Mapping the soluble human leukocyte antigen peptidome of pleural effusions reveals lung and tumor-associated antigens

Sofia Khazan-Kost, Gal Cafri, Dganit Melamed Kadosh, Navit Mooshayef, Sumit Chatterji, Dan Dominissini, , Sigal Manor, Bracha Zisser, Limor Broday, Efrosiniia Talalai, Anat Shemer, Oranit Zadok, Efrat Ofek, Amir Onn, Arie Admon, and Michael Peled

Supplemental Material

Supplementary Tables

Supplementary Table S7. CTA-deriving peptides identified in the pleural HLA peptidomes

Gene name	Score*	Peptide	Patients	Typing	Min rank**
HDGFL1	3	AEAERAAEA	P1	HLA-B50:01	0.0176
			P2	HLA-B41:01	0.1064
TDRD5	2	VLRTEGHAI	P1	HLA-C12:03	2.4599
TDRD1	4	MEQYCSIKI	P2	HLA-B41:01	0.6405
		YGNFEILSLMR	P3	HLA-A33:01	5.9686
TOPAZ1	1	QTVECLQSL	P2	HLA-A26:01	1.0153
SAGE1	4	APDNVLLTL	P3	HLA-B53:01	0.1094
			P9	HLA-B35:02	0.0338
			P11	HLA-B38:01	0.4833
			P13	HLA-B35:02	0.0338
		STRDLHSTV	P11	HLA-C12:03	0.3482
SPATA16	2	KIRADKIEK	P3	HLA-A03:01	0.1145
RBMXL2	3	KVAQATKP	P3	HLA-A03:01	38
			P4	HLA-A11:01	44.4
DNAH17	1	SLADLEAF	P3	HLA-C04:01	4.607
		VLYLKPLRI	P4	HLA-C04:01	0.6036
		EWKDGLFSTI	P9	HLA-A23:01	1.5053
		SPSTSIFFI	P13	HLA-B35:02	0.3594
C1orf94	2	SSSALVAKGP	P3	HLA-A03:01	70.625
			P4	HLA-A11:01	61.1111
			P11	HLA-A26:01	76.4286
RIMBP3	2	TAGSTLLEF	P3	HLA-B53:01	0.3784
ADIG	2	DILQVRRYRYD	P4	HLA-B44:02	76.6667
			P6	HLA-A33:01	6.6001
TSKS	1	LLDRALTSL	P4	HLA-C05:01	0.0312
			P14	HLA-C08:02	0.0305
DNAH8	2	VEIWLLDLL	P4	HLA-B40:01	0.1702
		YTFESAKKVC	P14	HLA-C08:02	19.5437
FAM186B	1	AELSLVPAP	P5	HLA-B39:01	21.3796
MS4A6E	1	LSLMLVSTV	P5	HLA-A02:01	8.1356
CNBD1	1	LKTIPDLTF	P8	HLA-C07:01	2.7206
			P12	HLA-C07:01	2.7206
			P13	HLA-C06:02	2.9534
GJA8	1	LSVASVSLFL	P8	HLA-B58:01	4.4833
		LSRGYQETL	P11	HLA-C12:03	0.5468
C1orf167	1	SEAICWQLL	P9	HLA-B49:01	0.8435
DPPA2	5	VEVITSAPGA	P10	HLA-B18:01	10.3617

FAM71D	1	APFVSPPML	P11	HLA-C12:03	1.4251
SMIM23	1	DEKQTLLAL	P11	HLA-B49:01	0.8556
PAPPA	1	EVIASYNQL	P11	HLA-A26:01	0.0533
		SHDLGLHVL	P11	HLA-B38:01	0.0045
ACRBP	2	EYERFFALL	P11	HLA-A23:01	0.4514
			P13	HLA-C04:01	0.2937
IQCA1L	1	VYREEEESL	P11	HLA-A23:01	0.5211
KCNU1	1	LDKDKVYGVA	P13	HLA-A01:01	34.3571
RBM44	1	LDLTGEMKNVE	P13	HLA-B57:01	83.4375
XAGE1A	2	MPEAGEEQPQV	P13	HLA-B35:02	2.4477
ACSBG2	2	PIGELYGL	P13	HLA-B35:02	2.0659
ACRV1	2	PSGEHLSGE	P13	HLA-B57:01	63.125
DMRT1	4	SSFTVTPVI	P13	HLA-B57:01	0.6472
PRAME	5	VLDGLDVLL	P13	HLA-C04:01	0.1152
CCDC70	1	ALLEGEKAL	P14	HLA-C08:02	0.7197
TCP11	2	PSLLNHTTK	P14	HLA-A11:01	2.9328

* Score refers to the number of sources [CTDatabase (http://www.cta.lncc.br/index.php)**;** Wang et al., 2016**;** da Silva et al. 2017**;** Djureinovic et al. 2016**;** Bruggeman et al. 2018**;** Qi et al. 2021] the gene was identified as cancer/testis antigen (CTA).

** Peptides were scored NetMHCpan 4.1 (http://www.cbs.dtu.dk/services/NetMHCpan/) according to their match to the patient's HLA alleles motif. Rank <2 represented increased he probability that the peptide was HLA ligand.

Supplementary Table S10. Number and percent of identified soluble and membranal peptides per HLA allele.

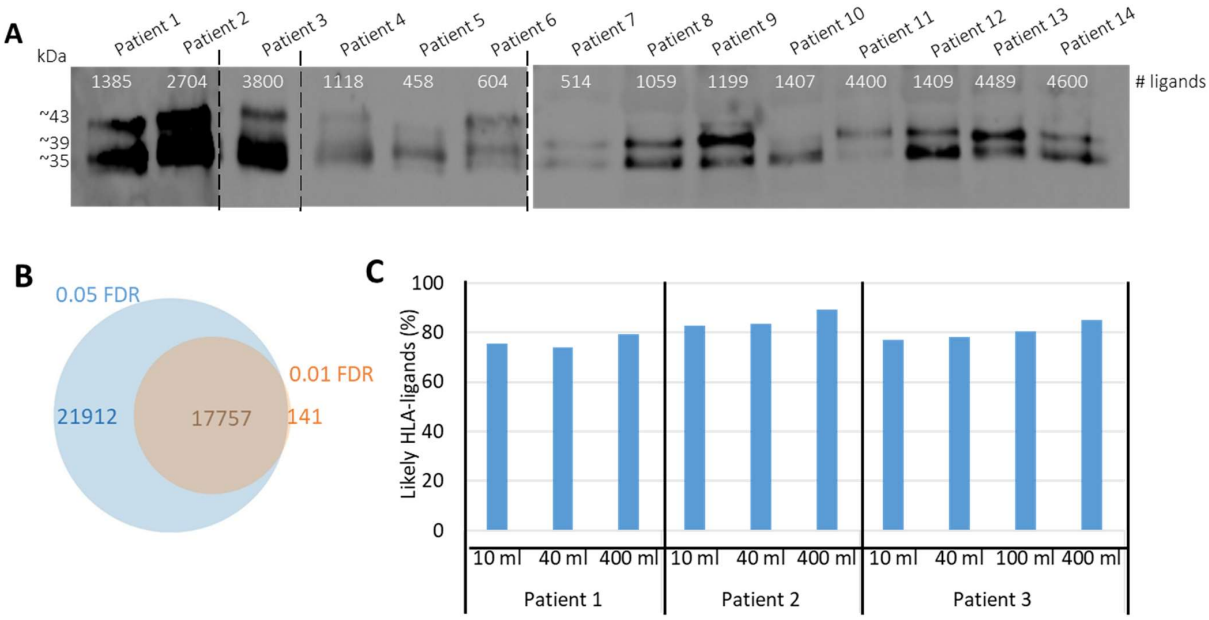
Patient	Allele	mHLA*		sHLA*	
		No.	%**	No.	%**
P1	HLA-A26:01	86	27	203	15
P1	HLA-B38:01	81	26	371	27
P1	HLA-A68:01	68	22	360	26
P1	HLA-B50:01	39	12	206	15
P1	HLA-C12:03	33	10	221	16
P1	HLA-C06:02	8	3	24	2
P2	HLA-A26:01	503	19	366	14
P2	HLA-B38:01	764	28	825	31
P2	HLA-B41:01	595	22	469	17
P2	HLA-A01:01	468	17	682	25
P2	HLA-C12:03	297	11	315	12
P2	HLA-C17:01	63	2	47	2
P3	HLA-B18:01	1421	35	974	26
P3	HLA-A03:01	1412	35	1161	31
P3	HLA-A33:01	520	13	831	22
P3	HLA-B53:01	272	7	534	14
P3	HLA-C05:01	265	7	146	4
P3	HLA-C04:01	164	4	154	4
P4	HLA-B40:01	1379	40	239	21
P4	HLA-A11:01	651	19	394	35
P4	HLA-B44:02	518	15	200	18
P4	HLA-A02:01	376	11	87	8
P4	HLA-C03:04	315	9	166	15
P4	HLA-C05:01	201	6	32	3
P5	HLA-A02:01	490	49	151	33
P5	HLA-B39:01	330	33	235	51
P5	HLA-B08:01	129	13	68	15
P5	HLA-C07:02	41	4	4	1
P6	HLA-A02:05	33	27	84	14
P6	HLA-B18:01	27	22	122	20
P6	HLA-C12:03	25	20	149	25
P6	HLA-A33:01	16	13	134	22
P6	HLA-B50:01	14	11	91	15
P6	HLA-C06:02	7	6	24	4
P8	HLA-A01:01	1037	67	674	64
P8	HLA-B58:01	470	30	366	35
P8	HLA-C07:01	45	3	19	2

P9	HLA-B49:01	1066	35	296	25
P9	HLA-B35:02	805	27	360	30
P9	HLA-A23:01	470	16	329	27
P9	HLA-A31:01	344	11	106	9
P9	HLA-C04:01	283	9	103	9
P9	HLA-C07:01	44	1	5	0
P10	HLA-B18:01	456	59	899	64
P10	HLA-A25:01	178	23	210	15
P10	HLA-A32:01	61	8	166	12
P10	HLA-C12:03	55	7	90	6
P10	HLA-C04:01	29	4	42	3
P11	HLA-B49:01	1834	34	1573	31
P11	HLA-B38:01	1287	24	1076	21
P11	HLA-A23:01	984	18	1570	31
P11	HLA-A26:01	859	16	389	8
P11	HLA-C12:03	422	8	449	9
P11	HLA-C07:01	67	1	47	1
P12	HLA-B49:01	385	35	442	31
P12	HLA-A01:01	363	33	534	38
P12	HLA-B15:17	217	20	245	17
P12	HLA-A32:01	124	11	174	12
P12	HLA-C07:01	16	1	14	1
P13	HLA-B35:02	1956	44	1706	38
P13	HLA-A24:02	1112	25	1309	29
P13	HLA-A01:01	525	12	592	13
P13	HLA-B57:01	394	9	504	11
P13	HLA-C04:01	366	8	234	5
P13	HLA-C06:02	139	3	175	4
P14	HLA-B35:01	1217	33	1920	42
P14	HLA-A11:01	1060	28	1026	22
P14	HLA-A03:01	590	16	569	12
P14	HLA-B14:02	386	10	326	7
P14	HLA-C08:02	263	7	285	6
P14	HLA-C04:01	210	6	474	10

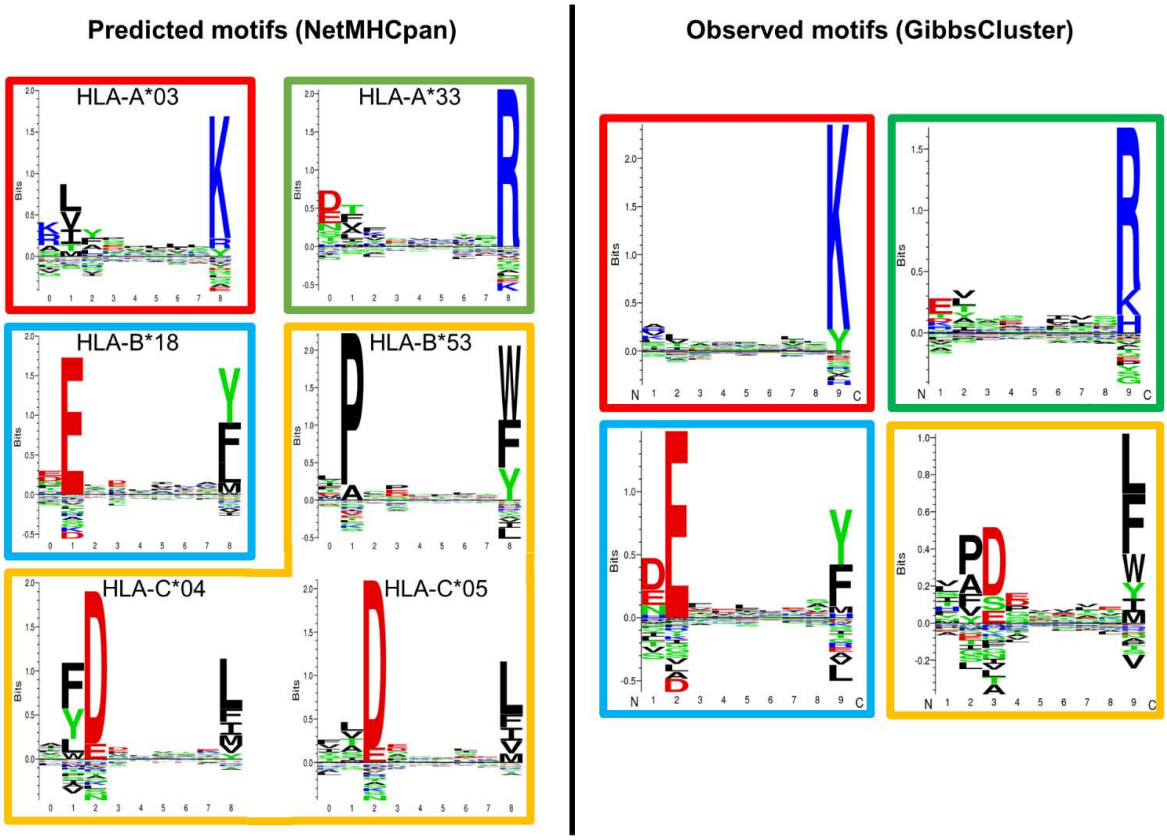
* Membranal HLA (mHLA)**;** soluble HLA (sHLA)

** Percent was calculated from the total number of identified peptides in each patient's soluble or membranal HLA peptidome.

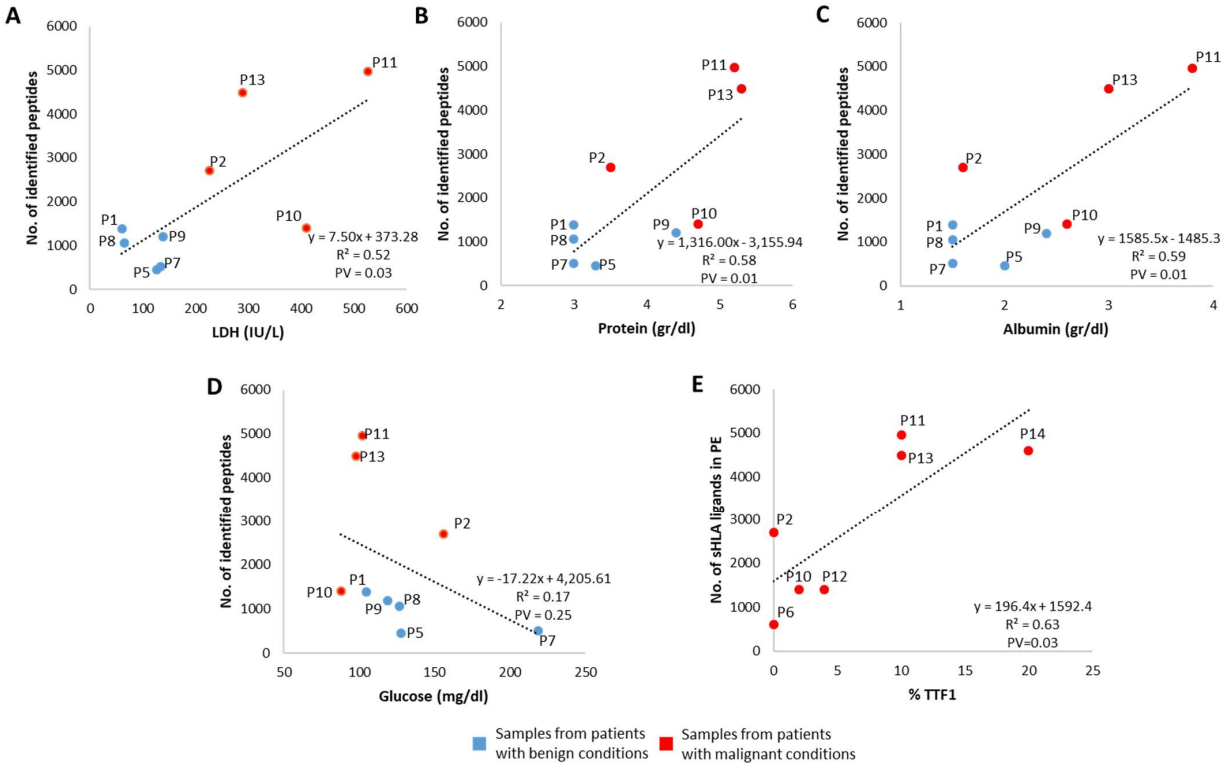
Supplementary Figures



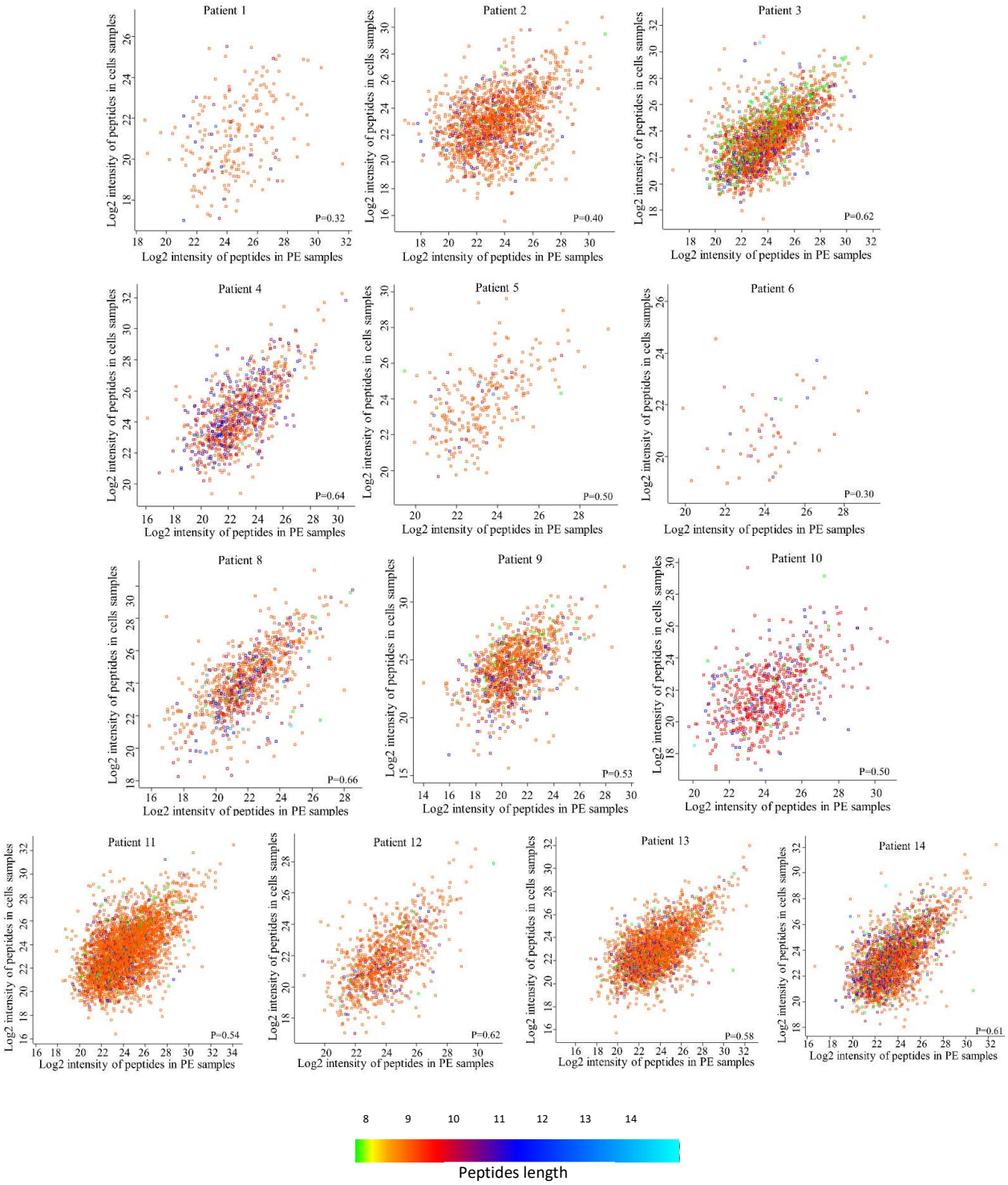
Supplementary Figure S1. Quality control and calibration of the HLA-peptide complexes immunoaffinity purification process. A) Western blot analyses of HLA molecules purified from 400 ml pleural effusions of Patients 1-14. The detection was with the anti-pan-HLA antibody (ab126237, Abcam, Cambridge, UK). Different western blots are separated by dotted lines. The number of likely-HLA ligands are indicated in white above the bands. **B)** Comparison of the total number of identified HLA ligands identified by a using 0.01 or 0.05 FDR. Database search was performed with MaxQuant (version 1.6.10.43). The list of identified peptides was filtered to retain only peptides with the length of 8-14 aa that fit the consensus sequence motifs of the HLA allotypes of each patient, and to remove known contaminants. **C)** Percentage of likely HLA ligands from different volumes of pleural effusions from Patients 1-3.



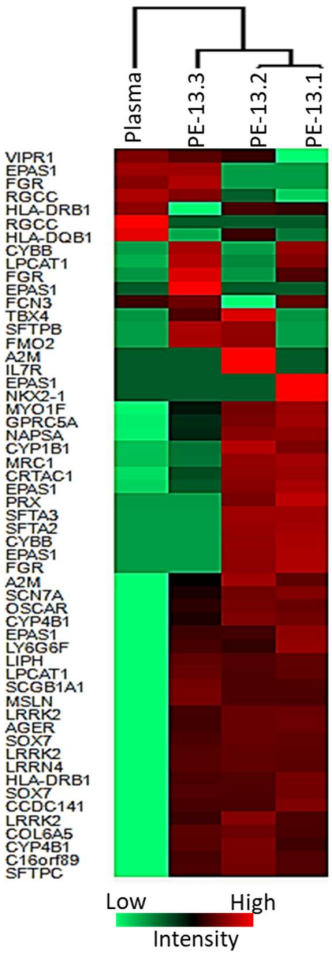
Supplementary Figure S2. Predicted vs. observed HLA peptides sequence motifs of Patient 3. The known HLA motifs were retrieved from NetMHCpan 4.0 (left). The observed motifs were obtained by GibbsCluster 2.0 analysis (right). In cases when no distinct motif was observed, the motifs were grouped (yellow boxes).



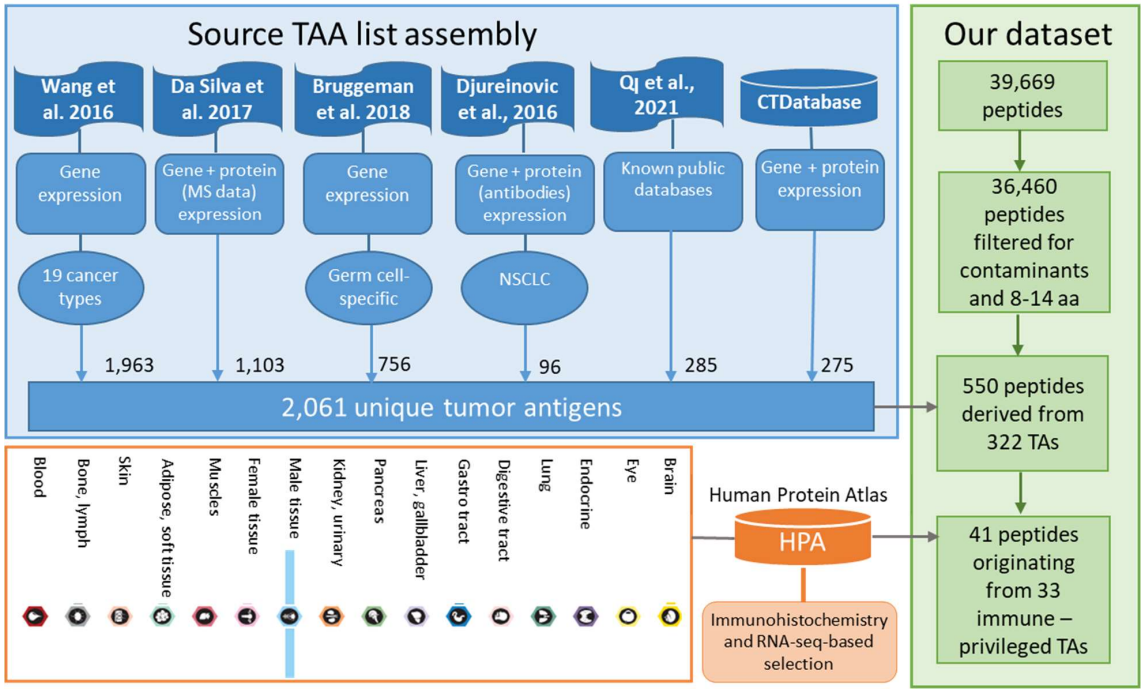
Supplementary Figure S3. Linear regression analyses between the numbers of identified peptides and the levels of LDH (A), protein (B), albumin (C), glucose (D), and TTF1 (E). In A-D the presented data is of Patients 1, 2, 5, 7, 8, 9, 10, 11, 13. In E, the percentage of TTF1 was calculated from paraffin-embedded cell blocks from Patients 2, 6, 10, 11, 12, 13, 14. R^2 and p-value are indicated in each graph.



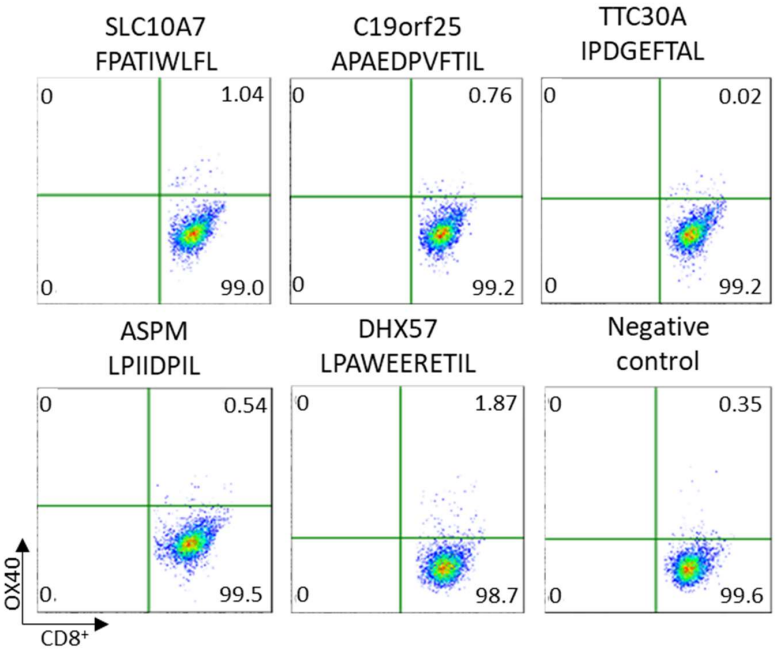
Supplementary Figure S4. Similarities between membranal and soluble peHLA peptidomes of each patient. Scatter plots on $\log_2$ scales of the LC-MS intensities of membranal and soluble peHLA peptidomes. Pearson correlation is indicated in each graph.



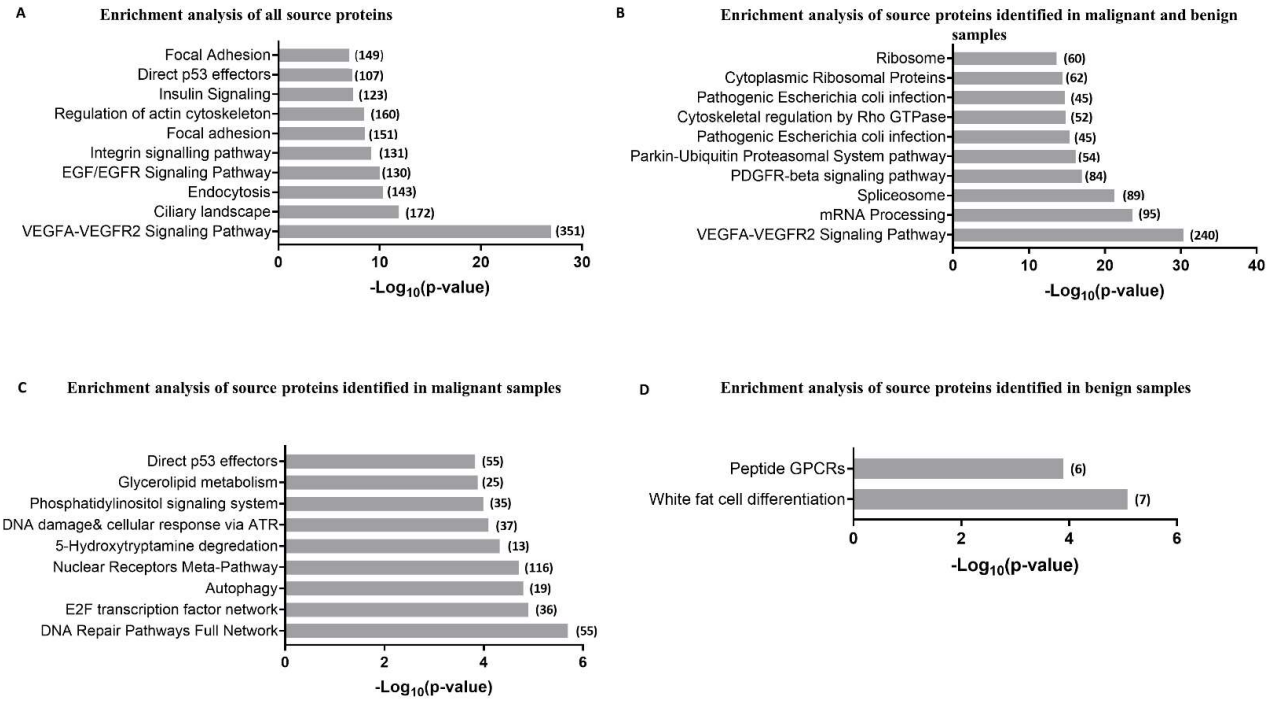
Supplementary Figure S5. HLA ligands derived from lung-associated proteins are more abundant in the peHLA peptidome than in the plasma sHLA peptidome. Hierarchical cluster analysis of lung-associated genes, as defined by the Human Protein Atlas (HPA, retrieved April 1, 2021), that were also the source for peptides in the soluble peHLA peptidome and the plasma sHLA peptidome. Three samples of pleural effusions from P13, taken from different visits to the clinic, are compared with a plasma sample from the same patient. The presented $\log_2$ intensities were normalized by Z-score.



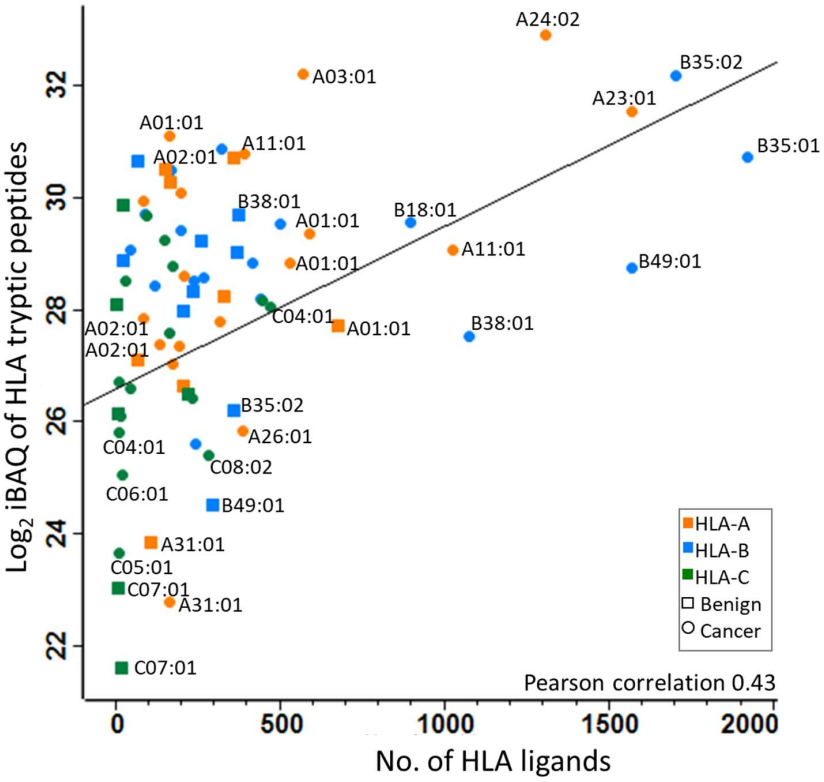
Supplementary Figure S6. Summary of the process for defining tumor antigens. Tumor antigens, including CTAs, were defined by using six publicly available databases: Supplementary Data 3C published data by Wang et al. (Wang et al., 2016); Supplementary Table 2 published by da Silva et al. (da Silva et al. 2017); Supplementary Table 1D published by Bruggeman et al. (Bruggeman et al. 2018); Supplementary Table 5 published by Djureinovic et al. (Djureinovic et al. 2016); and Supplementary Table S4 published by Qi et al. (Qi et al., 2021); CTDatabase (http://www.cta.lncc.br/index.php). Each gene was scored 1-6, according to the number of databases in which it was classified as a CTA. The source proteins of the identified peptides were next checked for restricted expression in tumors and testis according to the Human Protein Atlas (Uhlen et al. 2010).



Supplementary Figure S7. FACS analysis of OX40 expression. The assay was performed on purified naïve CD8+ T cells, from healthy HLA-B*35:02 positive donor, by co-cultured 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+/OX40+ cells.



Supplementary Figure S8. ToppGene pathway enrichment analysis of source proteins of the identified HLA peptides. Functional pathway enrichment analysis of the source proteins of the HLA peptides analyzed by ToppGene (https://toppgene.cchmc.org/) from **(A)** all the samples, **(B)** source proteins that were common to benign and malignant effusions, **(C)** source proteins that were unique to malignant effusions, and **(D)** source proteins that were unique to benign effusions. The top ten significantly enriched functional clusters (y-axis) are displayed over the corresponding −log10 p-value (x-axis). The p-value represents the probability of discovering at least x number of genes (x = number in brackets) out of the total number of the genes in the list (Supplementary Table S9) annotated to a particular functional cluster given the proportion of genes in the whole genome that are annotated to that functional cluster.



Supplementary Figure S9. The iBAQ intensity of HLA molecules purified from pleural fluids correlate with the number of putative HLA ligands presented by each allotype. HLA heavy chains from 80% acetonitrile fractions of the HLA immunoaffinity purification from pleural fluids of Patients 1-14 were digested with trypsin and the peptides were analyzed by LC-MS/MS. Colors represent the HLA allele, i.e. HLA-A (orange), HLA-B (blue), HLA-C (green). Circles stand for cancer and square for benign patients.