Journal for ImmunoTherapy of Cancer

Soluble HLA peptidome of pleural effusions is a valuable source for tumor antigens

Sofia Khazan-Kost,[1] Gal Cafri,[2] Dganit Melamed Kadosh,[1] Navit Mooshayef,[3] Sumit Chatterji,[3] Dan Dominissini,[4,5] Sigal Manor,[6] Bracha Zisser,[6] Limor Broday,[7] Efrosiniia Talalai,[3,5] Anat Shemer,[2] Oranit Zadok,[8] Efrat Ofek,[9] Amir Onn,[3,5] Arie Admon [ORCID],[1] Michael Peled[3,5]

Correspondence to
Professor Arie Admon; admon@technion.ac.il

Dr Michael Peled; Michael.peled@sheba.health.gov.il

ABSTRACT

Background Soluble human leucocyte antigen (sHLA) molecules, released into the plasma, carry their original peptide cargo and provide insight into the protein synthesis and degradation schemes of their source cells and tissues. Other body fluids, such as pleural effusions, may also contain sHLA-peptide complexes, and can potentially serve as a source of tumor antigens since these fluids are drained from the tumor microenvironment. We explored this possibility by developing a methodology for purifying and analyzing large pleural effusion sHLA class I peptidomes of patients with malignancies or benign diseases.

Methods Cleared pleural fluids, cell pellets present in the pleural effusions, and the primary tumor cells cultured from cancer patients' effusions, were used for immunoaffinity purification of the HLA molecules. The recovered HLA peptides were analyzed by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) and the resulting LC-MS/MS data were analyzed with the MaxQuant software tool. Selected tumor antigen peptides were tested for their immunogenicity potential with donor peripheral blood mononuclear cells (PBMCs) in an in vitro assay.

Results Mass spectrometry analysis of the pleural effusions revealed 39,669 peptides attributable to 11,305 source proteins. The majority of peptides identified from the pleural effusions were defined as HLA ligands that fit the patients' HLA consensus sequence motifs. The membranal and soluble HLA peptidomes of each individual patient correlated to each other. Additionally, soluble HLA peptidomes from the same patient, obtained at different visits to the clinic, were highly similar. Compared with benign effusions, the soluble HLA peptidomes of malignant pleural effusions were larger and included HLA peptides derived from known tumor-associated antigens, including cancer/testis antigens, lung-related proteins, and vascular endothelial growth factor pathway proteins. Selected tumor-associated antigens that were identified by the immunopeptidomics were able to successfully prime CD8+ T cells.

Conclusions Pleural effusions contain sHLA-peptide complexes, and the pleural effusion HLA peptidome of patients with malignant tumors can serve as a rich source of biomarkers for tumor diagnosis and potential candidates for personalized immunotherapy.

Key messages

What is already known on this topic

⇒ Metastatic lung cancer is a deadly disease, whose treatment can benefit from T cell immunotherapy and anticancer vaccines. However, the required analysis of human leucocyte antigen (HLA) peptidome by mass spectrometry depends on the availability of a large amount of tumor cells, while metastatic lung tumors are usually biopsied only by a needle. Since lung cancer is often accompanied by accumulation of pleural effusion that contains soluble HLA molecules, we developed a methodology to purify and analyze the malignant pleural effusion soluble HLA class I peptidomes, aiming to discover biomarkers and immunotherapeutic targets.

What this study adds

⇒ We have found that very large soluble HLA-peptidomes can be established from pleural effusions. These immunopeptidomes include many tumor antigens and immunogenic peptides, which can be applied to improve immunotherapeutics design, diagnostics, prognostics, and personalized treatment.

BACKGROUND

Lung cancer is a leading cause of cancer-related deaths worldwide, with most patients diagnosed at late stages, when survival is significantly shorter. Historically, these late-stage lung cancer patients were treated with chemoradiotherapy, which delays disease progression in some cases. New biological treatments, targeting driver mutations and immune checkpoints, have entered the arena, providing some hope of extending survival.[1 2] Antitumor vaccination with immunogenic peptides, or their encoding RNA or DNA sequences, including tumor-associated antigens (TAA), cancer/testis antigens (CTA), and neoantigens, is a potentially useful personalized immunotherapy

modality. CTA are a subset of TAA that are normally expressed in germline cells, embryonal cells, and adult immune-privileged tissues. They are generally not expressed in adult somatic cells, but are often present in tumor cells, and therefore, may elicit beneficial immune responses when administered as personalized immunotherapeutics.[3–6] To elicit effective anticancer T-cell immune responses, peptides derived from tumor antigens must be presented in the context of human leucocyte antigen (HLA) molecules. The HLA class I molecules are cell surface glycoproteins expressed on most nucleated cells, which present proteolytic products of both normal and abnormal cellular proteins to the immune cells. Due to the large polymorphism in the HLA genes, each HLA allomorph binds and presents a unique repertoire composed of thousands of different peptides.[7] Identification of the HLA peptidomes, using affinity purification of the HLA molecules followed by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) analysis of their bound peptides, can lead to the discovery of tumor antigens potentially useful for clinical treatments.[4 6 8 9] In addition to their membranal form, soluble HLA molecules (sHLA) are secreted into the plasma,[10] and since they are loaded with their original cargo peptides, they provide an opportunity for large-scale analysis of their bound peptidomes.[11] Indeed, plasma sHLA peptidomes have been found to be a rich source of tumor antigens, potentially useful for cancer diagnosis and immunotherapy.[11–13]

Pleural effusion is an abnormal excess of pleural fluid that accumulates within the pleural cavity in both malignant and benign diseases. In healthy individuals, about 10–20 mL of pleural fluid is present in the pleural cavity. In disease states, due to the imbalance of the hydrostatic and osmotic pressures in blood vessels in the parietal pleura, increased permeability, or lymphatic congestion (or a combination of these factors), pleural fluid can accumulate in large volumes.[14 15] Malignant pleural effusion, which is defined by the presence of tumor cells in the fluid, is observed in up to 50% of advanced non-small cell lung cancer (NSCLC) cases and is tightly associated with poor prognosis.[14] Since malignant fluids contain soluble molecules secreted by the tumor, stromal, and immune cells, their content represents the tumor microenvironment (TME). Hence, both soluble factors and tumor cells can be isolated and serve as useful sources for diagnostic/prognostic biomarkers and therapeutic targets.[16 17] Although large volumes are removed from the patients to relieve the dyspnea caused by the pleural effusion, only small fractions are needed for pathology examinations, and the rest can be used for additional analyses. While soluble HLA molecules have already been identified in pleural effusions,[18] their bound peptides cargo has not been studied yet. This work aimed to determine whether HLA molecules can be recovered from pleural effusions in sufficient quantities for analysis of their peptidomes and to look among the identified peptidomes for TAA and CTA associated peptides that can serve as targets for

immunotherapy and as potentially useful biomarkers, for improving lung cancer treatment and diagnosis.

MATERIALS AND METHODS
Sample collection and handling
Pleural fluids of 14 cancer and benign disease patients (table 1) were collected at Chaim Sheba Medical Center (Ramat Gan, Israel). Fresh cells were separated from the effusions by centrifugation (20 min, 800 ×g, at 4°C). All samples were snap-frozen and stored at −80°C until use. To establish primary cells lines, 10^7 cells were isolated from the pleural fluids and cultured in RPMI medium supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin-amphotericin B on rat type I collagen-coated plates (Corning Life Science, Tewksbury, Massachusetts, USA) in a humidified incubator, at 37°C in 5% CO_2.

Immunofluorescence staining
Cells were fixed with 4% paraformaldehyde and immunofluorescence staining was performed according to Cell Signaling (Danvers, MA, USA) protocol. Samples were incubated (overnight, 4°C) with the following primary antibodies: antiepithelial cell adhesion molecule (EpCAM) VL1D9 (Cell Signaling #2929; 1:800) and antithyroid transcription factor 1 (TTF1) (Cell Signaling #1273; 1:50), rinsed, and then incubated with anti-rabbit Alexa Fluor 488 or anti-mouse Alexa Fluor 594 (1:400, 2 hours, room temperature). 4',6-diamidino-2-phenylindole (DAPI) (1 µg/mL) was added during the final washes after incubation with the secondary antibodies. Immunofluorescence was visualized with a Nikon eclipse 80i microscope and imaged with DS-Fi3 digital camera.

Tissue processing, TTF1 immunohistochemistry and scoring
Cells' pellets were collected to Eppendorf tubes and suspended in 1 mL 4% paraformaldehyde. Tubes were placed at 4°C for overnight cell fixation and then centrifuged at 200 g for 5 min. After supernatant was removed, cell pellet was resuspended in 30–50 µL melted bio agar (05–9803S, Bio-Optica, Milano, Italy) and mixture was placed in small rings for a few minutes freezing. The formed round shaped gel was placed in a cassette in 4% paraformaldehyde jar and later on, was processed on automated tissue processing machine (Tissue-Tek VIP, Sakura Finetek, Torrance, California, USA). The processing started with dehydration of the tissue using a series of increasing concentrations of alcohol (70%, 80%, 90%, 95%, and 100%) and then removal of the dehydrant with Xylene. Finally, the tissue was infiltrated with paraffin and manually placed into a block. Tissues embedded in paraffin were sectioned into 3.5 µm slices and mounted on microscope slides. The slides were placed in an oven for 1 hour at 60 before proceeding to the staining steps. H&E stains were performed automatically on a ST5020 device (Leica Biosystems, Wetzlar, Germany) according to manufacturer's instruction. Immunohistochemistry

Table 1 Patients' characteristics

Patient	Age/gender	Disease	HLA-A	HLA-B	HLA-C	Sample	Total no of identified peptides*	No of putative HLA ligands†	% of HLA ligands
P1	68/M	Heart failure	26:01/68:01	38:01/50:01	06:02/12:03	Pleural fluids	1718	1385	81
						Bulk cells	893	315	35
P2	82/M	Lung adenocarcinoma	01:01/26:01	41:01/38:01	12:03/17:01	Pleural fluids	3001	2704	90
						Bulk cells	3183	2690	85
P3	42/F	Lung adenocarcinoma	03:01/33:01	18:01/53:01	04:01/05:01	Pleural fluids	4358	3800	87
						Bulk cells	4947	4054	82
						Cultured cells	3968	3606	91
P4	75/F	Endometrioid carcinoma	02:01/11:01	40:01/44:02	03:04/05:01	Pleural fluids	1350	1118	83
						Bulk cells	3886	3440	89
P5	92/M	Heart and kidney failure	02:01/02:01	08:01/39:01	07:02/07:02	Pleural fluids	672	458	68
						Bulk cells	1647	990	60
P6	79/M	Squamous cell carcinoma	02:05/33:01	18:01/50:01	06:02/12:03	Pleural fluids	776	604	78
						Bulk cells	540	122	23
P7	61/M	Heart and kidney failure	02:01/03:02	08:01/51:01	07:02/07:02	Pleural fluids	1187	514	43
P8	64/M	Heart failure	01:01/01:01	58:01/58:01	07:01/07:01	Pleural fluids	1613	1059	66
						Bulk cells	2018	1552	77
P9	73/M	Kidney failure	23:01/31:01	35:02/49:01	04:01/07:01	Pleural fluids	1590	1199	75
						Bulk cells	3641	3012	83
P10	92/M	Lung adenocarcinoma	25:01/32:01	18:01/43:01	04:01/12:03	Pleural fluids	2188	1407	64
						Bulk cells	1354	779	58
P11	60/F	Lung adenocarcinoma	23:01/26:01	38:01/49:01	07:01/12:03	Pleural fluids	6095	4965	81
						Bulk cells	7886	6740	85
						Cultured cells	4696	3670	78
P12	64/F	Lung adenocarcinoma	01:01/32:01	15:17/49:01	07:01/07:01	Pleural fluids	1981	1409	71
						Bulk cells	1218	1105	91
P13	48/F	Lung adenocarcinoma	01:01/24:02	35:02/57:01	04:01/06:02	Pleural fluids	5418	4492	83
						Bulk cells	5945	5146	87
						Medium	1066	719	67
						Cultured cells	5133	4598	90
						Plasma	4004	2955	74
P14	72/F	Lung adenocarcinoma	03:01/11:01	14:02/35:01	04:01/08:02	Pleural fluids	5605	4600	82
						Bulk cells	5477	3726	68
Total no of identified unique peptides							39 669	32 970	83

*The number of peptides identified from the purification of 400 mL pleural effusions.
†Putative ligands are peptides that fit the sequence consensus motifs of the patients' HLA allotypes.
F, female; HLA, human leucocyte antigen; M, male.

stains were performed with TTF1 antibody (1:50, M3575, Dako, Glostrup, Denmark) on a Benchmark XT staining module (Ventana Medical Systems., Oro Valley, Arizona, USA) using iVIEW DAB Detection Kit (catalog #: 760-091, Ventana Medical Systems). After immunostaining, sections were counterstained with hematoxylin (Ventana Medical Systems), rinsed in distilled water and finally dehydrated manually in graded ethanol (70%, 96%, and 100%). Then, the sections were cleared in xylene and covered with Entellan (Surgipath, Eagle River, Wisconsin, USA) for examination by light microscope. The quantification of TTF1+ cells' percentage out of the total cells' population was made based on counts in 10 representative high-power fields.

HLA peptidome affinity purification

HLA molecules were immunoaffinity-purified from the benign and malignant pleural effusions (10–400 mL) and from the available patient plasma sample (10 mL) as previously described.[11 13] Briefly, pleural or plasma fluids were centrifuged for 40 min, at 47,808 ×g at 4°C (Sorvall RC 6 Plus centrifuge, Thermo Fisher Scientific, Waltham, Massachusetts, USA). The HLAs in the supernatant were then subjected to immunoaffinity purification using pan-HLA-A, HLA-B, and HLA-C monoclonal antibodies (W6/32)[19] covalently bound to Protein A-Sepharose beads (A$_2$S Technologies, Yavne, Israel). Beads volume used for the purification varied according to the fluids volume, thus the 10 and 40 mL pleural effusion were purified using 0.2 mL beads, while the 100 and 400 mL of effusions, 0.5 mL beads were used. When HLA-peptide complexes were purified from the larger volumes (100 and 400 mL), the fluids were passed overnight at 4°C through a series of columns connected by tubes to allow for a consistent flow. First, to absorb aggregates, the fluids passed through a 50 mL plastic column containing ~2 mL Sepharose 4 Fast Flow (GE Healthcare, Chicago, Illinois, USA), then, using two sets of 10 mL Poly-Prep Chromatography Columns (BioRad, Hercules, California, USA) with Protein A, and Protein A cross-linked to the W6/32 mAb. All columns were preconditioned with two-column-volumes of 0.1 N acetic acid followed by 100 mM Tris-HCl (pH 8). After passing the fluids, the columns were washed twice with 20 mM Tris-HCl (pH 8) with 400 mM NaCl, followed by another wash with 20 mm Tris-HCl (pH 8). HLA molecules with their bound peptides were then eluted using five-column volumes of 1% trifluoroacetic acid (TFA) and loaded on disposable C18 microcolumns (Harvard Apparatus, Holliston, Massachusetts, USA). The HLA peptides were recovered by passing 30% acetonitrile in 0.1% TFA over the column, while the HLA molecules and associated proteins were eluted with 80% acetonitrile in 0.1% TFA. The peptide fraction was then dried by vacuum centrifugation, reconstituted in a minimal volume of 0.1% TFA, loaded on C18 StageTips[20] and eluted with 80% acetonitrile in 0.1% TFA. Samples were dried again by vacuum centrifugation and reconstituted with 0.1% formic acid. Samples were analyzed by
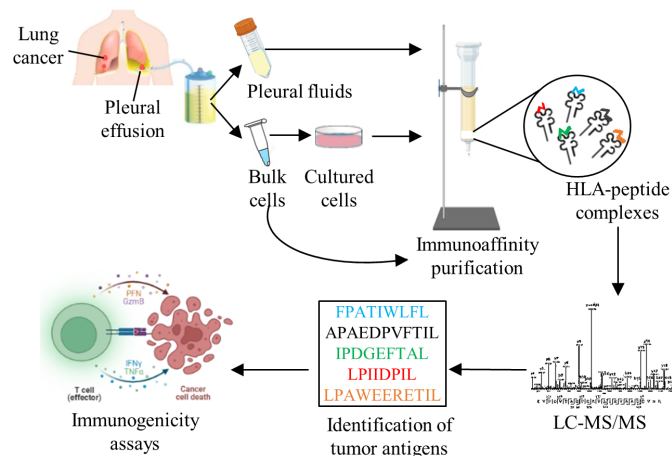

Figure 1 Experimental design. Pleural effusions were collected from cancer and non-cancer patients, cells were separated and cultures were grown, when possible. HLA-peptide complexes were immunoaffinity-purified, peptides were separeted and analyzed by mass spectrometer. Peptides were considered HLA ligands after filtration of known contaminants, and if they were 8–14 amino-acids-long and matched the HLA allele consensus motifs of each patient, based on their tissue typing and online analysis with the NetMHCpan V.4.1 server http://www.cbs.dtu.dk/services/NetMHCpan/. Tumor antigens were identified from the peptides lists and candidates were tested for an immunogenic response. The figure was created using BioRender (https://biorender.com/). HLA, human leucocyte antigen; LC-MS/MS, liquid chromatography and tandem mass spectrometry.

LC-MS/MS. The membranal HLA (mHLA) peptidomes were recovered after solubilizing the cell pellets in lysis buffer containing a mild detergent for 45 min at 4°C before loading on the immunoaffinity column. The lysis buffer contained 0.25% sodium deoxycholate, 0.2 mM iodoacetamide, 1 mM EDTA, 1:200 protease inhibitors cocktail (Sigma-Aldrich, St. Louis, Missouri, USA), 1 mM PMSF, and 1% octyl-D-glucopyranoside (Sigma-Aldrich) in cold PBS. The mHLA were immunoaffinity purified similarly to the soluble HLA, except that 0.1 mL of W6/32-conjugated Sepharose beads were used. For both soluble and membranal experiments, HLA purification efficacy was confirmed by western blot using 10% of the 80% acetonitrile sample (figure 1 and online supplemental figure S1A) with the anti-pan HLA class I antibody (ab126237, Abcam, Cambridge, UK), followed by incubation with peroxidase-conjugated anti-rabbit IgG (Jackson ImmunoResearch, West Grove, Pennsylvania, USA). The chemiluminescence was detected using the WesternBright ECL substrate kit (Advansta, Menlo Park, California, USA) and developed using the ImageQuant LAS4000 digital imaging system (GE Healthcare Life Sciences, Pittsburgh, Pennsylvania, USA).

Tryptic digestion of the HLA heavy chains purified from pleural fluids

To assess the amount of the purified HLA heavy molecules-peptide complexes and the level of copurifying

and interacting proteins, the dried 80% acetonitrile fraction was reconstituted in lysis buffer containing 8 M urea (Sigma-Aldrich), 400 mM ammonium bicarbonate (Sigma-Aldrich) and 10 mM dithiothreitol (Sigma-Aldrich). Half of the sample was reduced at 60°C for 30 min. Carbamidomethylation was performed in the dark for 30 min by 40 mM iodoacetamide (IAA, Sigma-Aldrich). Next, three volumes of HPLC water were added, followed by 0.2 µg trypsin (Promega, Madison, Wisconsin, USA) and an overnight incubation at 37°C. Another aliquot of 0.2 µg trypsin was added and incubated for 3 hours at 37°C. The digested samples were acidified to a final concentration of 0.1% TFA and desalted on C18 StageTips prior to MS analysis.

Mass spectrometry analyses

LC-MS/MS analyses of the HLA-bound peptides were performed with a Q-Exactive-Plus mass spectrometer fitted with an Easy nLC 1000 nano-capillary UHPLC (Thermo-Fisher Scientific) as in,[21] with slight modifications. Peptides were loaded on a reversed-phase capillary column of about 30 cm long, 75 µM inner diameter, home-packed with 3.5 µM silica ReproSil-Pur C-18-AQ resin (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany). Elution was performed with a linear gradient of 5%–35% of acetonitrile in 0.1% formic acid, for 2 hours, at a flow rate of 0.15 µL/min. The top 10 most intense ions in each full-MS spectrum, with single to triple-charged states, were selected for fragmentation by high-energy collision dissociation (HCD). No fragmentation was performed on peptides with unassigned precursor ion charge states, or charge states of four and above. The normalized collision energy was set to 25%. Full-scan MS spectra were acquired at a resolution of 70 000 at 200 m/z, with a target value of 3×10^6 ions. For MS2 the resolution was 17 500 at 200 m/z with a target value 1×10^5. The intensity threshold was set at 6.7×10^3. Maximum injection time of the fragmented masses for both MS1 and MS2 was set to 150 ms. The tryptic peptides from the 80% acetonitrile fraction described above were analyzed in a Q-Exactive-Plus mass spectrometer (Thermo Fisher Scientific) fitted with a capillary HPLC (Dionex UltiMate 3000 RSLCnano Pump). The tryptic peptides were eluted using a linear gradient of 5-28% acetonitrile in 0.1% formic acid. The flow rate of the gradient program was 0.15 µL/min, for 2 hours. The 10 most intense ions from the first scan event, with charge states of double to seven, were fragmented by HCD, at a relative collision energy of 25. Resolution and target values for both MS1 and MS2 were as described for the HLA peptidome. The intensity threshold was set at 1.7×10^4. Maximum injection time of the fragmented masses was set to 20 ms for MS1 and 60 ms for MS2.

Data analysis

Mass spectrometry data were analyzed by MaxQuant V.1.6.10.43 and searched with the Andromeda search engine[22] using the human UniProt database (https://www.uniprot.org/; downloaded on June 2019, containing

74,416 entries). We also included a list of mutations available from some patients (online supplemental table S1) to identify possibly mutated peptides (yet, no mutated peptides were identified). All samples derived from each patient were analyzed together, with the MaxQuant's match-between-runs option allowed per patient. Mass tolerance of 4.5 ppm for the precursor masses and 20 ppm for the fragments were allowed. Methionine oxidation, N-acetylation and carbamidomethylation were accepted as variable modifications. Minimal peptide length was seven amino acids. To maximize the number of identified HLA ligands, both 0.05 and 0.01 false discovery rates (FDR) were applied for the peptide spectrum matches, as in reference 21 (online supplemental figure S1B). Deamidation (Asn) was accepted as an additional variable modification in the 0.01 FDR analysis (online supplemental table S2). The resulting peptide lists were filtered to discard peptides identified in the reverse database, and common contaminants. Peptides were considered potential HLA class I ligands if they were 8–14 amino-acids-long and matched the HLA allele consensus motifs of each patient, based on their tissue typing and online analysis with the NetMHCpan 4.1 server (http://www.cbs.dtu.dk/services/NetMHCpan/).[23] Peptides that were ranked ≤2 were included in the subsequent data analyses. Additionally, the peptide sequences were clustered into motifs using the GibbsCluster (http://www.cbs.dtu.dk/services/GibbsCluster/)[24] and compared with the predicted motifs of the patient's HLA haplotypes. The raw MS files of the trypsin-digested 80% acetonitrile fractions were also analyzed by MaxQuant version 1.6.10.43 and searched with the Andromeda search engine[22] using the human UniProt database (https://www.uniprot.org/; downloaded on June 2019). To improve identification rates, each patient's sample was analyzed separately and the protein sequences of the HLA allomorphs were removed from the database. Only the patient-specific HLA amino acid sequences were inserted as a separate FASTA file obtained from the IPD-IMGT/HLA database (https://www.ebi.ac.uk/ipd/imgt/hla/). The search settings were trypsin-specific, with FDR set to 0.01 and decoy mode to revert. Methionine oxidation and N-acetylation were accepted as variable modifications, while carbamidomethylation was accepted as a fixed modification. Intensity-based absolute quantification was used to quantify the levels of purified HLA-peptide complexes (online supplemental table S3).

Immunogenicity assay

T cells were primed as previously described.[25] Briefly, naïve CD8+ T cells were isolated from the peripheral blood of healthy donor positive for HLA-B*35:02 by depletion of non-naive T cells and NK cells (using biotin-conjugated antibodies against CD45RO, CD56, CD57, and CD244), followed by positive selection using CD8 MicroBeads (Miltenyi Biotech, Bergisch Gladbach, Germany). Isolated naïve CD8+ T cells were incubated with IL-7 overnight. The next day, T cells were washed and mixed with

matured autologous dendritic cells (DCs), pulsed with the investigational peptides (GL Biochem, Shanghai, China) as indicated at 1 μg/mL. Interleukin (IL)-21 (30 ng/mL, Peprotech, Rocky Hill, NJ) was added to the co-culture on day 0 the next morning. IL-7 and IL-15 (5 ng/mL, Peprotech) were added on days 3, 5, and 7. On day 10, cells were restimulated with the original peptide for 2 hours. On day 11, T cells were harvested and analyzed for OX40 and 4-1BB positivity, and media was analyzed for IFNγ secretion by ELISA (Biolegend, San Diego, California, USA).

RESULTS

Development and optimization of pleural effusion soluble HLA peptidome analysis

Large volumes of pleural fluids (~500 mL per patient) were drained from nine cancer patients (mostly NSCLC) and five patients with benign conditions (heart and/ or kidney failure) during routine clinical treatment for shortness of breath. The HLA types, along with age, gender, and disease, are listed in table 1. The membranal HLA-bound peptides of the pelleted (bulk) cells from the pleural effusions and the cleared pleural effusion soluble HLA (peHLA)-bound peptides were extracted and analyzed by LC-MS/MS (figure 1). Primary tumor cell cultures were successfully established in three NSCLC cases (P3, P11 and P13), and their membranal HLA-bound peptides were also analyzed. Immunofluorescence staining to detect TTF1 and EpCAM, known markers for lung adenocarcinoma,[26] validated that the source of the cultured cells was indeed lung adenocarcinoma (figure 2A).

To analyze the soluble peHLA peptidome, we first optimized the pleural effusion volumes in order to maximize the number of identified HLA peptides, while keeping the number of contaminating non-HLA-ligand peptides to a minimum. The soluble peHLA molecules with their bound peptides were immunoaffinity purified from 10, 40, 100, and 400 mL of the pleural effusions. Western blot analysis of the purified peHLA-I confirmed their recovery and demonstrated extraction of HLA molecules containing the extracellular part only (39 kDa) or the entire HLA heavy chain, with their transmembrane and cytoplasmic domains (43 kDa) (figure 2B, online supplemental figure S1A), as previously observed for plasma sHLA.[11 27] The recovery of the HLA molecules from pleural effusions was also confirmed by proteomic analysis of the affinity purified HLA heavy chain fraction[21] (online supplemental table S3). While several hundreds of HLA peptides were identified in the smaller volumes of pleural effusions, thousands more were identified in the larger volumes (online supplemental table S4). The numbers of identified peptides positively correlated with the pleural effusion volumes, increasing by about 2.5-fold when the pleural effusion volumes were increased from 10 mL to 400 mL. Above 90% of the peptides identified in the smaller volumes were also identified in the larger
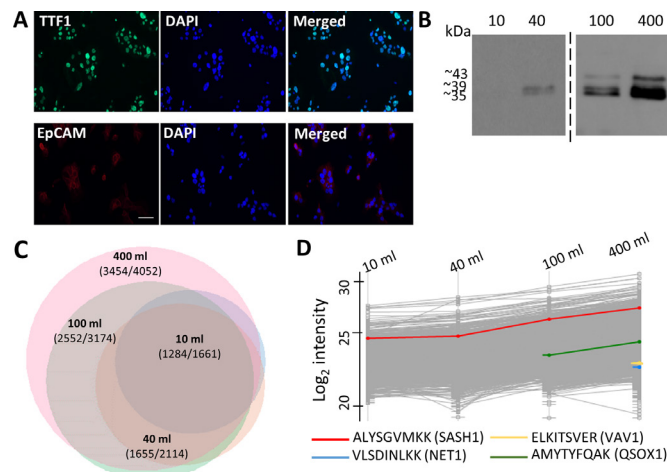

Figure 2 Optimization of pleural effusion volumes for HLA peptidome purification. (A) Primary tumor cells cultured from Patient 3 pleural effusion samples immunostained for TTF1 (green) and EpCAM (red). Nuclei were stained with DAPI (blue). (B) Western blot analysis of 10% of the protein fraction of 10–400 mL pleural effusion HLA purifications from patient 3, stained with anti-pan-HLA antibody. (C) Venn diagram of peptides identified from HLA purification of 10 (light blue), 40 (orange), 100 (green) and 400 (pink) mL of the pleural effusions collected from patient 3. The scaling represents total number of identified peptides. The number of putative HLA ligands and total number of identified peptides appear on the graph. The diagram was produced using the R package VennEuler (V.1.1–0). (D) Profile plot of log$_2$ liquid chromatography-mass spectrometry (LC-MS) intensities of peptides identified from different volumes of pleural effusions. The gray lines indicate the log$_2$ intensities of the identified peptides and the colored lines are selected peptides. EpCAM, epithelial cell adhesion molecule; HLA, human leucocyte antigen; TTF1, thyroid transcription factor 1.

samples (figure 2C, online supplemental table S4). Also, the number and the relative percentage of the identified likely-ligand peptides increased among the peptidomes recovered from the larger volumes (online supplemental figure S1C). Moreover, peptides derived from potential biomarkers of lung cancer, such as neuroepithelial cell transforming 1,[28] Vav guanine nucleotide exchange factor 1 (VAV1),[29] and quiescin sulfhydryl oxidase 1 (QSOX1),[30] were only detected in the larger volumes of pleural effusions (figure 2D). Furthermore, the LC-MS signal intensities of the peptides correlated with the pleural effusion volumes used for the analysis. For example, the SAM and SH3 domain containing 1 (SASH1) peptide (ALYSGVMKK), a potentially prognostic biomarker and therapeutic target of NSCLC,[31] was detected already in the 10 mL sample, but its signal intensity increased with increasing pleural effusion volumes (figure 2D). Thus, to obtain the largest and most informative peptide dataset, subsequent analyses employed peHLA purified from 400 mL pleural fluids.

Large and diverse HLA peptidomes can be recovered from both malignant and benign pleural effusions

Analysis of both soluble and membranal HLA peptidomes from pleural effusion patients with malignant

tumors (n=9) or with benign conditions (n=5) resulted in the identification of 39,669 different peptides, originating from 11,305 source proteins (online supplemental table S5). Of these peptides, 19,294 were identified exclusively in the malignant effusions, while 1784 were identified in the benign effusions only. Overall, the HLA peptides identified in each 400 mL patient sample ranged between ~500 and close to 7000 different peptides (table 1, online supplemental table S5). Furthermore, large increase in the numbers of identified peptides could be obtained by using 0.05 FDR for the peptide spectrum matches[21] in addition to the commonly used 0.01 FDR (online supplemental figure S1B). About 110 peptides were identified with deamidation of Asn, which is indicative of possible loss of N-linked glycosylation. Indeed, 42 of these were observed in known glycosylation sites of proteins (online supplemental table S2). As expected for HLA-I ligands,[7] most of the peptides were nine amino acids long (figure 3A), and most of the peptide sequences matched the consensus sequence motifs of the HLA haplotypes of each patient, as determined by the NetMHCpan[23] (figure 3B, online supplemental figure S2). Notably, some of the identified peptides matched more than one HLA allele and were observed in multiple patients. For example, the sarcoma antigen 1 (SAGE1) peptide (APDNVLLTL) was associated with HLA-B*53:01, HLA-B*35:02, and HLA-B*38:01, and was observed in patients 3, 9, 11, and 13 (online supplemental table S5). Importantly, pleural effusions from cancer patients contained more peHLA peptides than pleural effusions from patients with benign conditions (figure 3B), which may be the result of increased cellular content in the malignant pleural effusions. Indeed, malignant pleural effusions, unlike pleural effusions extracted from patients with benign conditions, are usually exudates, which are clinically defined by high lactate dehydrogenase (LDH), total protein, and albumin concentrations, and often show reduced glucose concentrations secondary to increased cellular content. As expected, positive correlations were observed between the number of soluble peHLA peptides and LDH, total protein, albumin concentrations, and %TTF1+ cells in the pleural effusions (online supplemental figure S3), implying that the numbers of soluble HLA peptides correlate with the total cellular content, specifically with the cancer cells' content of the pleural effusions, and with the patient's medical condition.[32 33]

Since we aimed to identify clinically relevant peptides, it was imperative to ensure consistency between the peHLA peptidomes of different visits of a single patient to the clinic, if the patient's medical condition was stable between these visits. Thus, when possible, samples of pleural effusions were collected during different visits of the same patient, and their peHLA peptidomes were analyzed. Indeed, as shown in figure 3C, high similarities were observed between pleural effusions, even when taken several months apart.

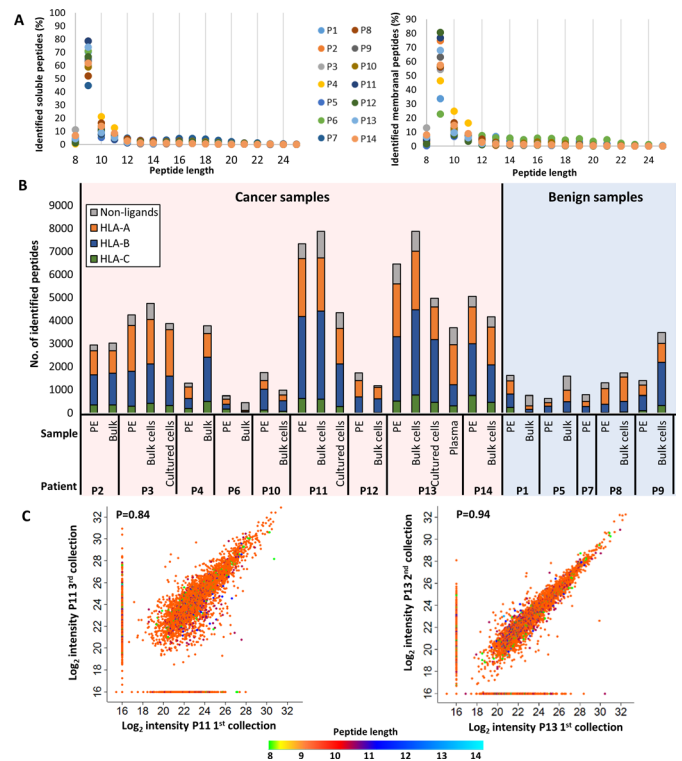While there was great variability between patients, a correlation was noted between the soluble peHLA

Figure 3 Characterization of the identified HLA peptidome. (A) Length distributions of the peptides identified in the pleural fluids (left) and bulk cells (right) of all patients. The different colors of the circles represent different patients. The percentages were calculated for each length before data filtration. (B) Distribution of the HLA peptides to their presenting allotypes. The gray color represents the peptides that are either unassigned to a specific HLA allele by NetMHCpan V.4.1, are known contaminants in MaxQuant, and/or are longer than 14 amino acids. The colorful bars indicate the number of ligands attributed to HLA-A, HLA-B or HLA-C, in orange, blue or green, respectively, according to NetMHCpan V.4.1. (C) Scatter plots of $\log_2$ liquid chromatography-mass spectrometry (LC-MS) intensities of peptides identified from different pleural effusions samples of Patients 11 (left) and 13 (right). The dot colors indicate the peptides length. Pearson correlation value is indicated in each graph. HLA, human leucocyte antigen.

peptidomes, the membranal HLA peptidomes of the pleural fluids cell pellets (bulk), and the membranal HLA peptidomes of the cultured cells from the same patients. Pearson correlations of their relative intensities were mostly above 0.5 (figure 4, online supplemental figure S4). As expected, due to their physical proximity, the HLA peptidomes of both bulk and cultured cells were more similar to the peHLA peptidomes (Pearson correlations of 0.58 and 0.51, respectively) than to plasma sHLA peptidomes (Pearson correlations 0.49 and 0.39, respectively; figure 4).

TAAs were discovered among the peHLA peptidomes

Unique sets of peptides, which originated from proteins that are specific to tumor cells in the lung or to lung tissue, were observed in the peHLA peptidomes but not in the plasma sHLA peptidomes (figure 4). For
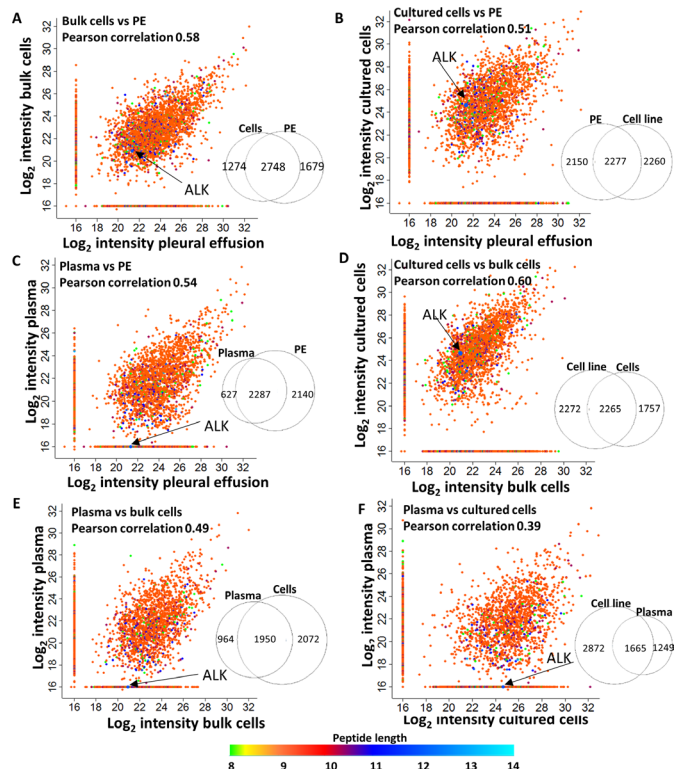
Figure 4 The intensities of the peptides identified in the pleural effusion, bulk cells, cultured cells, and plasma of Patient 13, are correlated to each other. Scatter plots of the log$_2$ LC-MS intensities of the peptides, identified from different sources, are shown. The colors represent the peptide length. Venn diagrams near each plot indicate the number of peptides identified in each sample. Arrows point to the intensities of the peptide MPYPSKSNQEVL, originating from anaplastic lymphoma kinase (ALK). Pearson correlations are indicated in each graph. LC-MS, liquid chromatography and mass spectrometry.
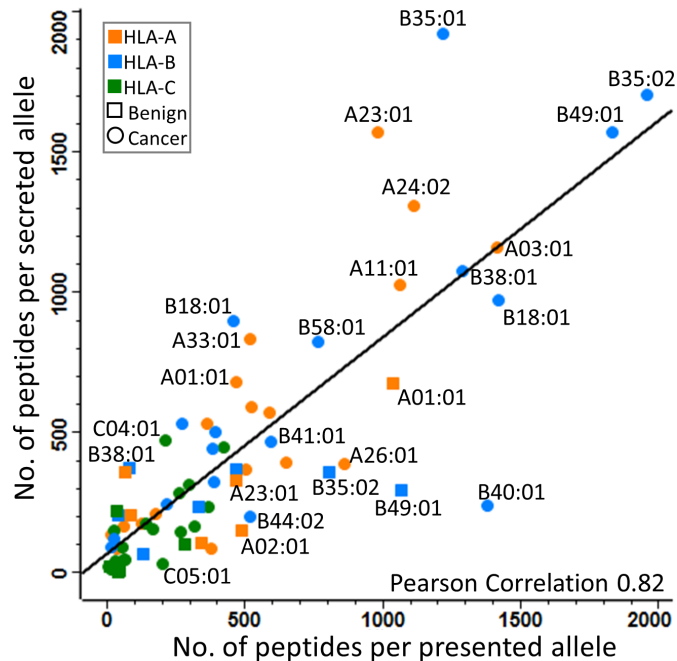

Figure 5 The number of peptides per HLA allotype, presented on the membranal HLA and the soluble HLA, are correlated to each other. The number of peptides associated with specific HLA allotype, from malignant and benign pleural effusions, is represented by circles and squares, respectively. Some of the allotypes are annotated (see online supplemental table S10 for the complete list). HLA, human leucocyte antigen.

example, the HLA-B*35:02 anaplastic lymphoma kinase (ALK) peptide (MPYPSKSNQEVL), was observed in the tumor cells and pleural fluid, but not in the plasma of Patient 13, whose tumor was positive for the echinoderm microtubule-associated protein-like 4-ALK (EML4-ALK) translocation. Furthermore, based on a database of lung-specific genes from the Human Protein Atlas (HPA) (https://www.proteinatlas.org/),[34] we found that relative to the plasma sHLA, the soluble peHLA peptidomes were enriched with HLA peptides derived from lung-specific genes (online supplemental figure S5). Thus, for tumors localized to the lungs, the peHLA soluble peptidome may serve as a useful source for identification of tumor-specific antigens.

The large size of the peHLA peptidomes suggests that they may be used as a source for TAA, as previously suggested for plasma sHLA peptidomes.[11–13] To facilitate identification of TAA and CTA in the peptidomes, a combined database of TAA and CTA was established from six TAA/CTA databases: The CTdatabase (http://www.cta.lncc.br/index.php) and five additional sources taken from recent publications[35–39] (online supplemental

figure S6, online supplemental table S6). Of the 2061 putative unique TAA/CTA genes, a total of 550 peptides, deriving from 322 TAA genes, were detected mostly in the pleural effusions or cells of the cancer patients (online supplemental table S5). Next, the HPA was used to assess the tissue expression patterns of the selected 322 TAA/CTA source genes of peHLA peptides detected here. As 'gold-standard' antigens, we used the well-investigated TAA - CTAG1 (NY-ESO-1) and MAGE Family Member A1 (MAGEA1)[3] antigens, which have very low mRNA and protein expression in most tissues, except for germ-line and immune-privileged tissues, such as testis and placenta. The selected CTA are listed in online supplemental table S7.

To assess the potential immunogenicity of identified HLA class I peptides, we focused on the HLA-B*35:02 allele, which is prevalent in the general population and presents many peptides in cancer Patient 13 in our dataset (figure 5). For peptides selection, we used the in-silico immunogenicity prediction tool IEDB (http://tools.iedb.org/immunogenicity/). Five HLA peptides with the highest IEDB score and/or originating from tumor antigens (online supplemental table S6) were selected for immunogenicity assay from the 3647 HLA ligands of Patient 13 (online supplemental table S8). Peptides immunogenicity was tested in an in vitro priming assay of naïve CD8$^+$ T cells and autologous DCs as stimulators from a healthy HLA-B*35:02 donor, to mimic a vaccination
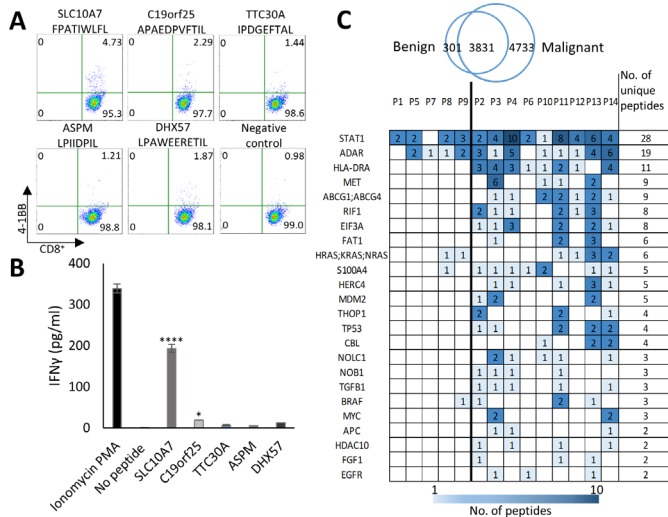
Figure 6 peHLA peptidome originate from tumor-associated genes and includes immunogenic peptides. (A) 4-1BB expression was assayed by FACS analysis of purified naïve CD8+ T cells, from healthy HLA-B*35:02 positive donor, following co-culturing with peptide-loaded autologous DCs. Pseudo-color dot plots display the responses for the five HLA peptides and the negative control (cells stimulated with empty DCs). upper numbers in the quadrants show percentages of CD8+/4-1BB+ cells. (B) The media from restimulated T cells was collected and assessed for the concentration of IFNγ by ELISA. P values were determined by one-way ANOVA test relative to the negative control. n=3, data are represented as mean±SEM *p<0.05, ****p<0.0001. (C) Examples of tumor antigens which are the source of many HLA peptides, including tumor antigens that were identified in the KEGG pathways in cancer database. The antigens are presented with the number of identified peptides derived from each source protein in each patient (see online supplemental table S9 for the complete list). ANOVA, analysis of variance; DC, dendritic cell; HLA, human leucocyte antigen; peHLA, pleural effusion soluble HLA.

scenario in vivo.[25] Two of the tested peptides, FPATIWLFL and APAEDPVFTIL that derived from the tumor antigens SLC10A7 and C19orf25, respectively, yielded an increase in T cells' activation markers (OX40, 4-1bb) and IFNγ secretion (figure 6A,B, online supplemental figure S7), confirming that the identified peptides can be recognized by naïve T cells.

In addition to TAA/CTA, we looked for peptides deriving from known oncogenes and tumor suppressor genes, which are not necessarily very specific to tumors, such as CTA, but are still preferentially expressed in tumors and thus may serve as diagnostic or prognostic markers. One hundred and ninety two different genes, which were the source of a total of 1081 peptides in the peHLA peptidomes, were found in the Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathways in Cancer database (https://www.genome.jp/kegg/pathway.html); of these, peptides from ninety genes were identified here exclusively in cancer patients (online supplemental table S9). Among the source genes identified were oncogenes and tumor suppressors such as epidermal growth factor

receptor (EGFR), proto-oncogene MYC (MYC), tumor protein P53 (TP53), and MET proto-oncogene (MET), each detected in more than two patients (figure 6C). In contrast, peptides from other known potential oncogenes, such as B-Raf proto-oncogene (BRAF) and Kirsten Rat Sarcoma (KRAS), were identified in samples from both cancer patients and patients with benign conditions (figure 6C, online supplemental table S9).

While finding TAA and CTA in tumors is critical for targeted immunotherapy, analyzing the pathways that are enriched in tumor cells and the TME is also therapeutically important, as it points towards the general pathways that should be targeted. Utilizing the ToppGene suite pathway enrichment analysis tool,[40] the vascular endothelial growth factor A (VEGF) pathway was found to be the most prominent pathway among the source proteins of both benign and malignant pleural effusion peptidomes (online supplemental figure S8A-B). Indeed, the VEGF pathway has been implicated in the pathogenesis of pleural effusions.[41] Prominent pathways that were observed exclusively in malignant peHLA peptidomes included the DNA repair pathways (online supplemental figure S8C), while the fat metabolism pathway was prominent in benign effusions (online supplemental figure S8D, online supplemental table S9).

Influence of the HLA haplotypes on the degree of peHLA secretion and numbers of identified peptides

Previous studies have suggested that the identity of the HLA allotype may impact the extent of HLA secretion[11 13] and presentation,[42] and consequently, the number of identified peptides and the potential to identify TAA in body fluids. Here, we noted only a moderate correlation between the number of peptides presented on the membranal HLA and the soluble peHLA (figure 5). This may imply that the HLA allotype may influence the size and diversity not only of the membranal HLA peptidome,[42] but also of the presented soluble peptidome. Indeed, in several allotypes, such as HLA B*35:01 and B*49:01, an increased number of peptides was observed in their soluble relative to their membranal form, suggesting that these allotypes have a higher tendency to be secreted (figure 5). Yet, some secreted HLA allotypes, such as A*24:01, A*23:02, B*35:01, and B*35:02 were observed with significantly larger diversity of bound peptides relative to the levels of secreted HLA heavy chains, while membranal HLAs such as A*01:01, A*02:01 were observed to present lower diversity of peptides in their secreted forms (figure 5 and online supplemental figure S9). Additionally, and as expected, the HLA-C peptidomes were small relative to the HLA-A and HLA-B peptidomes (figures 3B and 5, online supplemental table S10).[43]

DISCUSSION

Personalized immunotherapy is based on the identification of neoantigens and TAA that can be targeted by the immune system. The need for a large number of tumor

cells for detailed HLA peptidome analyses, aimed at identifying neoantigens and TAA,[8 44 45] limits the usefulness of HLA peptidomics in routine clinical practice. Specifically, metastatic lung tumors are usually biopsied only by a needle, and most of the tissue is required for pathological diagnosis. Furthermore, there is no healthy control tissue for comparative analysis. While the limited availability of cancer tissue for HLA peptidome analysis can be partially overcome by culturing tumor cells,[45] or growing tumors in patient-derived xenograft mice,[21] pleural effusions can provide a rich and easily accessible source of information as liquid biopsies. The peHLA peptidomes described here were collected directly from the patients and analyzed immediately thereafter, without requiring any intermediate steps of culturing or grafting, and therefore may provide a more accurate representation of the TMEs and their microenvironments.

Most of the identified peptides were likely authentic HLA ligands that fit the HLA consensus sequence motifs of the patients. The positive correlation found between the membranal HLA peptidome of the cells and the soluble peHLA peptidome, and the similar peHLA peptidomes that were detected in the pleural effusions of the same patient at different time points, further validated the potential usefulness of the peHLA peptidome analyses. Moreover, the number of peptides found in the peHLA peptidomes of fluids correlated with clinical parameters of pleural effusions, like LDH, and albumin concentrations, and with the percentage of cancer cells in the fluids, supporting the hypothesis that soluble HLA secretion is increased in tumors, potentially serving as an escape mechanism.[10 46]

While previous studies have pointed to the use of the plasma sHLA as a source of TAA,[11–13] the presented findings suggest that the soluble peHLA peptidome may be an advantageous source of TAA if the tumor is localized to the lungs. An interesting example is the tyrosine kinase receptor ALK, which is normally expressed primarily during late embryogenesis, but its fusion with other proteins leads to its ectopic expression and to the development of lung cancer. Therefore, ALK has become a drug target with several ALK-targeting drugs already in clinical use.[47] In our study, an HLA ligand, originating from the ALK protein, was identified in Patient 13 pleural fluid, bulk cells, and patient-derived cell line but not in the plasma. Although ALK was not classified as a CTA according to our database search, and therefore may be less suitable for vaccination, in our opinion, its elevated presence in pleural fluids may render it a biomarker for the presence of ALK fusion proteins and set the stage for ALK-targeted treatment. It should be emphasized that the plasma-soluble HLA peptidome may serve as a valuable source for TAA, also for lung cancer, as evidenced by the presence of TAAs both in the pleural fluids of Patient 13, and in his plasma, including SAGE1, PBK and ODF2.

Prioritization of TAA is complicated by the limited information on their differential expression in healthy versus tumor tissues. Neoantigens are expressed specifically by tumor cells and are therefore considered highly immunogenic and safe for clinical use.[48–50] The most useful and safe CTA for immunotherapy are likely those specifically expressed in tumor cells, germline cells, and other immune-privileged tissues.[6] Here, to prioritize TAA, information from six different CTA databases, each with its unique interpretation of tissue expression patterns, tumor specificity, and strategies for CTA definition, was integrated. The databases of Wang et al[35] and Bruggeman et al[37] analyzed mRNA expression patterns of testis-specific and germline-specific genes, while da Silva et al[36] added mass spectrometry-based proteomics data for the identification of CT proteins. Another list we used was of Qi et al[39] that compiled the data from the CTdatabase, HPA, and Cancer Antigenic Peptide database, to create a larger database that was used also for this study. Unfortunately, markedly different lists of putative CTAs are included in these databases, with most of the genes (77%) listed in one database only. Therefore, the present analysis included the entire list of peptides identified from all six TAA databases in order not to miss any putative tumor antigens.

Selection of TAA and HLA peptides for immunogenicity assay, and eventually for immunotherapy, can be based on their preferred expression in tumor cells and immune-privileged tissues, such as germline cells.[6] In addition, the potential immunogenic score of the peptides is highly relevant, and the IEDB database provides a useful measure for this property. The different CTA/TAA databases[35–37] compiled here for this purpose, provide a useful list of genes for selection of such potentially useful immunogenic peptides. In addition, the HPA database[34] provides information on the tissue expression patterns of the different genes and proteins, thus facilitating selection of peptides that have a high expression levels in the immune privileged tissues relative to the others.

Examples of interesting CTA candidates detected here were peptides from PRAME, a putative CTA, which is overexpressed in various cancers and induces specific immune responses in cancer patients,[51] and SAGE1,[52 53] from which two peptides were identified in three cancer patients, and in one patient with renal failure. Although SAGE1 is almost exclusively expressed in male tissues, some of the SAGE1 HLA peptides were identified in this study in female cancer patients (Patients 3, 11, and 13). Indeed, this is a known phenomenon of CTA that has already been described for the HLA peptidome of glioblastoma patients.[13]

Testing of potentially immunogenic peptides using the patients' PBMCs is extremely demanding due to exhaustion of the immune system. Therefore, the preliminary establishment of an immunogenicity potential can exploit donors' PBMCs for in vitro analysis. In this study, five selected peptides were used for the immunogenicity testing, utilizing PBMCs from a healthy donor carrying a matching HLA-B*35:02. The observation of a strong immune reaction to one of the selected peptides supports

the notion that the peHLA peptidome is potentially useful for future clinical uses.

Peptides derived from proteins related to the TME, are not tumor-specific and cannot be used as immunotherapeutic targets, yet, can be exploited as diagnostic and prognostic biomarkers. Examples include HLA peptides from the RAC family, and VAV1, both of which promote metastasis and tumor growth,[54 55] as well as the oncogenes KRAS, BRAF, and EGFR. Patient 3 cell cultures carried peptides deriving from TTF1 (AVYVPFWQK) and EpCAM (YVDEKAPEF), which are used for the pathological diagnosis of lung adenocarcinoma. In line with these findings, Patient 3 tissue stained positive for TTF1 and EpCAM, thus indicating that the peHLA peptidome can indeed be used as liquid biopsy.

Several factors may affect the number of sHLA-peptide complexes found in body fluids, including disease type,[27] the presentation levels and identities of the HLA allotypes. Previous studies have demonstrated a bias in HLA allotype preferences for secretion, such as HLA-A*24, which was the most secreted allotype in leukemia[11] and glioblastoma.[13] The preference for secretion of specific allotypes such as HLA-A*24, may be derived from their high membranal presentation, while the numbers of identified peptides associated with each allotype may relate to the larger diversity of peptides they bind.[42] Here, the number of identified soluble peHLA peptides was found to correlate with the number of membranal HLA peptides of the same patients, implying that generally, the identity of the HLA allotypes influences the numbers of both the membranal and soluble HLA peptides presented. However, a larger panel of samples is needed to assess the impact of HLA haplotypes on peptides presentation and secretion. Some HLA allotypes bind and present larger diversity of peptides relative to others (online supplemental figure S9) possibly due to larger numbers of peptides that fit consensus motif of these HLA allotypes, an issue that warrants further investigation. Cancer cells are known to release larger amounts of soluble HLA to the circulation, possibly as an escape mechanism from immune surveillance.[10 46] The larger numbers of HLA peptides recovered from the pleural effusions of the cancer patients relative to the patients with benign conditions are in par with this conclusion.

In summary, this proof-of-concept study demonstrates the valuable information that can be drawn from immunopeptidome analysis of pleural fluids and cells. Purification of the peHLA-peptidomes is technically simple and can be performed in most clinical laboratories, and the peptide samples can conveniently be sent to central mass spectrometry facilities for analysis. The derived 'big-data' can be combined with molecular profiling of pleural effusion DNA, RNA, proteins, and tumor cells[56–58] to identify tumor antigens that can be applied to improve immunotherapeutics design, diagnostics, prognostics, and personalized treatment.

Author affiliations
[1]Faculty of Biology, Technion Israel Institute of Technology, Haifa, Israel
[2]Chaim Sheba Medical Center, Ramat Gan, Israel
[3]Institute of Pulmonary Medicine, Chaim Sheba Medical Center, Ramat Gan, Israel
[4]Sheba Cancer Research Center, Chaim Sheba Medical Center, Ramat Gan, Israel
[5]Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
[6]Ezer Mizion Bone Marrow Donor Registry, Petah Tikva, Israel
[7]Department of Cell and Developmental Biology, Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
[8]Institute of Oncology, Chaim Sheba Medical Center, Ramat Gan, Israel
[9]Pathology Department, Chaim Sheba Medical Center, Ramat Gan, Israel

ORCID iD
Arie Admon http://orcid.org/0000-0003-0504-3950

REFERENCES
1 Hirsch FR, Scagliotti GV, Mulshine JL, *et al*. Lung cancer: current therapies and new targeted treatments. *Lancet* 2017;389:299–311.
2 Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature* 2018;553:446–54.

3 Gjerstorff MF, Andersen MH, Ditzel HJ. Oncogenic cancer/testis antigens: prime candidates for immunotherapy. *Oncotarget* 2015;6:15772–87.

4 Hu Z, Ott PA, Wu CJ. Towards personalized, tumour-specific, therapeutic vaccines for cancer. *Nat Rev Immunol* 2018;18:168–82.

5 Fritsch EF, Burkhardt UE, Hacohen N, *et al*. Personal neoantigen cancer vaccines: a road not fully paved. *Cancer Immunol Res* 2020;8:1465–9.

6 Rammensee H-G, Singh-Jasuja H. Hla ligandome tumor antigen discovery for personalized vaccine approach. *Expert Rev Vaccines* 2013;12:1211–7.

7 Rock KL, Reits E, Neefjes J. Present yourself! by MHC class I and MHC class II molecules. *Trends Immunol* 2016;37:724–37.

8 Fritsche J, Rakitsch B, Hoffgaard F, *et al*. Translating Immunopeptidomics to Immunotherapy-Decision-Making for patient and personalized target selection. *Proteomics* 2018;18:e1700284.

9 Chong C, Coukos G, Bassani-Sternberg M. Identification of tumor antigens with immunopeptidomics. *Nat Biotechnol* 2022;40:175–88.

10 Tabayoyong WB, Zavazava N. Soluble HLA revisited. *Leuk Res* 2007;31:121–5.

11 Bassani-Sternberg M, Barnea E, Beer I, *et al*. Soluble plasma HLA peptidome as a potential source for cancer biomarkers. *Proc Natl Acad Sci U S A* 2010;107:18769–76.

12 Ritz D, Gloger A, Weide B, *et al*. High-Sensitivity HLA class I peptidome analysis enables a precise definition of peptide motifs and the identification of peptides from cell lines and patients' sera. *Proteomics* 2016;16:1570–80.

13 Shraibman B, Barnea E, Kadosh DM, *et al*. Identification of tumor antigens among the HLA peptidomes of glioblastoma tumors and plasma. *Mol Cell Proteomics* 2019;18:1255–68.

14 Feller-Kopman D, Light R. Pleural disease. *N Engl J Med* 2018;378:740–51.

15 Walker S, Mercer R, Maskell N, *et al*. Malignant pleural effusion management: keeping the flood gates shut. *Lancet Respir Med* 2020;8:609–18.

16 Psallidas I, Kanellakis NI, Gerry S, *et al*. Development and validation of response markers to predict survival and pleurodesis success in patients with malignant pleural effusion (promise): a multicohort analysis. *Lancet Oncol* 2018;19:930–9.

17 Murthy P, Ekeke CN, Russell KL, *et al*. Making cold malignant pleural effusions hot: driving novel immunotherapies. *Oncoimmunology* 2019;8:e1554969.

18 Amirghofran Z, Sheikhi AK, Kumar PV, *et al*. Soluble HLA class I molecules in malignant pleural and peritoneal effusions and its possible role on NK and LAK cytotoxicity. *J Cancer Res Clin Oncol* 2002;128:443–8.

19 Barnstable CJ, Bodmer WF, Brown G, *et al*. Production of monoclonal antibodies to group A erythrocytes, HLA and other human cell surface antigens-new tools for genetic analysis. *Cell* 1978;14:9–20.

20 Ishihama Y, Rappsilber J, Mann M. Modular stop and go extraction tips with stacked disks for parallel and multidimensional peptide fractionation in proteomics. *J Proteome Res* 2006;5:988–94.

21 Rijensky NM, Blondheim Shraga NR, Barnea E, *et al*. Identification of tumor antigens in the HLA peptidome of patient-derived xenograft tumors in mouse. *Mol Cell Proteomics* 2020;19:1360–74.

22 Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008;26:1367–72.

23 Jurtz V, Paul S, Andreatta M, *et al*. NetMHCpan-4.0: improved peptide-MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. *J Immunol* 2017;199:3360–8.

24 Andreatta M, Alvarez B, Nielsen M. GibbsCluster: unsupervised clustering and alignment of peptide sequences. *Nucleic Acids Res* 2017;45:W458–63.

25 Wölfl M, Greenberg PD. Antigen-Specific activation and cytokine-facilitated expansion of naive, human CD8+ T cells. *Nat Protoc* 2014;9:950–66.

26 Stenhouse G, Fyfe N, King G, *et al*. Thyroid transcription factor 1 in pulmonary adenocarcinoma. *J Clin Pathol* 2004;57:383–7.

27 Zavazava N, Krönke M. Soluble HLA class I molecules induce apoptosis in alloreactive cytotoxic T lymphocytes. *Nat Med* 1996;2:1005–10.

28 Fang L, Zhu J, Ma Y, *et al*. Neuroepithelial transforming gene 1 functions as a potential prognostic marker for patients with non-small cell lung cancer. *Mol Med Rep* 2015;12:7439–46.

29 Qi Y, Kong F-M, Deng Q, *et al*. Clinical significance and prognostic value of Vav1 expression in non-small cell lung cancer. *Am J Cancer Res* 2015;5:2491–7.

30 Sung H-J, Ahn J-M, Yoon Y-H, *et al*. Quiescin sulfhydryl oxidase 1 (QSOX1) secreted by lung cancer cells promotes cancer metastasis. *Int J Mol Sci* 2018;19:3213.

31 Burgess JT, Bolderson E, Adams MN, *et al*. SASH1 is a prognostic indicator and potential therapeutic target in non-small cell lung cancer. *Sci Rep* 2020;10:18605.

32 Light RW. Clinical practice. pleural effusion. *N Engl J Med* 2002;346:1971–7.

33 Rodríguez-Panadero F, López Mejías J. Low glucose and pH levels in malignant pleural effusions. diagnostic significance and prognostic value in respect to pleurodesis. *Am Rev Respir Dis* 1989;139:663–7.

34 Uhlen M, Oksvold P, Fagerberg L, *et al*. Towards a knowledge-based human protein atlas. *Nat Biotechnol* 2010;28:1248–50.

35 Wang C, Gu Y, Zhang K, *et al*. Systematic identification of genes with a cancer-testis expression pattern in 19 cancer types. *Nat Commun* 2016;7:10499.

36 da Silva VL, Fonseca AF, Fonseca M, *et al*. Genome-Wide identification of cancer/testis genes and their association with prognosis in a pan-cancer analysis. *Oncotarget* 2017;8:92966–77.

37 Bruggeman JW, Koster J, Lodder P, *et al*. Massive expression of germ cell-specific genes is a hallmark of cancer and a potential target for novel treatment development. *Oncogene* 2018;37:5694–700.

38 Djureinovic D, Hallström BM, Horie M, *et al*. Profiling cancer testis antigens in non–small-cell lung cancer. *JCI Insight* 2016;1:1–18.

39 Qi YA, Maity TK, Cultraro CM, *et al*. Proteogenomic analysis unveils the HLA class I-Presented immunopeptidome in melanoma and EGFR-mutant lung adenocarcinoma. *Mol Cell Proteomics* 2021;20:100136.

40 Chen J, Bardes EE, Aronow BJ, *et al*. ToppGene suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* 2009;37:W305–11.

41 Thickett DR, Armstrong L, Millar AB. Vascular endothelial growth factor (VEGF) in inflammatory and malignant pleural effusions. *Thorax* 1999;54:707–10.

42 Abelin JG, Keskin DB, Sarkizova S, *et al*. Mass spectrometry profiling of HLA-associated peptidomes in mono-allelic cells enables more accurate epitope prediction. *Immunity* 2017;46:315–26.

43 Johnson DR. Differential expression of human major histocompatibility class I loci: HLA-A, -B, and -C. *Hum Immunol* 2000;61:389–96.

44 Creech AL, Ting YS, Goulding SP, *et al*. The role of mass spectrometry and Proteogenomics in the advancement of HLA epitope prediction. *Proteomics* 2018;18:e1700259.

45 Freudenmann LK, Marcu A, Stevanović S. Mapping the tumour human leukocyte antigen (HLA) ligandome by mass spectrometry. *Immunology* 2018;154:331–45.

46 Campoli M, Ferrone S. Tumor escape mechanisms: potential role of soluble HLA antigens and NK cells activating ligands. *Tissue Antigens* 2008;72:321–34.

47 Du X, Shao Y, Qin H-F, *et al*. ALK-rearrangement in non-small-cell lung cancer (NSCLC). *Thorac Cancer* 2018;9:423–30.

48 Schumacher TN, Scheper W, Kvistborg P. Cancer neoantigens. *Annu Rev Immunol* 2019;37:173–200.

49 Gros A, Parkhurst MR, Tran E, *et al*. Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients. *Nat Med* 2016;22:433–8.

50 Anagnostou V, Smith KN, Forde PM, *et al*. Evolution of neoantigen landscape during immune checkpoint blockade in non-small cell lung cancer. *Cancer Discov* 2017;7:264–76.

51 Xu Y, Zou R, Wang J, *et al*. The role of the cancer testis antigen PRAME in tumorigenesis and immunotherapy in human cancer. *Cell Prolif* 2020;53:e12770.

52 Gordeeva O. Cancer-Testis antigens: unique cancer stem cell biomarkers and targets for cancer therapy. *Semin Cancer Biol* 2018;53:75–89.

53 Yang P, Meng M, Zhou Q. Oncogenic cancer/testis antigens are a hallmarker of cancer and a sensible target for cancer immunotherapy. *Biochim Biophys Acta Rev Cancer* 2021;1876:188558.

54 Sebban S, Farago M, Rabinovich S, *et al*. Vav1 promotes lung cancer growth by instigating tumor-microenvironment cross-talk via growth factor secretion. *Oncotarget* 2014;5:9214–26.

55 Cooke M, Baker MJ, Kazanietz MG. Rac-GEF/Rac signaling and metastatic dissemination in lung cancer. *Front Cell Dev Biol* 2020;8:118.

56 Dong Y, Wang Z, Shi Q. Liquid biopsy based single-cell transcriptome profiling characterizes heterogeneity of disseminated tumor cells from lung adenocarcinoma. *Proteomics* 2020;20:e1900224.

57 Guo Z, Xie Z, Shi H, *et al*. Malignant pleural effusion supernatant is an alternative liquid biopsy specimen for comprehensive mutational profiling. *Thorac Cancer* 2019;10:823–31.

58 White R, Pulford E, Elliot DJ, *et al*. Quantitative mass spectrometry to identify protein markers for diagnosis of malignant pleural mesothelioma. *J Proteomics* 2019;192:374–82.