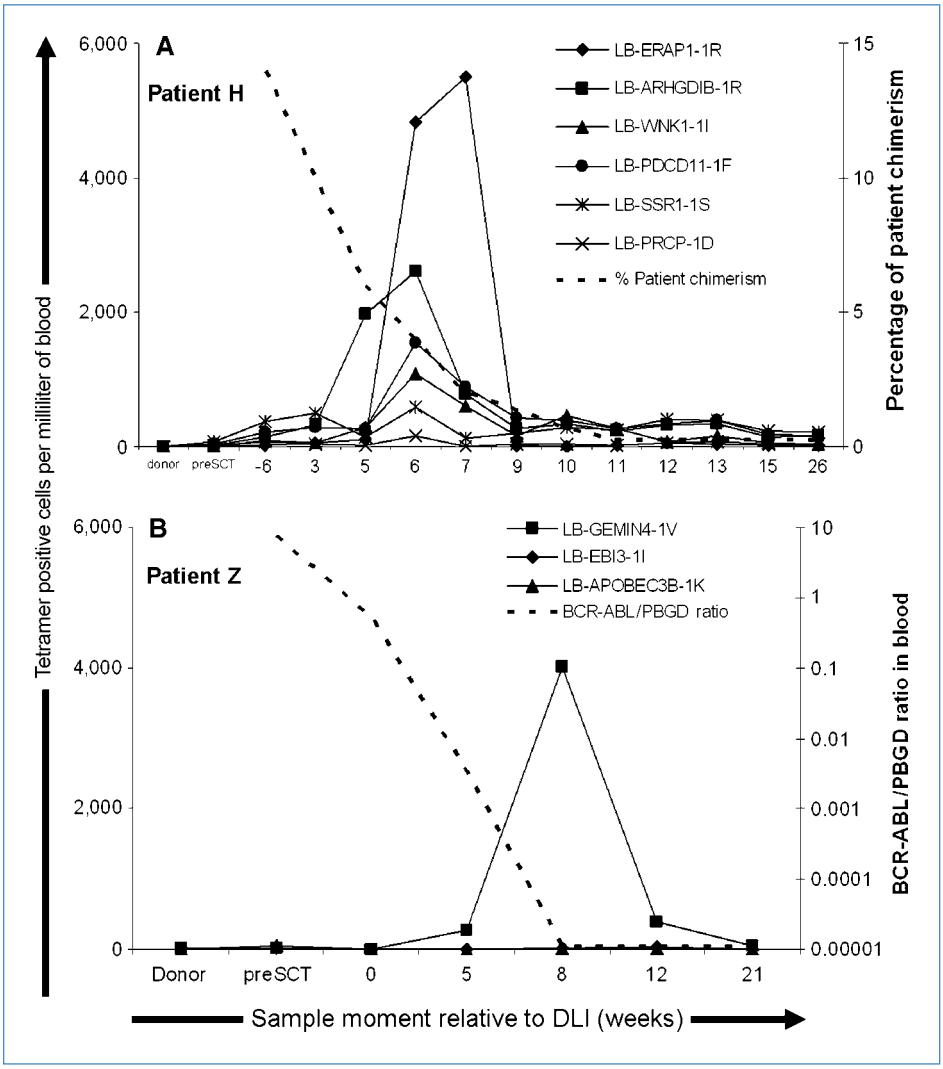

Figure 4. MiHA-specific T-cell responses in patients H and Z. Tetramers were constructed to measure MiHA-specific $CD8^+$ T cells in peripheral blood samples obtained at several time points, as indicated on the *X* axis. On the primary *Y* axis, the absolute number of tetramer-positive cells per milliliter of blood is depicted. The secondary *Y* axis depicts donor chimerism in patient H (A) and disease status in patient Z (B).

responses in the patients. Despite staining of the respective T-cell clones (data not shown), no tetramer-positive T cells were detected in the samples of patient Z using LB-APOBEC3B-1K and LB-EBI3-1I tetramers. Intermediate TCR avidity, as expressed by the intermediate EC_{50} value for recognition of *APOBEC3B*-derived peptides, may cause the lack of detection of the LB-APOBEC3B-1K–specific T cells. The absence of tetramer-positive T cells for LB-EBI3-1I may be explained by a low frequency of circulating T cells. Whereas the individual contribution of each MiHA-directed T-cell response cannot be determined, the induction of strong GVT reactivity and the absence of severe GvHD in both patients suggest involvement of one or more MiHAs with selective or predominant expression in malignant hematopoietic cells.

In conclusion, we present an efficient WGAs-based strategy for broad characterization of MiHAs. Our WGAs approach is valuable as a high-throughput method for detailed analysis of the specificity and diversity of T-cell responses in patients with different clinical responses as well as for selective identification of clinically relevant MiHAs with hematopoietic restricted expression.

Acknowledgments

We thank Lourdes Sampietro for help with the bead array analysis, Quinta Helmer for data file processing, Menno van der Hoorn and Guido de Roo for flow cytometric cell sorting, and Dirk van der Steen for tetramer production.

Grant Support

Dutch Cancer Society grant no. 05-3267 (C.E. Rutten) and European Union 6th Framework Programme, Allostem project no. 503319 (M. Griffioen).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 05/27/2010; revised 09/08/2010; accepted 09/13/2010; published OnlineFirst 11/09/2010.

Published OnlineFirst November 9, 2010; DOI: 10.1158/0008-5472.CAN-10-1832

References

1. Appelbaum FR. The current status of hematopoietic cell transplantation. Annu Rev Med 2003;54:491–512.
2. Kolb HJ. Graft-versus-leukemia effects of transplantation and donor lymphocytes. Blood 2008;112:4371–83.
3. Feng X, Hui KM, Younes HM, Brickner AG. Targeting minor histocompatibility antigens in graft versus tumor or graft versus leukemia responses. Trends Immunol 2008;29:624–32.
4. Falkenburg JHF, van de Corput L, Marijt EWA, Willemze R. Minor histocompatibility antigens in human stem cell transplantation. Exp Hematol 2003;31:743–51.
5. Mullally A, Ritz J. Beyond HLA: the significance of genomic variation for allogeneic hematopoietic stem cell transplantation. Blood 2007; 109:1355–62.
6. Dierselhuis M, Goulmy E. The relevance of minor histocompatibility antigens in solid organ transplantation. Curr Opin Organ Transplant 2009;14:419–25.
7. Marijt WA, Heemskerk MH, Kloosterboer FM, et al. Hematopoiesis-restricted minor histocompatibility antigens HA-1- or HA-2-specific T cells can induce complete remissions of relapsed leukemia. Proc Natl Acad Sci U S A 2003;100:2742–7.
8. Slager EH, Honders MW, van der Meijden ED, et al. Identification of the angiogenic endothelial cell growth factor-1/thymidine phosphorylase as a potential target for immunotherapy of cancer. Blood 2006; 107:4954–60.
9. Ferrara JL, Levine JE, Reddy P, Holler E. Graft-versus-host disease. Lancet 2009;373:1550–61.
10. den Haan JM, Sherman NE, Blokland E, et al. Identification of a graft versus host disease-associated human minor histocompatibility antigen. Science 1995;268:1476–80.
11. den Haan JM, Meadows LM, Wang W, et al. The minor histocompatibility antigen HA-1: a diallelic gene with a single amino acid polymorphism. Science 1998;279:1054–7.
12. Brickner AG, Warren EH, Caldwell JA, et al. The immunogenicity of a new human minor histocompatibility antigen results from differential antigen processing. J Exp Med 2001;193:195–206.
13. Spierings E, Brickner AG, Caldwell JA, et al. The minor histocompatibility antigen HA-3 arises from differential proteasome-mediated cleavage of the lymphoid blast crisis (Lbc) oncoprotein. Blood 2003;102:621–9.
14. Brickner AG, Evans AM, Mito JK, et al. The PANE1 gene encodes a novel human minor histocompatibility antigen that is selectively expressed in B-lymphoid cells and B-CLL. Blood 2006;107: 3779–86.
15. van Bergen CAM, Kester MGD, Jedema I, et al. Multiple myeloma-reactive T cells recognize an activation-induced minor histocompatibility antigen encoded by the ATP-dependent interferon-responsive (ADIR) gene. Blood 2007;109:4089–96.
16. Dolstra H, Fredrix H, Maas F, et al. A human minor histocompatibility antigen specific for B cell acute lymphoblastic leukemia. J Exp Med 1999;189:301–8.
17. Murata M, Warren EH, Riddell SR. A human minor histocompatibility antigen resulting from differential expression due to a gene deletion. J Exp Med 2003;197:1279–89.
18. Warren EH, Vigneron NJ, Gavin MA, et al. An antigen produced by splicing of noncontiguous peptides in the reverse order. Science 2006;313:1444–7.
19. Kawase T, Akatsuka Y, Torikai H, et al. Alternative splicing due to an intronic SNP in HMSD generates a novel minor histocompatibility antigen. Blood 2007;110:1055–63.
20. Griffioen M, van der Meijden ED, Slager EH, et al. Identification of phosphatidylinositol 4-kinase type II ¦ as HLA class II-restricted target in graft versus leukemia reactivity. Proc Natl Acad Sci U S A 2008;105:3837–42.
21. Stumpf AN, van der Meijden ED, van Bergen CA, Willemze R, Falkenburg JH, Griffioen M. Identification of 4 new HLA-DR-restricted minor histocompatibility antigens as hematopoietic targets in antitumor immunity. Blood 2009;114:3684–92.
22. Akatsuka Y, Nishida T, Kondo E, et al. Identification of a polymorphic gene, BCL2A1, encoding two novel hematopoietic lineage-specific minor histocompatibility antigens. J Exp Med 2003;197:1489–500.
23. Rijke BD, Horssen-Zoetbrood A, Beekman JM, et al. A frameshift polymorphism in P2X5 elicits an allogeneic cytotoxic T lymphocyte response associated with remission of chronic myeloid leukemia. J Clin Invest 2005;115:3506–16.
24. Kawase T, Nannya Y, Torikai H, et al. Identification of human minor histocompatibility antigens based on genetic association with highly parallel genotyping of pooled DNA. Blood 2008;111:3286–94.
25. Spaapen RM, Lokhorst HM, van den OK, et al. Toward targeting B cell cancers with $CD4^+$ CTLs: identification of a CD19-encoded minor histocompatibility antigen using a novel genome-wide analysis. J Exp Med 2008;205:2863–72.
26. Kamei M, Nannya Y, Torikai H, et al. HapMap scanning of novel human minor histocompatibility antigens. Blood 2009;113:5041–8.
27. Spaapen RM, de Kort RA, van den OK, et al. Rapid identification of clinical relevant minor histocompatibility antigens via genome-wide zygosity-genotype correlation analysis. Clin Cancer Res 2009;15: 7137–43.
28. Warren EH, Fujii N, Akatsuka Y, et al. Therapy of relapsed leukemia after allogeneic hematopoietic cell transplant with T cells specific for minor histocompatibility antigens. Blood 2010;115:3869–78.
29. Miller G, Shope T, Lisco H, Stitt D, Lipman M. Epstein-Barr virus: transformation, cytopathic changes, and viral antigens in squirrel monkey and marmoset leukocytes. Proc Natl Acad Sci U S A 1972;69:383–7.
30. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.
31. Buus S, Lauemoller SL, Worning P, et al. Sensitive quantitative predictions of peptide-MHC binding by a 'Query by Committee' artificial neural network approach. Tissue Antigens 2003;62:378–84.
32. Burrows SR, Kienzle N, Winterhalter A, Bharadwaj M, Altman JD, Brooks A. Peptide-MHC class I tetrameric complexes display exquisite ligand specificity. J Immunol 2000;165:6229–34.
33. Arden B, Clarck S, Kabelitz D, Mak T. Human T-cell receptor variable gene segment families. Immunogenetics 1995;42:455–500.
34. Meadows L, Wang W, den Haan JM, et al. The HLA-A*0201-restricted H-Y antigen contains a posttranslationally modified cysteine that significantly affects T cell recognition. Immunity 1997;6:273–81.

AACR American Association for Cancer Research

Cancer Research

The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

High-Throughput Characterization of 10 New Minor Histocompatibility Antigens by Whole Genome Association Scanning

Cornelis A.M. Van Bergen, Caroline E. Rutten, Edith D. Van Der Meijden, et al.

Cancer Res 2010;70:9073-9083. Published OnlineFirst November 9, 2010.