Supporting Information

Schuster et al. 10.1073/pnas.1707658114

SI Methods

Patient Samples. All tissue samples were collected at the University Hospital of Tübingen after obtaining patient informed consent in accordance with the principles of the Declaration of Helsinki. The local institutional review board (Ethics Committee at the Medical Faculty and at the University Hospital of Tübingen) approved all study protocols. If not stated otherwise, samples were stored at −80 °C until further use. No samples were excluded unless histologic diagnosis did not match to the investigated cohort (e.g., benign ovarian cystadenoma) or MS analysis yielded less than 100 MHC class I or class II peptides. EOC patient characteristics are presented in Dataset S1. Two-digit HLA typing was performed by sequence specific primer PCR using the HLA-Ready Gene System (Innotrain) and evaluated by SCORE Software (Olerup) at the Department of Transfusion Medicine of the University Hospital of Tübingen. High-resolution four-digit HLA typing was performed by next generation sequencing on a GS Junior Sequencer using the GS GType HLA Primer Sets (both Roche). High-resolution HLA class I typing was checked or complemented for samples with available RNA-Seq data using Optitype (52).

Tissue Dissociation. EOC, benign ovary, and fallopian tube tissues were freshly collected from patients undergoing tumor resection/debulking or salpingoophorectomy. Tissues were minced into small pieces <2 mm^3 and transferred into an enzymatic dissociation solution containing 400 U/mL collagenase type IV, 5 U/mL dispase (both Life Technologies) and 0.1 mg/mL DNase (Roche) in DMEM (Life Technologies) with 10% FCS (Lonza). Dissociation was performed on a rotating shaker (Infors HT) for 3 h at 37 °C. Remaining tissue fragments (typically <1% of initial weight) were removed using a 100-μm cell strainer (BD). Single-cell suspensions were washed twice with PBS and erythrocytes were lysed using ammonium chloride lysis buffer.

HLA Surface Molecule Quantification. HLA surface expression was determined using QIFIKIT quantification flow cytometric assay (Dako) according to the manufacturer's instructions. Cells were stained with either pan-HLA class I-specific monoclonal antibody W6/32, HLA-DR–specific L243 or respective isotype control. Discrimination of cell types was based on surface marker staining with fluorescently labeled antibodies directed against CD45 (AmCyan clone 2D1; BD), CD31 (PeCy7, clone WM59; BioLegend), EpCAM (APC, clone HEA125; Miltenyi), and CD34 (APCCy7, clone 581; BioLegend). Next, 7-AAD (BioLegend) was added as viability marker immediately before analysis on a LSR SORP Fortessa instrument (BD). Triplicates were recorded for each sample with median fluorescence intensities used for calculation of surface molecule expression.

Cell Separation. Cell separation was performed using two consecutive MACS protocols, according to the manufacturer's instructions (Miltenyi). Separations were performed using XS columns and a superMACS separator (both Miltenyi). The first separation aimed at positive selection of $CD45^+$ leukocytes. The negative fraction was subsequently enriched for $EpCAM^+$ tumor cells. The remaining $CD45^-$ $EpCAM^-$ fraction was assumed to represent the stroma cell fraction.

HLA Ligand Isolation. HLA class I and II molecules were isolated by standard immunoaffinity purification as described previously (22). Pan-HLA class I-specific mAb W6/32 was employed for HLA class I isolation and pan-HLA class II mAb Tü39, as well as HLA-DR–specific mAb L243, were used for HLA class II isolation. Between 0.5 and 3 g of tissue samples or 2.5×10^8 to 1×10^9 of enriched cell fractions were employed for lysis and HLA ligand isolation.

Immunopeptidome Analysis by LC-MS/MS. Immunopeptidome analysis was performed on an LTQ Orbitrap XL mass spectrometer (Thermo Fisher) equipped with a nanoelectron spray ion source and coupled to an Ultimate 3000 RSLC Nano UHPLC System (Dionex). Peptide samples were loaded with 3% of solvent B (20% H_2O, 80% acetonitrile, and 0.04% formic acid) on a 2-cm PepMap 100 C18 Nanotrap column (Dionex) at a flowrate of 4 μL/min for 10 min. Separation was performed on a 50-cm PepMap C18 column with a particle size of 2 μm (Dionex) mounted in a column oven running at 50 °C. The applied gradient ranged from 3 to 30% solvent B within 140 min at a flow rate of 175 nL/min. (Solvent A: 99% H_2O, 1% ACN and 0.1% formic acid; Solvent B: 20% H_2O, 80% ACN and 0.04% formic acid.) MS analysis was performed in data-dependent acquisition mode employing a top five method (i.e., during each survey scan the five most abundant precursor ions were selected for fragmentation). Survey scans were recorded in the Orbitrap at a resolution of 60,000. MS/MS analysis was performed by collision induced dissociation (CID, normalized collision energy 35%, activation time 30 ms, isolation width 1.3 *m/z*) with subsequent analysis in the LTQ. Mass range for HLA class I ligands was limited to 400–650 *m/z* with possible charge states 2+ and 3+ selected for fragmentation. For HLA class II mass range was set to 300–1,500 *m/z*, allowing for fragmentation with positive charge states ≥2.

HLA class I samples were analyzed in two to five technical replicates, while for HLA class II samples one to three technical replicates were typically acquired. Initial runs were performed without dynamic exclusion, whereas for consecutive runs a dynamic exclusion of 5 s or more was enabled.

MS Data Processing and Analysis. MS data analysis was carried out using Proteome discoverer 1.3 (Thermo Fisher). Peak lists were searched against the human proteome as comprised in the UniProtKB/Swiss-Prot database (www.uniprot.org, released September 27th 2013; including 20,279 reviewed protein sequences) using Mascot search engine (Mascot 2.2.04; Matrix Science). Mass tolerance for processing was 5 ppm for precursor ions and 0.5 Da for fragment ions. No cleavage specificity was selected and the only dynamic modification allowed was oxidized methionine. Peptide confidence was determined using a percolator algorithm with a target value of $q \leq 0.05$ (5% FDR). Additional postprocessing filters were search engine rank = 1 and peptide length of 8–12 amino acids for HLA class I ligands and 12–25 amino acids for HLA class II ligands. Protein grouping was disabled to ensure multiple annotations of peptides, if sequences map into multiple proteins due to conservation. HLA annotation was performed using HLA prediction algorithms hosted at SYFPEITHI (www.syfpeithi.de) and NETMHC 3.4 (www.cbs.dtu.dk/services/NetMHC/). In case of ambiguous results, multiple alleles are mentioned.

For comparative profiling, "one-hit wonders" (i.e., peptides only presented on one source) with a PSM count ≤ 5 were removed from the datasets.

Label-free quantitation of peptides on tumor vs. $CD45^+$ and tumor vs. stroma cells was performed using Sieve 2.1 (Thermo Fisher). At least three replicates of MS raw files for each cell-enriched fraction as well as results from whole-tissue MHC precipitations were aligned together with a maximum retention

time shift of 2.5 min. Frames were generated based on MS^2 scan events with a maximum retention time width of 3.5 min and 5 ppm mass tolerance. Identifications were imported from Proteome discoverer using Mascot search results (see above). Total ion current chromatogram normalization was used to accommodate for differences in sample intensities.

RNA-Seq Analysis. RNA-Seq analysis was performed by an external service provider (CeGaT GmbH). In brief, RNA was isolated from cryopreserved ovarian cancer as well as benign fallopian tube tissue using the RNeasy Mini kit (Qiagen). One-hundred nanograms of isolated RNA was employed for library preparation using the TruSeq Stranded mRNA Kit (Illumina). Single-end sequencing was performed on a HiSEq. 2500 instrument with a target read length of 100 bp and a read depth of at least 50 million reads. Mapping of sequenced reads to the reference genome (hg19) was completed using STAR (v2.4.0). Additional data processing and counting of mapped reads were computed with Cufflinks Tool Suite (v2.1.1) Resulting FPKM values were calculated using Cuffdiff employing a pooled-variance model and geometric normalization with enabled multiread-correction.

Immunogenicity Analysis of HLA Class I Ligands. Priming of peptide-specific cytotoxic lymphocytes was conducted using an established protocol involving aAPCs (38), consisting of streptavidin-coated polystyrene beads (5.6-μm diameter; Bangs Laboratories). Beads were resuspended at 2×10^6 particles per milliliter and incubated with 10 nM biotinylated peptide-MHC complexes and 10 nM stimulating anti-CD28 antibody (clone 9.3 derived from ATCC) each for 30 min at ambient temperature. T cells were isolated from whole blood of healthy donors using a CD8 magnetic cell isolation kit (Miltenyi). One million T cells per well were cultured in 96-well plates (Corning) and stimulated with 200,000 loaded aAPCs in the presence of 5 ng/mL IL-12 (PromoCell). T cells were stimulated three times in total with a weekly stimulation interval. Then, 40 U/mL IL-2 and 5 ng/mL IL-7 was added 2 d after each stimulation. The induction and expansion of peptide-specific T cells was assessed by MHC-multimer staining 1 wk after the last stimulation round. For each peptide and donor, a minimum of 30 wells were screened. Cultures were considered positive if a clearly distinct tetramer positive population of at least 0.5% of $CD8^+$ cells could be observed in one of the cultured wells.

Construction of TMA. Consecutive paraffin-embedded tumor samples of patients with high-grade serous carcinoma of the ovary or fallopian tube (EOC) with at least FIGO stage II/III and operated at the University Women's Hospital in Tübingen between 1999 and 2008 were retrieved from the archives of the Institute of Pathology after confirmation of histological subtype and grading according to published criteria (53). One-hundred fifty-four cases were initially included in the study. A TMA was constructed as described previously (54). We used six cores of 0.6-mm diameter of each patient (maximum three cores each from two different sites of the primary tumors, at least two separate cores). In addition, we constructed a TMA using paraffin-embedded tissue from the primary tumors of the prospectively collected cases for ligandome analysis. Three-micrometer-thick sections were cut, rehydrated, and subjected to specific pretreatment for IHC. In total, 23 cases were evaluable for immunoscoring and correlation with immunopeptidome data.

IHC. The following primary antibodies and dilutions were used for IHC: CD3 (1:100, rat monoclonal SP7; DCS), CD8 (1:200, mouse monoclonal C8/144B; DAKO), MUC16 (1:450, mouse monoclonal M11; DAKO), MSLN (1:100, mouse monoclonal SPM143; GeneTex), HCA2 (1:100, mouse monoclonal, in-house production), and L243 (1:50, mouse monoclonal, in-house production). The tissue sections, except for HCA2 and L243, were pretreated with EDTA-buffer solution (pH 8.6) at 95 °C for 36 min. IHC staining was performed on an automated immunostainer according to the manufacturer's instructions using the iView DAB detection kit (both Ventana). HCA2 and L243 sections were pretreated with citrate buffer solution (pH 6) at 100° for 5 min. IHC staining was performed on a semiautomated Tecan immunostainer using the ZytochemPlus HRP-Polymer kit according to the manufacturer's instructions (Tecan; Zytomed).

Immunoscoring. Quantification of TILs was carried out by first assessing the average number of immunostained cells per high-power field (HPF = 400×) by counting at least two HPF for each core. In a second step, the average number of lymphocytes per HPF for the left and right triple core set was calculated, as well as for all cores together. This bilateral average count was used for further calculations.

For expression of MUC16, MSLN, or MHC class I (HCA2) and class II (L243) staining intensity was graded from 0 to 3 and multiplied by a score from 1 to 4 for the percentage of reactive cells (1: 0–10%; 2: 10–50%; 3: 50–80%; 4: 80–100%). For all parameters, the cases were separated in quartiles and the best separation between two quartiles defined as cut-off value between high and low expression. Of the 154 cases on the TMA, 71 patients had undergone documented optimal tumor debulking (<1 cm residual tumor mass) and could be successfully evaluated for TILs and expression of proteins. Immunoscoring and clinical data analysis were performed by independent investigators.

Statistical Analysis/Visualization. If not mentioned otherwise, all figures and statistical analyses were generated using GraphPad Prism 6.0 (GraphPad software) or Microsoft Office 2010 (Microsoft). Word clouds were created using an online applet (www.wordle.net). Kaplan–Meier analysis was performed using SPSS statistical software (v21; IBM Corp.). *P* values less than 0.05 were considered statistically significant. A D'Agostino and Pearson test was used to verify normality and the *F*-test was used to verify equal variance. For Fig. 1 and Fig. S5, nonparametric Mann–Whitney-test was used because normal distribution could not be assessed in all cases due to small sample sizes in individual test groups. Spearman correlation was used to correlate IHC scores of MSLN and MUC16, as the datasets were not showing normal distribution. *P* values comparing two Kaplan–Meier survival curves in Fig. S5 were calculated using the log-rank (Mantel–Cox) test in GraphPad Prism. All *P* values were corrected for multiple testing using Benjamini–Hochberg correction if not mentioned otherwise.

Data Availability. BAM files of RNA-Seq data are available through National Center for Biotechnology Information Sequence Read Archive project PRJNA398141. MS raw data for epithelial ovarian cancer samples have been deposited to the ProteomeXchange Consortium via the PRIDE partner (52) repository with the dataset identifier PXD007635 or can be requested from the corresponding author.

Fig. S1. Gene-expression levels of HLA-DP and HLA-DQ in EOC and benign fallopian tubes. Gene-expression levels of HLA-DP and HLA-DQ genes in 15 EOC and 15 benign fallopian tube samples as analyzed by RNA-Seq. Data points represent values of individual samples expressed as FPKM. Mann–Whitney test was used to test for significance (**$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$). Horizontal lines indicate mean values ± SD.

Fig. S1

Fig. S2. Assay principle and gating strategy for flow cytometry based MHC quantification. (*A*) Viable single cells were gated for $CD45^+$ leukocytes. From the $CD45^-$ population, tumor and benign epithelial cells were selected based on their expression of EpCAM. Endothelial cells were determined by their expression of CD31 with most cells further coexpressing CD34. (*B*) Representative examples of flow cytometry analyses for OvCa 80 (dark gray) and OvN 25 (light gray) of the employed quantification assay showing histogram distributions of respective $EpCAM^+$-gated populations. QIFIKIT beads were stained with the same secondary antibody as used for staining of primary cells and acquired directly after the primary cells using the same instrument settings. Mean fluorescent intensity (MFI) values were calculated for each bead population (1–5), which are labeled with different defined numbers of antibodies. Based on these results a calibration curve was generated which was subsequently used to infer MHC surface expression from MFI values of respective primary cell populations.

Fig. S2

Fig. S3. Saturation analysis of HLA class I and II ligand source proteins. Saturation analysis of HLA class I (*A*) and II (*B*) ligand source protein identifications in ovarian cancer, as well as different healthy tissues. The mean number of unique source proteins has been calculated for each source count by 1,000 random samplings from the entirety of respective tissue samples [i.e., the number of source proteins of randomly picked samples was summed up for each source count (number of sources: 1, 2, 3, 4. . .) and this process was repeated 1,000 times before calculating the average]. Exponential regression was employed to extrapolate the maximal attainable coverage of source protein accessions for each tissue.

Fig. S3

Fig. S4. Presentation frequency of selected EOC exclusive HLA class I (*Upper*) and class II (*Lower*) ligand source proteins. Frequency of presentation and number of unique ligands (color coded) of selected EOC-exclusive HLA ligand source proteins among ovarian cancer samples in comparison with the housekeeping protein β-actin.

Fig. S4

Fig. S5. (*A*) Photomicrographs of IHC stainings of high-grade serous ovarian carcinomas for MUC16 (CA-125) and MSLN with different immunoreactivity scores. (*B*) Prognostic relevance of MUC16 and MSLN. IHC stainings were performed on TMAs with 71 high-grade serous EOC samples from patients with documented optimal tumor debulking. (*Upper*) Kaplan–Meier plot depicting the influence of MUC16 expression (*Left*, low expression score < 7, $n = 41$; high expression score ≥ 7, $n = 30$) and MSLN expression (*Right*, low expression < 6, $n = 15$; high expression ≥ 6, $n = 52$) on overall survival. (*Lower*) Impact of CD3 T cell infiltration into the intraepithelial compartment (*Left* CD3E, low infiltration < 7 cells/HPF, $n = 13$; high infiltration ≥ 7, $n = 57$) on overall survival of patients. Subgroup analysis of combined CD3 and MSLN staining (all scoring cutoffs as above) for intraepithelial CD3 T cells (low MSLN/high CD3E, $n = 11$; low MSLN/low CD3E, $n = 40$; high MSLN/low CD3E, $n = 14$; high MSLN/high CD3E, $n = 1$). (*C*) Analysis of HLA ligand presentation and source protein expression of MUC16 ($n = 23$) and MSLN ($n = 16$) were analyzed by IHC staining. Serum marker analysis of CA-125 ($n = 30$) was performed at the day of surgery. High MUC16 and MSLN IHC immunoreactivity scores, as well as CA-125 serum level are significantly associated with positive HLA ligand presentation. Mann–Whitney test was used to test for significance (*$P < 0.05$; **$P < 0.01$, ***$P < 0.001$) because of rejected normality test (D'Agostino and Pearson). Please note that for CA-125 log transformation of values was performed before testing.

Fig. S5

Fig. S6. GTEX gene-expression profile. Gene-expression profile of TOP EOC exclusively presented antigens according to the publicly available database GTEx. (Date of accession: February 12, 2016.)

Fig. S6

Fig. S7. Multimer analyses of in vitro primed CD8 T cells. $CD8^+$ T cells from healthy blood donors were primed in vitro using aAPCs. Graphs show single, viable cells stained for CD8 and PE-conjugated multimers of indicated specificity. (*A*) Representative MHC-peptide multimer stainings of one donor primed with HLA-A*02 restricted KLK10 derived peptide RALAKLLPL. MHC-peptide multimer staining was performed after three consecutive stimulations with aAPCs. A peptide from DDX5 (YLLPAIVHI) was employed as negative control. T cells from the same patient were routinely tested for the absence of preexisting memory T cells directly ex vivo, as well as after 12-d recall stimulation. (*B*) MHC-peptide multimer stainings of additional allotypes demonstrating expansion of specific T cells: EVTSSGRTSI (A*25; MUC16), TYSEKTTLF (A*24; MUC16), SPHPVTALL (B*07; MUC16), and NPKAFFSVL (B*07; IDO1).

Fig. S7

Other Supporting Information Files

Dataset S1 (XLSX)
Dataset S2 (XLSX)
Dataset S3 (XLSX)
Dataset S4 (XLSX)
Dataset S5 (XLSX)
Dataset S6 (XLSX)