

# Soluble plasma HLA peptidome as a potential source for cancer biomarkers

Michal Bassani-Sternberg<sup>a</sup>, Eilon Barnea<sup>a</sup>, Ilan Beer<sup>b</sup>, Irit Avivi<sup>c</sup>, Tami Katz<sup>c</sup>, and Arie Admon<sup>a,1</sup>

<sup>a</sup>Department of Biology, Technion-Israel Institute of Technology, Haifa 32000, Israel; <sup>b</sup>IBM Haifa Research Laboratory, Haifa 31905, Israel; and <sup>c</sup>Bruce Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Department of Hematology and Bone Marrow Transplantation, Rambam Health Care Campus, Haifa 31096, Israel

This Feature Article is part of a series identified by the Editorial Board as reporting findings of exceptional significance.

Edited by Peter Cresswell, Yale University School of Medicine, New Haven, CT, and approved September 10, 2010 (received for review June 16, 2010)

The HLA molecules are membrane-bound transporters that carry peptides from the cytoplasm to the cell surface for surveillance by circulating T lymphocytes. Although low levels of soluble HLA molecules (sHLA) are normally released into the blood, many types of tumor cells release larger amounts of these sHLA molecules, presumably to counter immune surveillance by T cells. Here we demonstrate that these sHLA molecules are still bound with their authentic peptide repertoires, similar to those of the membranal HLA molecules (mHLA). Therefore, a single immunoaffinity purification of the plasma sHLA molecules, starting with a few milliliters of patients' blood, allows for identification of very large sHLA peptidomes by mass spectrometry, forming a foundation for development of a simple and universal blood-based cancer diagnosis. The new methodology was validated using plasma and tumor cells of multiple-myeloma and leukemia patients, plasma of healthy controls, and with cultured cancer cells. The analyses identified thousands of sHLA peptides, including some cancer-related peptides, present among the sHLA peptidomes of the cancer patients. Furthermore, because the HLA peptides are the degradation products of the cellular proteins, this sHLA peptidomics approach opens the way for investigation of the patterns of protein synthesis and degradation within the tumor cells.

Identification of serum biomarkers for diseases, such as cancer and autoimmunity is one of the major goals of modern proteomics. Because proteins released from live or dying cancer cells are readily degraded or cleared from the serum, the degradation products of these proteins attract significant attention as peptide serum biomarkers (1). However, because small peptides are unstable in the serum, the search for disease-specific serum markers was redirected to peptides adsorbed to the "molecular sponge" serum albumin, known to carry with it a variety of peptides (2, 3). "Immuno-MS" of serum albumin with its adsorbed peptides was thus proposed as a potentially promising alternative for separation of subsets of peptides from human sera for analysis by MALDI-MS (1, 4, 5). However, serum albumin nonspecifically adsorbs many peptides and small molecules, which are not necessarily derived from the cancer cells.

The HLA class I molecules transport peptides to the cell surface after degradation of their proteins of origin. These HLA peptidomes include a diversity of peptides, derived from both normal and abnormal proteins expressed in the cells. Thus, the HLA peptidomes were studied mostly in order to identify cancer specific peptides, for development of tumor immunotherapeutics (6, 7) and as a source of information about the protein synthesis and degradation schemes within the cancer cells (8).

It has been known for a while that the levels of the soluble HLA class I molecules (sHLA-I) are elevated in the serum of people inflicted with diseases, such as cancer (9–13), autoimmunity (14), allergy (15), and viral infections (16). Such elevated levels of sHLA were even proposed to serve as indicators of poor prognosis (11, 17), yet their bound peptidomes have never been analyzed before. We postulated that if indeed a significant portion of the plasma sHLA molecules are actually released from

the diseased cells, and if these molecules carry with them their original peptide cargo, then analysis of the sHLA peptidomes may provide useful information about the tumor cells and an ideal source for disease biomarkers, serving as a high-sensitivity diagnostics for diseases, such as cancer.

Here we describe a pilot study aimed at confirming the feasibility of analysis of the plasma sHLA peptidomes as useful approach for identification of cancer biomarkers. This notion is based on our finding that in cancer patients, the sHLA peptidomes are similar to the membranal HLA (mHLA) peptidomes. The sHLA-peptidomes were collected from fresh plasma of multiple myeloma (MM) and leukemia [acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL)] patients, as well as from healthy controls, using single-step immunoaffinity purification, followed by identification of the bound peptides by microcapillary chromatography combined with tandem mass spectrometry ( $\mu$ LC-MS/MS).

## Results

**Isolation and Analysis of the Plasma sHLA Molecules.** After immunoaffinity purification of the intact sHLA complexes, with their bound peptide cargo, using 2–5 mL of human plasma, the eluted peptides were identified by  $\mu$ LC-MS/MS with an Orbitrap mass spectrometer (shown schematically in Fig. 1). The sHLA molecules, isolated from the plasma of the MM and AML and ALL leukemia patients and of the healthy controls (Table S1), were detected as the three expected molecular forms known to be released from human cells (Fig. 2A): a 35-kDa form, corresponding to the extracellular part of the HLA molecules, released by metalloproteases present at the cells' membrane (18, 19); a 39-kDa alternative spliced form, lacking its transmembrane domain (20); and a 43-kDa intact molecule, pinched off the cells with fragments of the membrane (21). The same three forms of sHLA molecules were also found to be secreted spontaneously from the three studied MM cell lines and were purified from the growth medium of the cells (Fig. 2B). As expected, only the full-size form of the HLA molecules (43 kDa) was observed in the mHLA preparations, from both the tumor cells and from MM cell lines (Fig. 2B).

**Analysis of the Plasma sHLA Peptidomes.** Because the sHLA molecules were purified by the pan-HLA A, B, and C monoclonal antibody (W6/32), which binds only the native HLA-peptide

Author contributions: M.B.-S. and A.A. designed research; M.B.-S., E.B., and I.B. performed research; M.B.-S., I.B., I.A., and T.K. contributed new reagents/analytic tools; M.B.-S., E.B., I.B., and A.A. analyzed data; and M.B.-S., E.B., I.B., and A.A. wrote the paper.

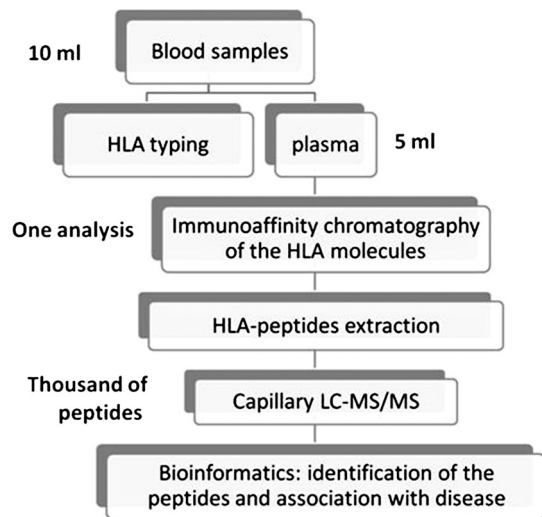
The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

See Commentary on page 18747.

<sup>1</sup>To whom correspondence should be addressed. E-mail: admon@tx.technion.ac.il.

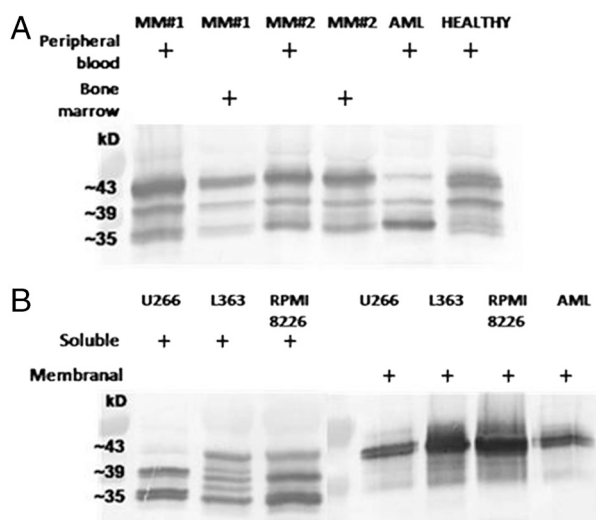
This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1008501107/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1008501107/-DCSupplemental) and [proteomecommons.org](http://proteomecommons.org).



**Fig. 1.** Schematic description of the new diagnostic methodology based on analysis of plasma soluble HLA peptidome.

complexes, we assumed that all of these forms are loaded with peptides and therefore contribute to the detected sHLA peptidomes. Indeed, the few micrograms of isolated sHLA molecules were sufficient for identification of hundreds to thousands sHLA peptides in each analysis (examples in Table 1 and larger lists in Tables S2–S4).

The identified sHLA peptides from the plasma samples (above 12,000 in this study) were filtered according to the degree of confidence in their identification scores by the different software tools (Pep-Miner, Sequest, and Mascot) and by the accuracy of the peptides mass measurements. We separated the peptides into two main groups according to the levels of “certainty” in the identifications. The high-certainty-identified HLA peptides were used for matching the peptides sequence to the consensus binding motifs of the different HLA haplotypes, and to search for cancer-related peptides. The identified “intermediate-certainty” HLA peptides are not listed, and their MS signals were used, in addition to the “high-certainty,” for correlation of signal intensities of sHLA peptidomes where the absolute identifications are not



**Fig. 2.** Western blot analysis of immunoaffinity purified HLA-I molecules. *A*: Plasma sHLA-I from peripheral blood and bone marrow from MM patient #1 and #2, AML-1 patient, and healthy 1 control. *B*: sHLA-I from the conditioned medium of MM cell lines: U266, L363, and RPMI8226, and the mHLA-I from the MM cell lines and from AML-1 cells.

essential. Table S2 lists the subset of high-certainty sHLA peptides, along with their  $\mu$ LC-MS/MS signal intensities, normalized retention times, identification scores, and fitness to HLA consensus binding motifs. Most of the identified peptides seem to be authentic HLA ligands, because their amino acid sequences fit the consensus binding motifs of the blood donors’ HLA (HLA A, B, and C, examples in Table 2). For example, the HLA haplotype of blood donor healthy 1 and healthy 6 are known (Table 2 and Table S1) each with his unique consensus binding motifs. Therefore, the high-certainty plasma sHLA peptides identified in these samples could be divided according to their fitness to the binding motif of their alleles. After removing from the list the known contaminating plasma-related peptides (described below), the remaining 98% of the high-certainty sHLA peptides of healthy 1 and of healthy 6 fit the consensus binding motifs of each of their alleles (summarized in Table 2). For example, the consensus binding motifs of one of the alleles of AML-1 and healthy 4 are not known, and therefore 12–13% of their high-certainty-identified peptides could not be assigned. As expected, the sHLA peptidomes of people with more unknown HLA consensus binding motifs contain even larger numbers of unassigned peptides. This analysis confirms that indeed the immunoaffinity purifications specifically enrich the sHLA peptidomes out of the highly complex protein-rich plasma.

Our sHLA peptidome preparations were contaminated to variable degrees by peptides derived from blood clotting and plasma proteins. Such peptides pose a problem only if they obscure the  $\mu$ LC-MS/MS signals of the authentic sHLA peptides. These peptides were easily singled out, because they did not fit the normal sizes of HLA peptides and they originate from known plasma proteins (examples in Table S5). To reduce the contamination of the sHLA peptidomes by these peptides, protease inhibitors were added to the plasma immediately after separating it from cells and platelets, and their levels were further reduced by an extra wash with nonionic detergents added to the immunoaffinity purifications procedure.

**Soluble HLA and the Membranal HLA Peptidomes of Cancer Cells Are Similar.** The sHLA and the mHLA peptidomes were compared after purifying the mHLA molecules by detergent solubilization of the cancer cells. First, the mHLA and the sHLA peptidomes, spontaneously secreted from three MM cultured cells, were compared (L363, RPMI8226, and U266) (Tables S6 and S7). Up to 40% of the sHLA peptides were detected also among the mHLA peptidomes of these cultured cells and the LC-MS signal intensities of the shared peptides of these sHLA and mHLA peptidomes were correlated (Fig. S1). Each point in Fig. S1 (and Figs. 3–6) indicates the normalized signal intensities of the shared peptides, detected in the different sHLA peptidomes. Both high-certainty ( $\square$ ) and intermediate-certainty peptides (+) were spread across the observed dynamic range of four orders of signal intensities. Next, the plasma sHLA peptidomes of leukemia patients were compared to the mHLA peptidomes isolated from each patient’s tumor cells, both obtained from one blood sample. Indeed, in advanced diseases, such as AML and ALL, up to 86% of the plasma sHLA peptides were detected also in the mHLA peptidomes of the patient’s tumor cells, and the LC-MS signal intensities of the shared peptides were also well correlated (Fig. 3, Table 3 and Table S3). These results strengthen our hypothesis that the plasma sHLA peptidome represents significantly the mHLA peptidome of the cancer cells.

**Cancer-Related sHLA Peptides in the Plasma.** The HLA peptidome represents degradation products of a significant part of the cellular proteome. Indeed, both the sHLA and mHLA peptidomes were derived from cellular proteins of different families and cellular functions. No preference was observed in the repertoires of the source proteins of the HLA peptidomes of the healthy and

**Table 1. Example for cancer-related HLA peptides identified from soluble HLA molecules of cancer patients**

Peptide	Access	Protein	Source	HLA haplotype score
YFIIIDHL	Q9NXF1	Testis-expressed sequence 10 protein	AML-1, healthy 1	Cw04:106, B35:20
GRWPWQGS	Q9Y6M0	Testisin	AML-1	B27:10000
SSVPRTAEL	Q99836	Myeloid differentiation primary response protein MyD88	MM-1	
GRAPGGLSL	P17275	Transcription factor jun-B	AML-1	B27:2000
IGVDFALKV	Q13637	Ras-related protein Rab-32	MM-1	B51:240
GTPAGGGFPR	Q07352	Butyrate response factor 1	MM-1, MM-2	A11:2
SVSNVVITK	P42566	Epidermal growth factor receptor substrate 15	MM-1, MM-2, AML-1, AML-3, ALL-1	A03:18, A11:4
VFDPVPVGV	Q08211	ATP-dependent RNA helicase A	AML-1	Cw04:45
GQLSVHTPK	Q9H334	Forkhead box protein P1	AML-1	B27:200
DGYVVKETI	Q15418	Ribosomal protein S6 kinase alpha-1	ALL-1	B51:532

Access, accession number in the UniProt database; HLA haplotype score, the best scores matching the HLA haplotype of the blood donors. The HLA types of the blood donors are listed in [Table S1](#).

diseased blood donors to membranal, cytoplasmic, or nuclear proteins. Examples for the gene ontology (GO) analyses (biological processes) of the identified sHLA and mHLA peptidomes are displayed in [Fig. S2 A and B](#), showing that both types of peptidomes include peptides derived from similar families and repertoires of proteins. Furthermore, the sHLA peptidomes, recovered from the patients' plasma, included peptides derived from cancer-related proteins, including tumor testis antigens, embryonic cancer antigens, and products of known tumor associated proteins (selected examples in [Table 1](#)). The listed putative tumor antigens are provided only as a proof of the concept, demonstrating that cancer-related peptides can be identified within the sHLA peptidomes of the patients. The sHLA peptidomes of cultured cancer cells contain also many more antigens that can be defined as cancer related according to the above

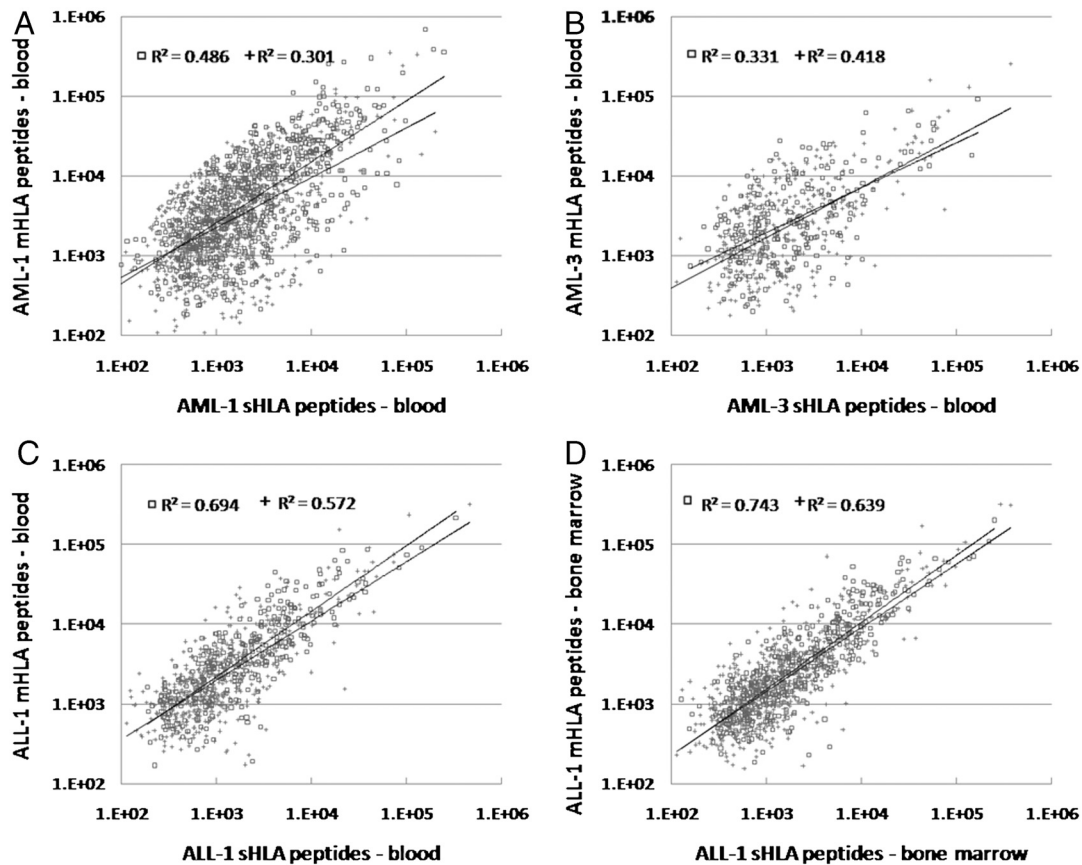
criteria ([Table S7](#)). Yet, their potential as clinically useful immunotherapeutics or biomarkers is not clear.

**Similar sHLA Peptidomes Were Observed in the Plasma Collected from the Peripheral Blood and from the Bone Marrow of Each Patient.** The plasma sHLA peptidomes of the bone marrow and of the peripheral blood samples of MM and ALL patients were compared. Up to 89% of the bone-marrow-derived plasma sHLA peptides were detected also in the plasma sHLA peptidomes of the peripheral blood, and their LC-MS signal intensities had strong correlation ([Fig. 4](#), [Table 3](#), and [Table S2](#)). These results indicate that sHLA molecules can provide useful information about the tumor micro-environment, because they circulate freely to the periphery where they can be easily collected.

**Table 2. The fitness of the sHLA peptides sequences to the HLA consensus binding motifs of the donors**

Subject	HLA allele	% high-certainty peptides	HLA-A+B supertype
Healthy 1; 250 high-certainty peptides	HLA-A01	37	A01
	HLA-A02	12	A02
	HLA-B35HLA-Cw04	49	B07*
	Unassigned motif	2	
Healthy 6; 180 high-certainty peptides	HLA-A02	15	A02
	HLA-A68	48	A03
	HLA-B18	24	B44
	HLA-B55HLA-Cw07	7	B07
	HLA-Cw03	4	*
	Unassigned motif	2	
Healthy 4; 200 high-certainty peptides	HLA-A24HLA-Cw07	42	A24
	HLA-A68	30	A03
	HLA-B49	12	B44
	HLA-Cw01	7	*
	Unassigned motif (HLA-B22)	12	
AML-1; 330 high-certainty peptides	HLA-A03	33	A03
	HLA-B27	6	B27
	HLA-B35HLA-Cw04	48	B07*
	Unassigned motif (HLA-Cw1502)	13	
Healthy 5; 244 high-certainty peptides	HLA-A26	17	A01
	HLA-A66	14	A03
	HLA-B38	4	B27
	HLA-B51	44	B07
	Unassigned motif (HLA-Cw12, HLA-Cw14)	21	
	MM-1; 250 high-certainty peptides	HLA-A02	12
HLA-A11	33	A03	
HLA-B51	20	B07	
HLA-B52	5	B62	
Unassigned motif (HLA-Cw02, HLA-Cw12)	30		
MM-2; 158 high-certainty peptides	HLA-A11	34	A03
	HLA-A26	13	A01
	HLA-B07	8	B07
	HLA-B38	1	B27
	Unassigned motif (HLA-Cw12, HLA-Cw15)	43	

Only sHLA peptidomes with above 150 high-certainty identified peptides were selected for this analysis. Fitness to the consensus binding motifs was assigned according to the HLA A and HLA B supertypes (22) and to HLA Cw (\*) according to (44).



**Fig. 3.** Correlation between the LC-MS signal intensities of the shared plasma sHLA peptidomes and cancer cells' mHLA peptidomes: The HLA peptidomes were isolated from the plasma and from the cancer cells of AML-1 (A), of AML-3 (B), from the plasma and from peripheral blood cells of ALL-1 (C); and from the bone marrow plasma and cells of ALL-1 (D). Square dots indicate high-score identified peptides with Pep-Miner score above 85, cross dots indicate intermediate-score peptides with Pep-Miner score of 70 to 85.

**Similar sHLA Peptidomes Are Detected in the Plasma of the Same Person Collected on Different Days.** The reproducibility of the recovered sHLA peptidomes was tested to assess the potential of this approach for biomarkers discovery. The initial validation of the method requires that consistent peptide repertoires should be present in each person's plasma sample collected on different days. For this purpose, plasma samples were collected from people on different days and the sHLA peptidomes were purified and analyzed separately. Indeed, very strong similarities between these peptidomes, with up to 91% identity in the identified peptides and similar signal intensities, were observed (Fig. 5, Table 3, and Table S4). Significantly, the similarities between the sHLA peptidomes collected on separate days (Fig. 5 A and B, and Table 3) were as high as the sHLA peptidome preparations divided into two and analyzed by two separate  $\mu$ LC-MS/MS analyses (Fig. 5C and Table 3).

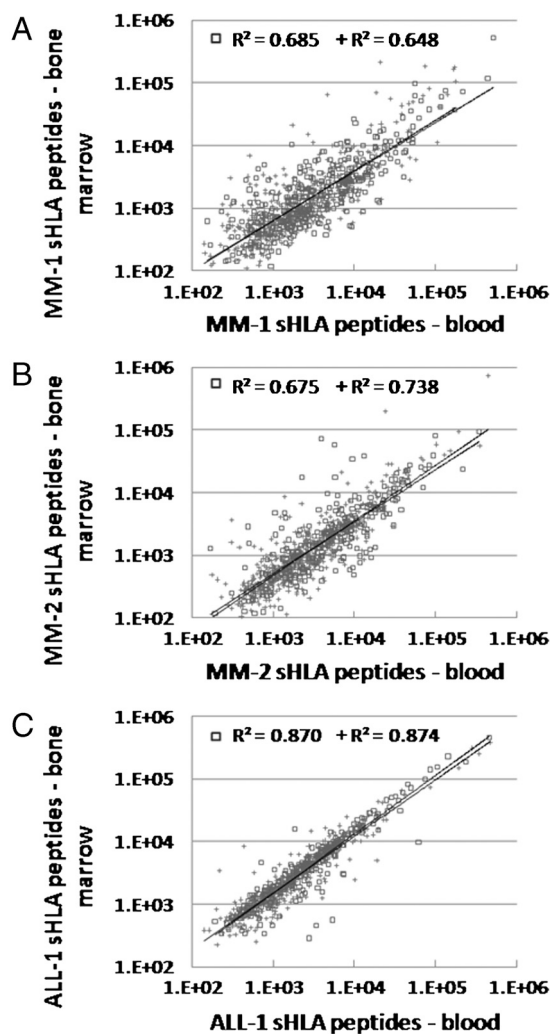
**HLA Haplotypes Define the sHLA Peptidomes.** Different people have different HLA haplotypes, and therefore their HLA peptidomes differ according to the peptide binding properties of their unique HLA alleles (22). Indeed, when the sHLA peptidomes of different people were compared, these peptidomes differed in accordance with the degree of difference between their HLA haplotypes. Large numbers of identical peptides were observed in peptidomes of people who have similar HLA haplotypes (Fig. 6 and Table S8). For example, as many as 1,531 identical peptides were observed in the peptidomes of AML-1 (out of total of 2,149 peptides) and healthy 1 (out of total of 2,070 peptides) sharing the alleles HLA-B35 and HLA-Cw4 (Fig. 6A and Table S8). In contrast,

AML-1 and healthy 4 (total 2,020 peptides), with no shared HLA alleles, had only 278 identical peptides (Fig. 6B and Table S8).

## Discussion

Immunoaffinity of the sHLA molecules with their bound peptides provides at least five orders of magnitude enrichment of the serum biomarkers. It bypasses the inherent difficulty associated with detection of cancer proteins or peptides present at diminishingly small concentrations in the otherwise protein-rich plasma. The sHLA molecules carry defined sets of peptides, largely derived from the cancer cells. This is in contrast to the nonspecific binding of peptides to the serum albumin (2, 3), previously suggested for analysis of its adsorbed peptides as a source for cancer biomarkers (1).

Because both the healthy and the diseased cells in the body release sHLA molecules, it is expected that, as the tumor loads of the patients' increase, larger fractions of the plasma sHLA peptidomes will originate from the tumor cells. Furthermore, cancer cells release larger amounts of sHLA to the circulation relative to healthy cells, possibly as an attempt to evade the anticancer immune response of T cells (23). However, the contributions of the diseased cells to the sHLA peptidomes are expected to vary not only according to the tumor size, but also according to its type and propensity to release sHLA to the circulation. Indeed, in advanced diseases, such as the AML and ALL studied here, where large numbers of cancer cells were present in the blood, large similarities of up to 86% were also observed between the cancer cells' mHLA and the plasma sHLA peptidomes of the same patients. Obviously, such comparisons between patients' plasma sHLA peptidomes and membrane HLA peptidomes of the cancer

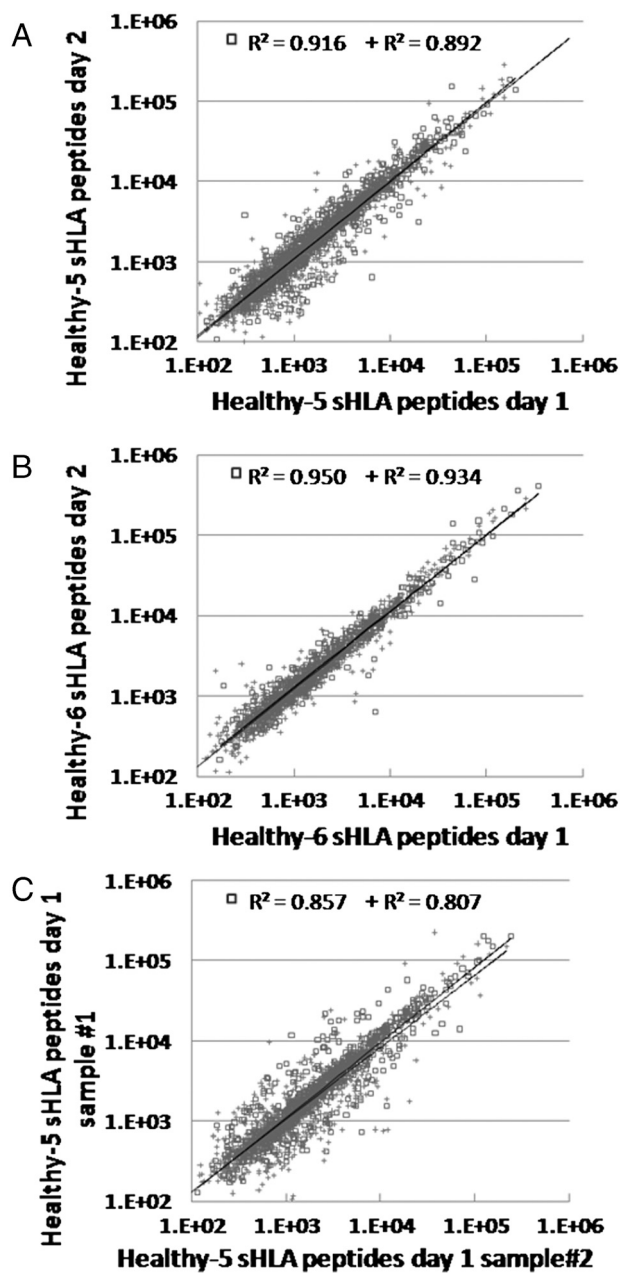


**Fig. 4.** Correlation between LC-MS signal intensities of the shared plasma sHLA peptides isolated from peripheral blood and from bone marrow samples from MM-1 (A), MM-2 (B) and ALL-1 (C). Square dots indicate high-score identified peptides with Pep-Miner score above 85, cross dots indicate intermediate-score peptides with Pep-Miner score of 70 to 85.

cells could be performed only with blood samples of advanced disease patients, containing very large numbers of cancer cells, and not with the healthy controls. The relatively small amounts of sHLA molecules, present in the plasma of healthy people, originate most likely from various tissues, which could not be sampled for these comparisons. Many types of cultured cells also maintain this propensity to release large amounts of sHLA molecules to their growth medium. Indeed when the mHLA and sHLA peptidomes of the human cancer cultured cells were compared, they resulted in similar peptidomes.

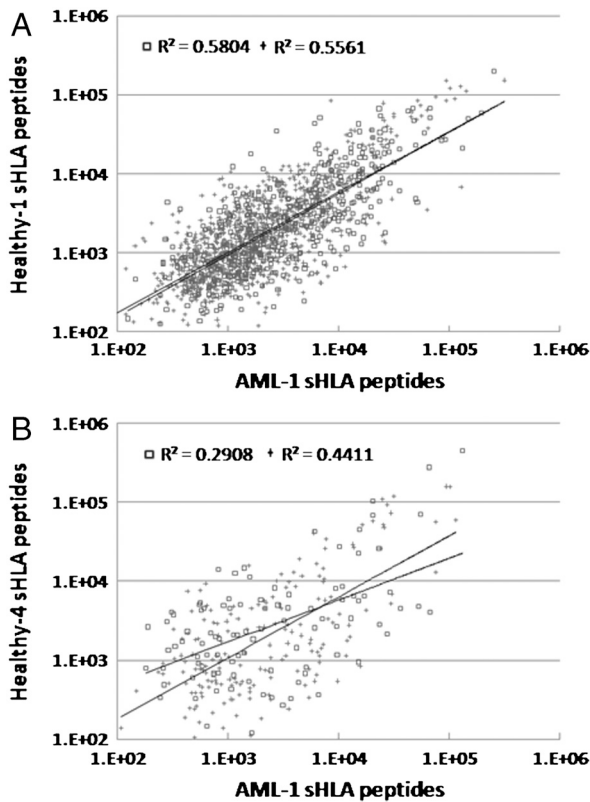
The current study also provides a unique point of view on the biology of the tumor cells within the body, their interaction with the immune system and with their microenvironment. The patients' cancer cells HLA peptidome are continuously scrutinized by the immune system, which unfortunately fails too often to eradicate the disease. An escape mechanism from immune surveillance by antitumor lymphocytes (T and NK cells) was already attributed to the presence of elevated levels of sHLA (24–28) (reviewed in refs. 21, 23), and the method described here may help identify the peptides involved with this escape mechanism.

Cancer cells are thought to express and rapidly degrade many short-lived proteins and defective ribosome products. Therefore, their HLA peptidomes are thought to be largely derived from such rapidly turning-over cellular proteins (29). These proteins



**Fig. 5.** Correlation between LC-MS signal intensities of the sHLA peptidomes collected on separate days from the same donor. The peptides extractions and analyses were performed independently. The sHLA peptides isolated from the plasma, taken on consecutive days, of healthy donors [healthy 5 (A), and healthy 6 (B)] and from one plasma sample of healthy 5 divided into two (C). Square dots indicate high-score identified peptides with Pep-Miner score above 85, cross dots indicate intermediate-score peptides with Pep-Miner score of 70 to 85.

are often too short lived to be detected in the cancer tissues by regular proteomics methodologies, but may become detectable through their proteolysis products, namely the HLA peptides (30). No major differences in the GO analysis results were observed between the sHLA of the healthy controls and of the cancer patients and between the membranar and the soluble HLA peptidomes (Fig. S2). These results are actually expected, because the HLA peptidomes contain large numbers of different peptides, derived from most of the cellular proteome and the cancer-related peptides are present only in small numbers, even among the diseased people's HLA peptidomes, and even these are distributed among different GO groups.



**Fig. 6.** Correlation between the LC-MS signal intensities of shared sHLA peptides isolated from different donors: correlation between sHLA the peptidomes of AML-1 and healthy 1 (A) and AML-1 and healthy 4 (B). Square dots indicate high-score identified peptides with Pep-Miner score above 85, cross dots indicate intermediate-score peptides with Pep-Miner score of 70 to 85.

The sHLA peptidome analyses resulted in the identification of thousands of peptides, among them some potential disease biomarkers. Selected subsets of disease associated HLA peptides should be derived from known cancer-related proteins, expressed exclusively, or in larger amounts, in the cancer cells and not in healthy tissues. Yet, due to the large polymorphism of the HLA haplotypes, confirmation that selected sHLA peptides may indeed serve as cancer biomarkers can be achieved only after analysis of the sHLA peptidomes from hundreds or even thousands of cancer patients and compared to healthy controls.

Most commonly studied tumor antigens are the tumor testis antigens, embryonic cancer antigens, products of known tumor associated proteins, and others (31, 32) (a few examples in Table 1). Examples for such peptides observed in this study include GTPAGGGFPR from Butyrate response factor 1 (TIS11b), which is known to induced myeloid cell proliferation in response to granulocyte colony-stimulating factor and to be fused with AML-1 in leukemic cells (33, 34). Another example is the peptide SVSNVVV-ITK, derived from epidermal growth factor receptor pathway substrate 15 (EPS15), which is often found fused with MLL gene in different leukemia (35). In addition, out-of-context expressed antigens, such as peptides derived from testis antigens, such as the peptide YPFIIDAHL derived from testis-expressed sequence 10 protein, detected in the sHLA peptidome of a female cancer patient and in the sHLA peptidome of a healthy male, are yet another group of potential biomarkers. Furthermore, even peptides normally found in the sHLA peptidomes of healthy people can become useful disease biomarkers if their levels are significantly elevated. The detection of numerous cancer-related sHLA peptides, released by cultured cancer cells (Table S7), is assumed to be caused merely due to the prolonged “life-in-plastic” of the cultured cell lines, which leads to expression of many aberrant antigens, likely of no medical significance, because they are not expressed in the authentic tumors. Thus, in the future, the identification of sHLA cancer-related antigens from the plasma of cancer patients may even provide candidate antigens for development of immunotherapeutics.

The strong similarities between sHLA peptidomes, taken on separate days from the same donors (both in the repertoires of the identified peptides and in their relative LC-MS signal intensities) point to the reproducibility of the technology, thus raising confidence in its potential usefulness for clinical uses (Fig. 5 and Table S4). Another level of knowledge about the individual sHLA peptidome may be provided by tissue typing of their HLA haplotypes (defining the repertoire of each person HLA alleles). Although the HLA alleles of the human population are very polymorphic, many of the alleles can be grouped into several supertypes according to their common consensus peptide binding motifs (recently updated in ref. 22). Thus, specific HLA peptides may be compared between patients and healthy controls harboring HLA alleles belonging to the same supertypes. Indeed, in this study, the HLA peptidomes of the different donors fitted subsets of HLA supertypes, which were later confirmed by tissue typing, to include the donors’ HLA alleles (examples in Table 2).

**Table 3. Measurement of similarity between LC-MS/MS analyses**

Donor	Peptides identified only in		Shared peptides	% of shared peptides
	sHLA	mHLA		
A				Out of sHLA peptides
AML-1	1060	2271	1691	61
AML-3	218	628	581	73
ALL-1 blood	141	642	859	86
ALL-1 bone marrow	178	541	1112	86
B	Blood	Bone marrow		Out of bone marrow
MM-1	736	236	859	78
MM-2	806	134	790	85
ALL-1	48	93	731	89
Healthy 4	1026	327	1269	79
C	Day 1	Day 2		Out of total peptides
Healthy 5	107	199	2233	88
Healthy 6	51	84	1450	91
Healthy 5	Day 1 sample #1	Day 1 sample #2		
	203	143	1946	85

The percentage of similarity is the number of the shared peptides, out of the sHLA peptidome, between sHLA and mHLA peptidomes (A), the shared peptides, out of the bone marrow sHLA peptidome, between the plasma sHLA-peptidomes from blood and from bone marrow (B), and for the reproducibility assay, the number of shared pairs divided by the sum of shared pairs and the peaks that could not be paired (C).

Because the analysis of sHLA peptidomes is relatively rapid and simple, selected cancer-related sHLA peptides may be used for designing patient specific immunotherapeutics and as a potential source for surrogate disease biomarkers. The sHLA peptidome data may even be used for personalizing the treatments, by exploiting the vast information content available in the patients' sHLA peptidomes. In the future, healthy individuals may prefer to define their "normal" sHLA peptidomes, to be followed by periodical analyses aimed at detecting the telltale changes associated with the onset of diseases. Mass spectrometry analysis is becoming less expensive and time consuming, with potential to be transformed into more routine clinical diagnosis tool. Yet, as sHLA peptidome analyses will become more commonplace, new technologies and instrumentations will be needed to reduce the processing time and to increase sensitivity. Future diagnosis based on sHLA peptidomics may also rely on protein arrays, composed of recombinant T cell receptors (TCR) (36) or TCR-like antibodies (37). Such arrays, specific to selected sHLA-peptide complexes, can provide a higher throughput and a lower cost alternative to mass spectrometry.

It can be concluded that the plasma sHLA peptidomics approach described here provides a source of information about the tumor cells within the human body, possibly even as rich in its information content as cancer transcriptomics, proteomics, or metabolomics. It establishes a foundation for a unique diagnostic approach for diseases such as cancer, autoimmunity, allergy, and viral diseases. Yet, it is expected that the ultimate clinical validation of sHLA peptides as useful tumor biomarkers will only be possible after analyses of sHLA peptidomes of large cohorts of patients and healthy controls.

## Materials and Methods

**Cell Lines and Antibodies.** The human MM cancer cell lines RPMI8226, U266, and the anti-HLA class I (W6/32) hybridoma were obtained from the American Type Culture Collection. The cells were maintained in RPMI medium 1640 with 2 mM L-glutamine, 20% heat-inactivated FCS, and 1% penicillin-streptomycin. The W6/32 monoclonal antibodies were affinity purified from mouse ascites fluid using Protein-A Sepharose (Amersham Biosciences).

**Patient Characterization.** Peripheral blood (PB) and bone marrow samples from MM, AML, and ALL patients, peripheral blood from healthy controls, or leukopheresis samples from AML patient (Rambam Hospital Institutional Review Board Committee, Study No. 2183) were provided by the Department of Hematology and Bone Marrow Transplantation, Rambam Hospital. The HLA alleles for the patients and healthy controls, the gender, age, and stage of disease are listed in Table S1.

**Plasma and Cells Collection.** PB, bone marrow, or leukopheresis samples were cleared of cells by centrifugation for 10 min at  $1,200 \times g$  at room temperature. A 1:1,000 Protease Inhibitor Cocktail (Sigma) was added to the plasma followed by centrifugation at  $4^\circ\text{C}$  for 10 min at  $12,000 \times g$ . The cleared plasma samples were stored at  $-20^\circ\text{C}$  until use for sHLA purification.

AML and ALL cells from the blood sample were layered on a density gradient (Lymphoprep, Axis-shield), and centrifuged at  $1,200 \times g$  for 20 min at  $20^\circ\text{C}$ , the mononucleated cells were collected, washed three times with PBS, and frozen at  $-80^\circ\text{C}$  until use for membranous HLA class I purification.

**Affinity Purification of the HLA Molecules.** Membranous HLA were purified by immunoaffinity from the cells after lysis with 0.25% sodium deoxycholate, 0.2 mM iodoacetamide, 1 mM EDTA, 1:200 Protease Inhibitors Cocktail (Sigma), 1 mM PMSF, 1% octyl- $\beta$ -D glucopyranoside (Sigma) in PBS at  $4^\circ\text{C}$  for 1 h. The lysate was cleared by 30 min centrifugation at  $40,000 \times g$ . HLA-I molecules from cleared lysate, growth medium of cultured cells, or from fresh human plasma were immunoaffinity purified with the mAb W6/32 bound to AminoLink Agarose (Pierce). Binding the W6/32 mAb covalently to AminoLink reduced contamination of the HLA preparation by serum antibodies relative to protein-A columns previously used for the same purpose. The affinity column was washed first with 10 column volumes of lysis buffer, followed by 10 column volumes of 150 mM NaCl, 20 mM Tris  $\cdot$  HCl (buffer A), 10 column volumes of 400 mM NaCl, 20 mM Tris  $\cdot$  HCl, 10 volumes of buffer A again, and finally with seven column volumes of 20 mM Tris  $\cdot$  HCl, pH 8.0. The sHLA-I molecules were eluted at room temperature by

adding 500  $\mu\text{L}$  of 0.1 N acetic acid. Small aliquots of each elution fraction were analyzed by 12% SDS-PAGE and by Western blotting developed with the monoclonal rabbit anti-human HLA class I antibody EP1359Y and the secondary anti-rabbit alkaline phosphatase antibody (A7359, Sigma) to evaluate the yield and purity of the eluted HLA.

**Purification of HLA Peptides.** Fractions containing the HLA molecules were heated to  $95^\circ\text{C}$  for 10 min, cooled on ice for 5 min, and centrifuged at  $18,000 \times g$  1 min to recover the bound peptides. The denatured protein subunits of the HLA molecules and the contaminating antibodies were separated from the small peptides by ultrafiltration, through a 3-kDa cutoff Microcon (Millipore) by centrifugation at  $10,000 \times g$ . The filters were pre-washed with 0.1 N acetic acid and 10% acetonitrile to remove contaminants interfering with the mass spectrometry.

**Peptide Concentration and Purification.** Recovered peptides mixtures (after the ultrafiltration) were concentrated by Micro-Tip reversed-phase columns ( $C_{18}$ , 200  $\mu\text{L}$ , Harvard Apparatus). The  $C_{18}$  tips were washed with 80% acetonitrile in 0.1% TFA, equilibrated with 0.1% TFA, and then loaded with the peptide mix. The tip was then washed by an additional 0.1% TFA volume and the peptides were eluted with 80% acetonitrile in 0.1% TFA. Peptides samples were then concentrated to about 18  $\mu\text{L}$  using vacuum centrifugation.

**Analysis by Capillary Chromatography and Mass Spectrometry.** Recovered HLA peptides were analyzed by  $\mu\text{LC-MS/MS}$  using an Orbitrap<sup>XL</sup> mass spectrometer (Thermo-Fisher) fitted with a capillary HPLC (Eksigent). The peptides were resolved by reversed-phase chromatography on 0.075  $\times$  200-mm fused silica capillaries (J&W) self-packed with Reprosil  $C_{18}$  (Maisch, GmbH, Germany) (38). The peptides were eluted at flow rates of 0.25  $\mu\text{L}/\text{min}$ , with linear gradients of 5–45% acetonitrile and 0.1% formic acid, during 90 min, followed by 15 min at 95% of acetonitrile and 0.1% formic acid. Spectra were collected in the orbitrap mass analyzer using full ion scan mode over the  $m/z$  range 400–2,000, set at 60,000 resolutions. The most intense seven masses from each full mass spectrum, with singly, doubly, and triply charge states, were selected for fragmentation by collision-induced disintegration in the linear ion-trap.

**Database Searches.** Pep-Miner (39) was used for peak-list generation of the  $\mu\text{LC-MS/MS}$  data. The peaks were identified using multiple search engines: Pep-Miner, Proteome Discoverer 1.0 SP1 (Thermo) combining the search results of Sequest (Thermo-Fisher) (40) and Mascot (server 2.2, Matrix Science) (41), using the human part of the Uniprot database (<http://www.uniprot.org>, January, 2009) including 20,332 proteins. The search was not limited by enzymatic specificity, the peptide tolerance was set to 0.01 Da, and the fragment ion tolerance was set to 0.5 Da. Oxidized methionine was searched as a variable modification. Peptides listed in Tables S2–S4 and S7 were filtered as followed: Pep-Miner score of above 85 and a Mascot score above 20 and a Sequest Xcorr above 2 and delta-CN below 0.1, mass below 1,350 Da, and mass accuracy of 0.003 Da. These search filtration parameters were set to achieve a false discovery rate for the HLA peptides of 0.05 using a randomized sequence databank as a decoy.

**HLA Peptides Report.** The peptides identifications based on the MS/MS data were combined into a final report, which also contained the fitness score for the consensus binding motifs for each of the HLA alleles, according to [http://www.bimas.cit.nih.gov/molbio/hla\\_bind/](http://www.bimas.cit.nih.gov/molbio/hla_bind/) (42), the normalization of the retention times of the different MS spectra and the peak areas calculations. As expected, some peptide identities remained inconclusive even though their masses were determined at accuracies better than 3 mDa (for example, peptides NLSQILQV and VQNDTLLQV from cluster number 43026 in Table S2). In order not to omit from the list potentially important peptides, all the different peptide sequences that fit the same observed mass and the same MS/MS fragmentation patterns are included in the list.

The retention times of all the HPLC runs were normalized to a reference run. First, the MS/MS spectra of two runs were clustered by spectrum similarity, whereas all spectra in a resulting cluster are assumed to be derived from the same peptide (39), and only clusters that include at least one spectrum from each run were normalized. Each MS/MS spectrum was mapped to the MS peak that contained its precursor ion. The retention time of the peak was used rather than the MS/MS retention time, which is in an arbitrary time point along the peak. Then the two vectors of retention times were normalized using a linear fit.

Peak masses were calibrated using the 445.1200 Da polydimethylcyclohexanes ion (43). Peak envelopes were identified in the three-dimensional MS space (time  $\times$  mass  $\times$  intensity). Mass/charge ( $m/z$ ) signals that appear in ad-

acent MS scans and did not differ from each other by more than 0.005 Da were used for calculating the peak volume, which was the sum of intensities of the  $m/z$  that were at least 30% as high as the largest  $m/z$  in the peak envelope.

**Measurement of Similarities Between HLA Peptidomes.** The purpose of this measurement was to assess the degree of similarity between two HLA peptidomes, analyzed by independent  $\mu$ LC-MS/MS runs. The similarity was based on matching pairs of related  $\mu$ LC-MS peaks that belong to the same peptides in the two runs. In order to take into consideration only peptides, and not contaminants, peaks, in which at least one MS/MS spectrum was identified by Pep-Miner, with score of at least 70, were included. Peaks that could be paired with the identified peaks, in terms of mass (mass difference of up to 5 ppm) and normalized retention time (difference of up to 0.25 min), were included, even if they were not identified or fragmented.

- Liotta LA, Petricoin EF (2006) Serum peptidome for cancer detection: Spinning biologic trash into diagnostic gold. *J Clin Invest* 116:26–30.
- Lopez MF, et al. (2005) High-resolution serum proteomic profiling of Alzheimer disease samples reveals disease-specific, carrier-protein-bound mass signatures. *Clin Chem* 51:1946–1954.
- Lowenthal MS, et al. (2005) Analysis of albumin-associated peptides and proteins from ovarian cancer patients. *Clin Chem* 51:1933–1945.
- Villanueva J, et al. (2006) Serum peptidome patterns that distinguish metastatic thyroid carcinoma from cancer-free controls are unbiased by gender and age. *Mol Cell Proteomics* 5:1840–1852.
- Villanueva J, et al. (2006) Differential exoprotease activities confer tumor-specific serum peptidome patterns. *J Clin Invest* 116:271–284.
- Rammensee HG (2006) Peptides made to order. *Immunity* 25:693–695.
- Purcell AW, McCluskey J, Rossjohn J (2007) More than one reason to rethink the use of peptides in vaccine design. *Nat Rev Drug Discov* 6:404–414.
- Admon A, Barnea E, Ziv T (2003) Tumor antigens and proteomics from the point of view of the major histocompatibility complex peptides. *Mol Cell Proteomics* 2:388–398.
- Albitar M, et al. (2007) Levels of soluble HLA-I and beta2M in patients with acute myeloid leukemia and advanced myelodysplastic syndrome: Association with clinical behavior and outcome of induction therapy. *Leukemia* 21:480–488.
- Albitar M, et al. (2007) Clinical relevance of soluble HLA-I and beta2-microglobulin levels in non-Hodgkin's lymphoma and Hodgkin's disease. *Leuk Res* 31:139–145.
- Leleu X, et al. (2005) Total soluble HLA class I and soluble HLA-G in multiple myeloma and monoclonal gammopathy of undetermined significance. *Clin Cancer Res* 11:7297–7303.
- Nocito M, Montalban C, Gonzalez-Porque P, Villar LM (1997) Increased soluble serum HLA class I antigens in patients with lymphoma. *Hum Immunol* 58:106–111.
- Shimura T, et al. (2001) Clinical significance of soluble form of HLA class I molecule in Japanese patients with pancreatic cancer. *Hum Immunol* 62:615–619.
- Tsuchiya N, Shiota M, Yamaguchi A, Ito K (1996) Elevated serum level of soluble HLA class I antigens in patients with systemic lupus erythematosus. *Arthritis Rheum* 39:792–796.
- Moore C, et al. (1997) Elevated levels of soluble HLA class I (sHLA-I) in children with severe atopic dermatitis. *Ann Allerg Asthma Im* 79:113–118.
- Migliaresi S, et al. (2000) Increased serum concentrations of soluble HLA-class I antigens in hepatitis C virus related mixed cryoglobulinaemia. *Ann Rheum Dis* 59:20–25.
- Schutt P, et al. (2008) The clinical significance of soluble human leukocyte antigen class-I, ICTP, and RANKL molecules in multiple myeloma patients. *Hum Immunol* 69:79–87.
- Demaria S, Bushkin Y (2000) Soluble HLA proteins with bound peptides are released from the cell surface by the membrane metalloproteinase. *Hum Immunol* 61:1332–1338.
- Demaria S, Schwab R, Gottesman SR, Bushkin Y (1994) Soluble beta 2-microglobulin-free class I heavy chains are released from the surface of activated and leukemia cells by a metalloprotease. *J Biol Chem* 269:6689–6694.
- Krangel MS (1986) Secretion of HLA-A and -B antigens via an alternative RNA splicing pathway. *J Exp Med* 163:1173–1190.
- Tabayoyong WB, Zavazava N (2007) Soluble HLA revisited. *Leuk Res* 31:121–125.
- Sidney J, et al. (2008) HLA class I supertypes: A revised and updated classification. *BMC Immunol* 9:1.
- Campoli M, Ferrone S (2008) Tumor escape mechanisms: Potential role of soluble HLA antigens and NK cells activating ligands. *Tissue Antigens* 72:321–334.

Gene ontology analysis was done by the PANTHER Classification System. For each donor, the peptidome containing the largest number of identified HLA peptides was selected for this analysis.

The data associated with this manuscript may be downloaded from [proteomecommons.org](http://proteomecommons.org) Tranche using the following hash: PAA8GYJy25k9X4z8klk9g91S1uiFLpQT27+Txn47ubkzIVlxjO5rLZN86JftettKKOTUSHYJQu5RYiyXb+Fc/SIoAAAAAAADEQ==. The hash may be used to prove exactly what files were published as part of this manuscript's data set, and the hash may also be used to check that the data have not changed since publication.

**ACKNOWLEDGMENTS.** This work was supported by the Israel Science Foundation Grant 916/05 and by the Greta Koppel Small Cell Lung Carcinoma fund, Mitchell Family Foundation, and Israel Cancer Association. This approach was filed as a PCT/IL2010/000212.

- Jinushi M, et al. (2005) Impairment of natural killer cell and dendritic cell functions by the soluble form of MHC class I-related chain A in advanced human hepatocellular carcinomas. *J Hepatol* 43:1013–1020.
- Puppo F, et al. (2000) Soluble human MHC class I molecules induce soluble Fas ligand secretion and trigger apoptosis in activated CD8(+) Fas (CD95)(+) T lymphocytes. *Int Immunol* 12:195–203.
- Spaggiari GM, et al. (2002) Soluble HLA class I molecules induce natural killer cell apoptosis through the engagement of CD8: Evidence for a negative regulation exerted by members of the inhibitory receptor superfamily. *Blood* 99:1706–1714.
- Zavazava N, Kronke M (1996) Soluble HLA class I molecules induce apoptosis in alloreactive cytotoxic T lymphocytes. *Nat Med* 2:1005–1010.
- Bangia N, Ferrone S (2006) Antigen presentation machinery (APM) modulation and soluble HLA molecules in the tumor microenvironment: Do they provide tumor cells with escape mechanisms from recognition by cytotoxic T lymphocytes? *Immunol Invest* 35:485–503.
- Yewdell JW, Anton LC, Bennink JR (1996) Defective ribosomal products (DRiPs): A major source of antigenic peptides for MHC class I molecules? *J Immunol* 157:1823–1826.
- Milner E, Barnea E, Beer I, Admon A (2006) The turnover kinetics of major histocompatibility complex peptides of human cancer cells. *Mol Cell Proteomics* 5:357–365.
- Novellino L, Castelli C, Parmiani G (2005) A listing of human tumor antigens recognized by T cells: March 2004 update. *Cancer Immunol Immunother* 54:187–207.
- Parmiani G, De Filippis A, Novellino L, Castelli C (2007) Unique human tumor antigens: Immunobiology and use in clinical trials. *J Immunol* 178:1975–1979.
- Baou M, et al. (2009) Involvement of Tis11b, an AU-rich binding protein, in induction of apoptosis by rituximab in B cell chronic lymphocytic leukemia cells. *Leukemia* 23:986–989.
- Shimada H, et al. (2000) Analysis of genes under the downstream control of the t(8;21) fusion protein AML1-MTG8: Overexpression of the TIS11b (ERF-1, cMG1): Gene induces myeloid cell proliferation in response to G-CSF. *Blood* 96:655–663.
- Meyer C, et al. (2009) New insights to the MLL recombinome of acute leukemias. *Leukemia* 23:1490–1499.
- Li Y, et al. (2005) Directed evolution of human T-cell receptors with picomolar affinities by phage display. *Nat Biotechnol* 23:349–354.
- Cohen CJ, et al. (2003) Direct phenotypic analysis of human MHC class I antigen presentation: Visualization, quantitation, and in situ detection of human viral epitopes using peptide-specific, MHC-restricted human recombinant antibodies. *J Immunol* 170:4349–4361.
- Ishihama Y, Rappsilber J, Andersen JS, Mann M (2002) Microcolumns with self-assembled particle frits for proteomics. *J Chromatogr A* 979:233–239.
- Beer I, Barnea E, Ziv T, Admon A (2004) Improving large-scale proteomics by clustering of mass spectrometry data. *Proteomics* 4:950–960.
- Eng J, McCormack A, Yates J (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectr* 5:976–989.
- Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20:3551–3567.
- Parker KC, Bednarek MA, Coligan JE (1994) Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *J Immunol* 152:163–175.
- Andreas Schlosser RV-E (2003) Volatile polydimethylcyclosiloxanes in the ambient laboratory air identified as source of extreme background signals in nanoelectrospray mass spectrometry. *J Mass Spectrom* 38:523–525.
- Rammensee H, et al. (1999) SYFPEITHI: Database for MHC ligands and peptide motifs. *Immunogenetics* 50:213–219.