

## Article

# Deciphering the immunopeptidome in vivo reveals new tumour antigens

<https://doi.org/10.1038/s41586-022-04839-2>

Received: 30 June 2021

Accepted: 6 May 2022

Published online: 15 June 2022

 Check for updates

Alex M. Jaeger<sup>1</sup>, Lauren E. Stopfer<sup>1,2</sup>, Ryuhjin Ahn<sup>1,2</sup>, Emma A. Sanders<sup>1</sup>, Demi A. Sandel<sup>1</sup>, William A. Freed-Pastor<sup>1,3</sup>, William M. Rideout III<sup>1</sup>, Santiago Naranjo<sup>1</sup>, Tim Fessenden<sup>1</sup>, Kim B. Nguyen<sup>1,4</sup>, Peter S. Winter<sup>1,5,7</sup>, Ryan E. Kohn<sup>1</sup>, Peter M. K. Westcott<sup>1</sup>, Jason M. Schenkel<sup>1,6</sup>, Sean-Luc Shanahan<sup>1</sup>, Alex K. Shalek<sup>1,5,7,8,9</sup>, Stefani Spranger<sup>1,4</sup>, Forest M. White<sup>1,2</sup> & Tyler Jacks<sup>1,4</sup>✉

Immunosurveillance of cancer requires the presentation of peptide antigens on major histocompatibility complex class I (MHC-I) molecules<sup>1–5</sup>. Current approaches to profiling of MHC-I-associated peptides, collectively known as the immunopeptidome, are limited to in vitro investigation or bulk tumour lysates, which limits our understanding of cancer-specific patterns of antigen presentation in vivo<sup>6</sup>. To overcome these limitations, we engineered an inducible affinity tag into the mouse MHC-I gene (*H2-KI*) and targeted this allele to the *Kras*<sup>LSL-G12D/+</sup> *Trp53*<sup>fl/fl</sup> mouse model (KP/K<sup>b</sup>Strep)<sup>7</sup>. This approach enabled us to precisely isolate MHC-I peptides from autochthonous pancreatic ductal adenocarcinoma and from lung adenocarcinoma (LUAD) in vivo. In addition, we profiled the LUAD immunopeptidome from the alveolar type 2 cell of origin up to late-stage disease. Differential peptide presentation in LUAD was not predictable by mRNA expression or translation efficiency and is probably driven by post-translational mechanisms. Vaccination with peptides presented by LUAD in vivo induced CD8<sup>+</sup> T cell responses in naive mice and tumour-bearing mice. Many peptides specific to LUAD, including immunogenic peptides, exhibited minimal expression of the cognate mRNA, which prompts the reconsideration of antigen prediction pipelines that triage peptides according to transcript abundance<sup>8</sup>. Beyond cancer, the K<sup>b</sup>Strep allele is compatible with other Cre-driver lines to explore antigen presentation in vivo in the pursuit of understanding basic immunology, infectious disease and autoimmunity.

The success of cancer immunotherapy has led to an explosion of interest in understanding immune recognition of cancer<sup>1,4</sup>. Studies using preclinical models and patient samples have revealed that responses to immunotherapy depend on the presentation of peptide antigens on MHC-I<sup>4,5</sup>. MHC-I is a heterotrimeric complex that consists of a heavy chain (H2-K and H2-D in C57BL/6 mice, HLA-A, HLA-B and HLA-C in humans), a light chain (β2 microglobulin) and a peptide (generally 8–11 amino acids in length). Peptides presented by MHC-I are derived from the proteolysis of intracellular proteins, giving rise to a diverse array of peptide MHC-I complexes (pMHCS), known collectively as the immunopeptidome<sup>6</sup>. Expression of non-self proteins, such as those present in cancer or virally infected cells, results in the presentation of foreign peptides or neoantigens that are recognized by CD8<sup>+</sup> T cells to drive antitumour immune responses<sup>9</sup>.

Numerous groups have contributed to the understanding of MHC-I trafficking, peptide loading onto MHC-I and the biochemical features of the immunopeptidome<sup>1,8,10–13</sup>. Notably, mass spectrometry analyses have characterized endogenously presented peptides in mouse and human cells and tissues, which has resulted in improved prediction

algorithms for peptide–MHC binding and designs of personalized, neoantigen vaccines<sup>10,11,14–16</sup>. Although these advances have improved the design of antigen-specific therapies, our understanding of the dynamic, context-dependent immunopeptidome is still lacking.

Proteomics studies interrogating the immunopeptidome have been carried out almost universally using antibody or affinity immunoprecipitation of MHC-I peptides from cells grown in culture, which lack the microenvironmental and/or tissue-specific stimuli that are certain to affect the repertoire of peptides presented by cancer cells<sup>10,14,15</sup>. Significantly less focus has been paid to understand the immunopeptidome in vivo, for which existing studies generally profile bulk tissue or tumour lysates and are obscured by heterogenous cell mixtures.

A more precise comparison of tumour and normal immunopeptidomes is likely to uncover new cancer-specific epitopes. Unfortunately, the field currently lacks specialized tools to precisely isolate MHC-I peptides from cells of interest in vivo. Genetically engineered mouse models (GEMMs)—such as the *Kras*<sup>LSL-G12D/+</sup> *Trp53*<sup>fl/fl</sup> (KP) model—recapitulate the histopathological features of human cancer and are

<sup>1</sup>David H. Koch Institute for Integrative Cancer Research, Cambridge, MA, USA. <sup>2</sup>Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA.

<sup>3</sup>Department of Medical Oncology, Dana Farber Cancer Institute, Harvard Medical School, Boston, MA, USA. <sup>4</sup>Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA. <sup>5</sup>Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, MA, USA. <sup>6</sup>Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA. <sup>7</sup>Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA, USA. <sup>8</sup>Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA.

<sup>9</sup>Ragon Institute of MGH, Massachusetts Institute of Technology and Harvard, Cambridge, MA, USA. ✉e-mail: tjacks@mit.edu

tractable systems for studying tumour progression in the native tissue microenvironment<sup>7</sup>. Therefore, GEMMs represent an underappreciated tool to interrogate the tumour immunopeptidome at distinct stages of tumour progression and to uncover features of tumour antigen presentation in vivo that have thus far remained elusive<sup>17</sup>.

Here we report a new GEMM that enables the specific purification of pMHCs from cells of interest in vivo. Using this tool, we show that the cancer immunopeptidome reflects a loss of cell identity through tumour evolution and provide compelling evidence that patterns of cancer-specific antigen presentation are not driven by changes in RNA abundance or translational efficiency. Finally, we identify immunogenic epitopes that are presented on cancer cells in vivo and provide evidence that the universe of targetable antigens in cancer is potentially broader than currently appreciated.

### Inducible affinity tagged MHC-I GEMM

We engineered a Cre-inducible exon that encodes the highly specific affinity tag StrepTagII into intron 1 of the *H2-K1* locus (which encodes the H2-K<sup>b</sup> alloantigen of MHC-I) (K<sup>b</sup>Strep; Fig. 1a). This design results in the Cre-dependent incorporation of the StrepTagII onto the mature H2-K<sup>b</sup> protein. We targeted the *K<sup>b</sup>Strep* allele to embryonic stem cells derived from mice harbouring the KP genotype. Thus, Cre recombinase activates oncogenic *Kras*<sup>G12D</sup> with simultaneous biallelic deletion of *Trp53* and activation of the StrepTagII, which enables the specific purification of MHC-I complexes from autochthonous tumours in vivo (Fig. 1b and Extended Data Fig. 1a–c).

To validate this system, we derived pancreatic organoids from KP mice and KP/K<sup>b</sup>Strep mice and confirmed that the *K<sup>b</sup>Strep* allele was Cre-inducible, presented on the cell surface and enabled the purification of intact MHC-I complexes (Extended Data Fig. 1d–j). We next validated our system in vivo and confirmed that the StrepTagII was specifically detected on cancer cell nests in the tumour microenvironment (Extended Data Fig. 2a–c).

For immunopeptidome profiling, we isolated peptides from cells grown in a monolayer (two-dimensional) and from orthotopically transplanted KP/K<sup>b</sup>Strep organoids in vivo (Ortho). We also isolated peptides from autochthonous tumours initiated through retrograde pancreatic ductal instillation of adenovirus expressing Cre recombinase into KP/K<sup>b</sup>Strep mice (Auto) and into KP/K<sup>b</sup>WT mice (Auto-WT)<sup>18</sup>. We extracted peptides with the expected characteristics<sup>12</sup> of H2-K<sup>b</sup> binders in length, predicted affinity and amino acid content from all samples, except for the negative control Auto-WT samples (Extended Data Fig. 2d–j).

We next compared peptides isolated from orthotopic and autochthonous pancreatic ductal adenocarcinoma (PDAC) using our K<sup>b</sup>Strep system with those identified in normal pancreas tissue from a previous study<sup>19</sup>. Antibody-based immunoprecipitation of H2-K<sup>b</sup> revealed numerous peptides specific to PDAC tissue (Extended Data Fig. 2k). Moreover, analysis of source protein gene expression in single-cell RNA sequencing (scRNA-seq) data indicated that normal pancreas MHC-I peptides were nonspecifically sampled across all cell types in the pancreas. By contrast, peptides identified in orthotopic and autochthonous samples were enriched for ductal cell features that reflected the expected cellular phenotype of malignant cells in PDAC (Tabula Muris; Extended Data Fig. 2l–m). Collectively, these results validate that the KP/K<sup>b</sup>Strep system enables high-resolution interrogation of cell-type-specific immunopeptidomes in vivo.

### Profiling the LUAD immunopeptidome in vivo

We next applied the KP/K<sup>b</sup>Strep model to autochthonous LUAD (Fig. 1c–h and Extended Data Fig. 3). Using antibody-based immunoprecipitation, we isolated H2-K<sup>b</sup> peptides from healthy lung (Normal-Ab) and 16-week tumour-bearing lung (Tumour-Ab) and compared the peptides to those captured by cancer-cell-specific StrepTactin affinity purification from 16-week KP/K<sup>b</sup>Strep tumours (Tumour-Strep) or

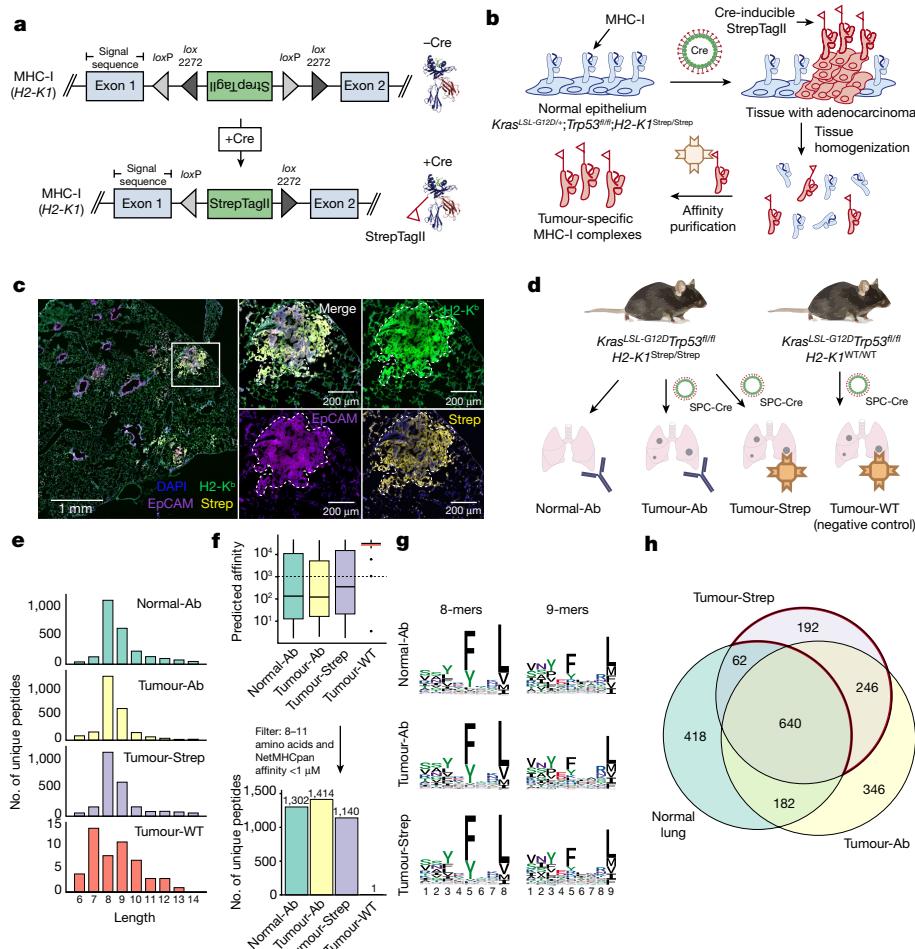
negative control KP/K<sup>b</sup>WT tumours (Tumour-WT) (Fig. 1d). Peptides isolated from all samples had length distributions, predicted affinities and amino acid motifs that reflect K<sup>b</sup> binding, except for the negative control Tumour-WT samples (Fig. 1e–g and Extended Data Fig. 3j–m). In vitro biochemical experiments also demonstrated that StrepTagII incorporation did not affect MHC-I stability, trafficking or recognition by T cells (Extended Data Fig. 4).

A comparison of the identities of peptides derived from the Normal-Ab, Tumour-Ab and Tumour-Strep samples uncovered 438 peptides that were specific to tumours (Fig. 1h, red outline). Although 246 of these peptides were also identified using the traditional antibody approach in tumour-bearing lungs, 192 were specific to Tumour-Strep samples, which suggests that cell-type-specific isolation of MHC-I peptides can provide deeper coverage of the immunopeptidome for cells of interest. A total of 346 peptides were specific to the Tumour-Ab samples, but given the lack of specificity with this method, we could not conclude whether these peptides were derived from neoplastic cancer cells.

### The immunopeptidome reflects cell identity

We next evaluated whether in vivo immunopeptidomes capture specific cellular identities or cell states within native tissue microenvironments. We applied gene signatures derived from peptides identified from healthy lung (Normal), antibody purification from tumour-bearing lung (Ab) or affinity purification of MHC-I specifically from cancer cells (Strep) to scRNA-seq data from healthy mouse lung<sup>20</sup> (Tabula Muris; Fig. 2a and Extended Data Fig. 5a,b). Direct comparison of the Strep and Normal signatures across all cell types revealed a highly significant enrichment for an alveolar type 2 (AT2) phenotype in the Strep signature. This result was consistent with tumour initiation driven by Cre recombinase expressed from a *SPC* promotor specific for AT2 cells (Fig. 1d and Fig. 2b, left). AT2 enrichment was not as pronounced when comparing the Ab and Normal signatures (Fig. 2b, middle). Notably, comparison of the Strep and Ab signatures also resulted in strong enrichment for an AT2 phenotype in Strep samples, even though both datasets were derived from identical tumour-bearing lung tissues (Fig. 2b, right). This result suggests that peptides isolated using non-specific antibodies from MHC-I complexes obtained from bulk tumour lysates are contaminated by peptides from tumour-infiltrating immune cells and other stromal cells in the microenvironment.

Recent data have demonstrated that KPLUAD tumours progressively lose AT2 identity, which prompted us to directly compare LUAD and AT2 immunopeptidomes<sup>21–23</sup>. To accomplish this, we crossed mice harbouring the *K<sup>b</sup>Strep* allele to *Sftpc*<sup>creERT2</sup> mice<sup>24</sup>, which enabled tamoxifen-inducible incorporation of the StrepTagII specifically on AT2 cells in healthy lung tissue (Fig. 2c and Extended Data Fig. 5c–e). MHC-I peptides isolated from AT2 cells in vivo strongly reflected an AT2 cell identity (Extended Data Fig. 5f,g). We next evaluated the LUAD immunopeptidome with respect to tumour progression at 8 weeks (early), 12 weeks (mid), and 16 weeks (late) and compared these data with bulk lung and AT2 peptides (Fig. 2c and Extended Data Fig. 5c–e). Comparison of MHC-I peptides isolated from early-stage, mid-stage and late-stage tumours to those found in normal AT2 cells or lung tissue revealed that the tumour immunopeptidome progressively diverges from normal (Fig. 2d,e). We next derived signatures from genes encoding peptides identified in early-stage, mid-stage and late-stage tumours and applied them to AT2, club and bronchioalveolar stem cell (club/BASC), and basal cell subsets from the healthy lung scRNA-seq data. The immunopeptidome signatures exhibited progressive decline in signal within the AT2 compartment throughout tumour progression (Fig. 2f). By contrast, club/BASC cells and basal cells exhibited slightly increased association with mid-stage and late-stage signatures, which indicates that as tumour cells lose AT2 identity and adopt alternative cellular phenotypes, a parallel alteration is observed in the tumour immunopeptidome<sup>25</sup> (Fig. 2f).



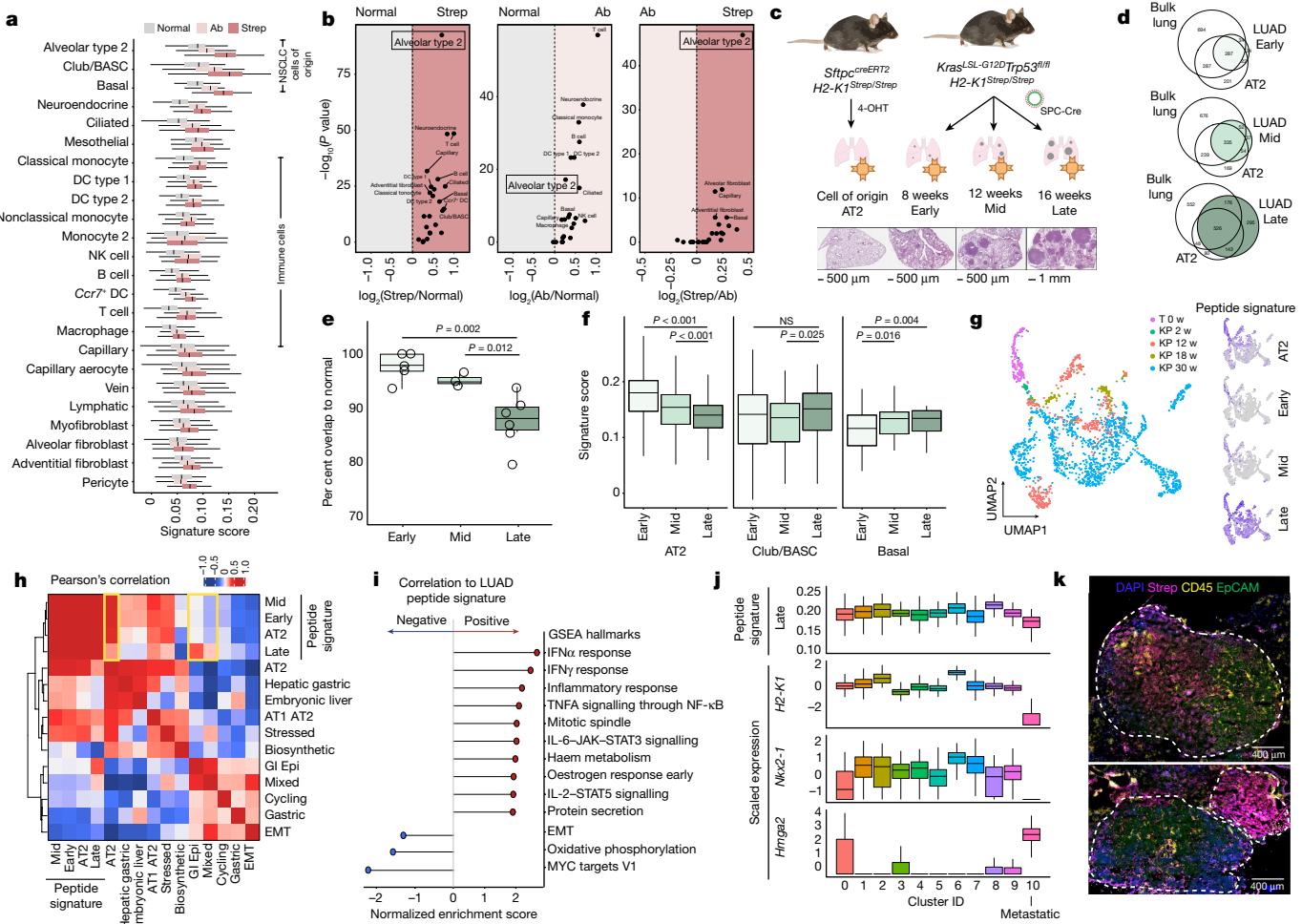
**Fig. 1 | Design and validation of the KP/K<sup>b</sup>Strep mouse model.** **a**, A Cre invertible exon encoding the StrepTagII epitope was inserted into intron 1 of the *H2-KI* gene (top). Cre recombination induces incorporation of the StrepTagII onto the amino-terminus of MHC-I (bottom). **b**, Schematic depicting how Cre activation of K<sup>b</sup>Strep enables tumour-specific isolation of MHC-I in autochthonous tumours. **c**, Left, multiplex immunofluorescence of a representative KP/K<sup>b</sup>Strep lung 8 weeks after tumour initiation. Right, zoomed images of the white box in the left image. **d**, Schematic depicting the different *in vivo* sample types that were compared in downstream analyses. **e**, Length distribution of peptides isolated from healthy lung (Normal-Ab), KP/K<sup>b</sup>Strep tumour-bearing lung with anti-H2-K<sup>b</sup> antibody (Tumour-Ab), KP/K<sup>b</sup>Strep

tumour-bearing lung with StrepTactin affinity purification (Tumour-Strep) and wild-type KP tumour-bearing lung with StrepTactin affinity purification (Tumour-WT, negative control). **f**, Number of unique peptides identified in each sample type after filtering for length (8–11 amino acids) and NetMHCPan predicted affinity (<1,000 nM). Boxes depict first, second and third quartiles and whiskers show range excluding outliers. **g**, Peptide motifs of 8-mers and 9-mers isolated from Normal-Ab, Tumour-Ab and Tumour-Strep samples. **h**, Venn diagram showing the relationship between peptides identified in normal lung, tumour-bearing lung with antibody (Tumour-Ab) and tumour-specific MHC-I purification with StrepTactin (Tumour-Strep). Peptides specific to cancer cells are outlined in red.

## Tumour evolution alters the immunopeptidome

To further understand the LUAD immunopeptidome in the context of tumour evolution, we applied AT2, early, mid and late peptide-derived gene signatures to published scRNA-seq data from normal AT2 cells (T 0 w) and KP tumour cells throughout progression (KP 2 w, KP 12 w, KP 18 w and KP 30 w)<sup>22</sup> (Fig. 2g). We observed dynamic changes in the correlation of peptide signatures with previously described expression modules in the KP model<sup>22</sup> (Fig. 2g). Notably, the late-stage signature exhibited decreased correlation to the AT2 module and increased correlation to the gastric epithelial module and the highly mixed transcriptional module, which were previously found to be associated with phenotypic plasticity in late-stage KP tumour evolution (Fig. 2h and Extended Data Fig. 5j–l). Moreover, gene set enrichment analysis revealed that inflammatory cytokine signalling was highly correlated with the late peptide signature, whereas Myc signalling, metabolic processes and epithelial-to-mesenchymal transition (EMT) were inversely correlated with the late signature (Fig. 2i). Additionally,

the metastatic cluster of KP tumour cells (cluster 10, low *Nkx2-1*, high *Hmga2*) exhibited low expression of *H2-KI* and was the cluster with the lowest expression of the late signature (Fig. 2j). Heterogeneity of antigen presentation across cancer cell states was further supported by multiplexed immunofluorescence of late-stage tumours, which revealed significant intratumour and intertumour heterogeneity in MHC-I presentation (Fig. 2k). Consistent with this notion, dampening tumour inflammation with CD8<sup>+</sup> immune cell depletion decreased the number of unique MHC-I peptides we recovered. By contrast, inducing CD8<sup>+</sup> T cell infiltration with an agonistic CD40 antibody and a FLT-3 ligand (CD40/FLT3L) increased previously unknown peptide identification, highlighting that antigen presentation is responsive to inflammatory cues in the tumour microenvironment<sup>26</sup> (Extended Data Fig. 5m,n). Collectively, these results demonstrate that antigen presentation by cancer cells is a highly integrative and complex process, subject to the evolution of cell identity and to microenvironmental cues *in vivo*. Importantly, our system captures each of these important features.



**Fig. 2 | The LUAD immunopeptidome is dynamic and heterogeneous throughout tumour evolution.** **a**, Comparison of the relative expression of gene signatures derived from Normal, Tumour-Ab or Tumour-Strep peptides across all cell types detected by scRNA-seq from healthy lung tissue (data from Tabula Muris). DC, dendritic cell; NK natural killer. **b**, Volcano plots for pairwise comparisons of peptide signatures for each cell type are shown on the right. **c**, Schematic showing the samples used for comparison of the immunopeptidome from the AT2 cell of origin up to tumour progression. **d**, Overlap between peptides identified in early-stage, mid-stage and late-stage tumours versus bulk lung tissue and AT2 cells. **e**, Quantification of the per cent overlap from early-stage ( $n = 5$ ), mid-stage ( $n = 3$ ), and late-stage ( $n = 6$ ) tumour peptides with normal peptides (bulk lung plus AT2). **f**, *P* values calculated using two-sided Student's *t*-test. **g**, Relative expression of signatures derived from peptides identified in early-stage (E), mid-stage (M) and late-stage (L) tumours in AT2, club/BASC and basal cells in the healthy lung. **h**, *P* values calculated using

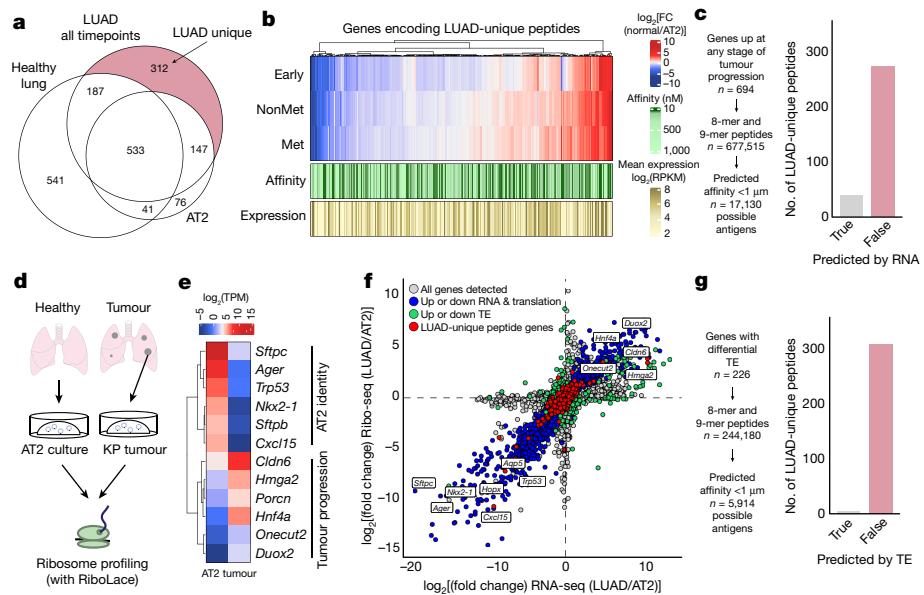
Mann–Whitney *U*-test. **g**, Uniform manifold approximation and projection (UMAP) embedding of re-analysed scRNA-seq data from ref.<sup>22</sup> showing all AT2 (T0w) and KP cells (KP 2–30w) throughout tumour progression (left) and expression of the AT2, early-stage, mid-stage and late-stage peptide signatures (right). **h**, Pearson's correlation of AT2, early-stage, mid-stage and late-stage signatures versus all gene modules described in ref.<sup>22</sup>. GI Epi, gastric epithelial. **i**, Gene set enrichment analysis (GSEA) of genes ranked according to their correlation to the late-stage tumour peptide signature. Gene sets positively correlated to the late-stage signature are shown in red, and negatively correlated are shown in blue. **j**, Relative expression of the late-stage signature, *Nkx2-1*, *Hmg2* or *H2-K1* across clusters found in KP scRNA-seq data. **k**, Multiplexed immunofluorescence depicting tumour-specific MHC-I expression (Strep, magenta) of a single late-stage KP/K<sup>b</sup>Strep tumour (top) or two adjacent tumours (bottom). Dashed lines outline individual tumours. Boxes depict first, second and third quartiles and whiskers show range excluding outliers (a, e, f, j).

## Features of LUAD-unique peptides

Although we noted that gene signatures derived from the totality of the immunopeptidome reflected broad cellular identities as measured by RNA-seq, we next sought to understand the features of peptides that were specifically presented in KP tumours compared to healthy bulk lung or AT2 cells (Fig. 3a, LUAD-unique). To understand the relationship between transcription and presentation of LUAD-unique peptides, we used published bulk RNA-seq data from sorted normal AT2 cells or from cells sorted from early-stage, mid-stage or late-stage KP tumours<sup>27</sup> (Fig. 3b). Genes encoding LUAD-unique peptides exhibited varied patterns of RNA expression across timepoints, which did not correlate with either mean mRNA expression levels or predicted peptide affinity. This result indicates that

changes in mRNA expression cannot exclusively explain LUAD-unique presentation of individual peptides. To corroborate this finding, we identified genes that were upregulated at any stage of tumour progression and predicted the affinity of all possible 8-mer and 9-mer peptides from this list (Fig. 3c). This analysis identified 17,130 peptides predicted to be enriched in tumour presentation based on mRNA expression and predicted affinity alone. However, this pipeline only identified 39 out of 312 total 8-mer and 9-mers in our LUAD-unique peptide list, which reinforces the importance of evaluating cell-type-specific and tissue-specific MHC-I presentation with empirical, mass spectrometry analysis (Fig. 3c).

We next assessed whether protein synthesis in tumour or normal cells could better predict the presentation of LUAD-unique peptides. To this end, we derived organoid cultures of normal AT2 cells and KP



**Fig. 3 | Transcription and translation of LUAD-unique peptides.** **a**, Venn diagram depicting the identification of LUAD-unique MHC-I peptides. **b**, Heatmap showing the relative mRNA expression (red, white, blue), mean mRNA expression (gold) and predicted affinity (green) of genes encoding LUAD-unique peptides throughout tumour progression compared with normal AT2 cells (data from ref. <sup>27</sup>). Met, metastatic; NonMet, non-metastatic. **c**, Workflow depicting the in silico approach used to predict tumour-specific peptides based on RNA expression and predicted affinity. Histograms show the number of LUAD-unique peptides according to whether they were predicted (grey) or not predicted (red) by RNA and affinity analysis. **d**, Experiment outline for Ribo-seq using RiboLace. **e**, Heatmap showing the relative

translation intensity (TPM) for AT2 identity genes and genes associated with KP tumour progression. **f**, Comparison of RNA-seq abundance (x axis) and Ribo-seq abundance (y axis) in tumour organoids versus AT2 organoids. Genes coordinately upregulated or downregulated in both RNA-seq and Ribo-seq are shown in blue, genes exhibiting differential TE are shown in green, and genes encoding for LUAD-unique peptides are shown in red. **g**, Workflow depicting the approach used to predict LUAD-unique peptides based on differential TE in LUAD versus AT2 and predicted affinity. Histograms show the number of LUAD-unique peptides according to whether they were predicted (grey) or not predicted (red) by differential TE.

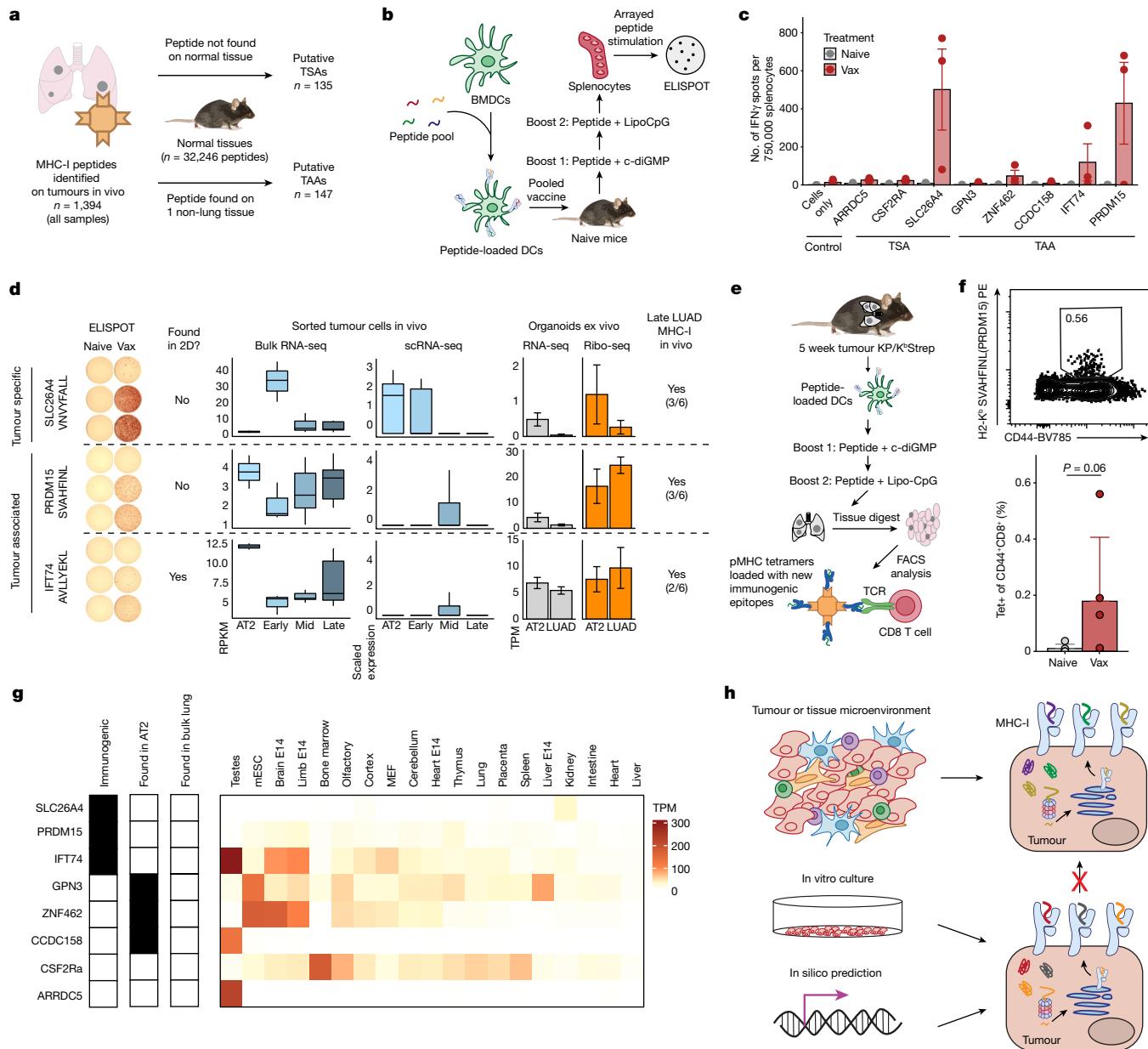
LUAD cells and performed ribosome profiling (Ribo-seq) and RNA-seq ex vivo<sup>28–30</sup> (Fig. 3d and Extended Data Fig. 6a–d). Genes associated with AT2 identity (*Sftpc*, *Sftpb*, *Ager*, *Nkx2-1* and *Cxcl15*) exhibited higher translation rates in the AT2 samples, whereas genes associated with KP LUAD progression (*Onecut2*, *Duo2*, *Hmg2*, *Porcn*, *Cldn6* and *Hnf4a*) were translated at higher rates in LUAD samples (Fig. 3e). Integrated analysis of RNA-seq and Ribo-seq data revealed numerous genes that were coordinately upregulated or downregulated at the mRNA and ribosome protected fragment (RPF) levels (Fig. 3f, blue) or exhibited differential translation efficiency (TE) (Fig. 3f, green, and Extended Data Fig. 6e–g). However, genes encoding LUAD-unique peptides generally exhibited no difference in mRNA or RPF abundance (Fig. 3f, red), and prediction of peptides based on differential TE resulted in only 4 out of 312 empirically identified, LUAD-unique peptides (Fig. 3g).

The inability of mRNA abundance or TE to fully describe the LUAD-unique immunopeptidome prompted us to consider post-translational features of tumour-specific presentation<sup>31</sup>. Although source protein length and thermal stability were similar in normal and LUAD-unique immunopeptidomes, we observed differences in source protein localization and a trend towards decreased protein half-life in source proteins giving rise to LUAD-unique peptides (Extended Data Fig. 7). We next sought to perturb post-translational processes through inhibition of the molecular chaperone heat shock protein 90 kDa (HSP90), which can reshape the tumour immunopeptidome and stimulate antitumour immune responses<sup>32</sup>. Treatment with an HSP90 inhibitor (HSP90i) increased the number of specific MHC-I peptides we identified, including those derived from HSP90 client proteins with lower thermal stability (Extended Data Fig. 8a–k). Taken together, these data demonstrate that the immunopeptidome can be shaped through post-translational mechanisms and, more broadly, underscore the potential of the KP/K<sup>b</sup>Strep model to discover treatment-induced changes in the tumour immunopeptidome in vivo (Extended Data Fig. 8l).

## Discovery of tumour antigens in KP LUAD

Given our ability to discern the LUAD immunopeptidome from that of normal tissues in vivo, we evaluated new tumour epitopes in the KP/K<sup>b</sup>Strep model. We found 135 peptides that were recurrently presented on LUAD in vivo but not observed in healthy tissues from our study or previously published data<sup>19</sup>, which we termed putative non-mutated tumour-specific antigens (TSAs) (Fig. 4a and Extended Data Fig. 9a). In addition, we nominated an additional 147 peptides that were only found on one tissue (not lung) and determined these peptides to be putative tumour-associated antigens (TAAs) (Fig. 4a and Extended Data Fig. 9b). Seeking to evaluate the immunogenicity of both classes of antigens, we vaccinated naïve mice with a pooled peptide vaccine (three TSAs, five TAAs) using bone-marrow derived dendritic cells<sup>33</sup> (Fig. 4b). Interferon-γ (IFNγ) enzyme-linked immune adsorbent spot (ELISPOT) assays of splenocytes from control and vaccinated mice revealed three immunogenic peptides (Fig. 4c and Extended Data Fig. 10). Notably, two out of three peptides were not presented by KP/K<sup>b</sup>Strep cells in vitro, and comparative RNA or RPF expression in LUAD compared to AT2 cells would not have led to their prioritization (Fig. 4d and Extended Data Fig. 9c–f). Collectively, these data reinforce the importance of examining the immunopeptidome in vivo and challenge the notion that differential expression is required for tumour-specific antigen presentation.

To further explore the immunogenicity of these epitopes, we vaccinated 5-week KP tumour-bearing animals and measured CD8<sup>+</sup> T cell reactivity with custom peptide–MHC-I tetramers<sup>34</sup> (Fig. 4e). We detected CD8<sup>+</sup> T cells that recognized a peptide derived from the TAA PRDM15 (SVAHFNL) in tumour-bearing lung tissue in vaccinated mice, but not control mice (Fig. 4f and Extended Data Fig. 9h). Seeking to understand the lack of immunogenicity of the remaining five out of eight vaccine peptides, we cross-referenced our AT2 immunopeptidome data and



**Fig. 4 | Discovery of new tumour antigens in LUAD.** **a**, Workflow for identifying putative non-mutated TSAs or TAAs. **b**, Schematic of the pooled peptide vaccine strategy in naive mice. **c**, Quantification of IFN $\gamma$  ELISPOT data of splenocytes from naive mice or mice vaccinated (Vax) with pooled peptides. Each peptide in the pool was used individually to stimulate splenocytes before ELISPOT. Data are mean  $\pm$  s.e.m.,  $n = 3$  mice per group. **d**, Comparison of in vivo bulk RNA-seq (data from ref.<sup>27</sup>), scRNA-seq (data from ref.<sup>22</sup>) and ex vivo RNA-seq and Ribo-seq for genes encoding immunogenic peptides. Identification of peptides by in vitro and in vivo immunopeptidomics is also indicated. For in vivo peptide identification, the fraction of late-stage tumour samples for which the peptide was identified is indicated. Boxes depict first, second and third quartiles and whiskers show range excluding outliers.

Histograms depict mean  $\pm$  s.e.m., ex vivo organic data include  $n = 3$  for AT2 and  $n = 4$  for LUAD. **e**, Pooled vaccine strategy for KP tumour-bearing animals. **f**, Flow cytometry plots depicting pMHC-I tetramer staining for a representative TAA (SVAHFNL, PRDM15) in the lung tissue of naive and vaccinated, tumour-bearing mice.  $P$  values calculated using one-sided Mann-Whitney test. Data are mean  $\pm$  s.e.d.,  $n = 5$  mice in each group. **g**, Transcript abundance across healthy mouse tissues for peptides included in the pooled vaccine. Detection of the peptide on AT2 cells or bulk lung tissue is also indicated. E14, embryonic day 14. **h**, Model depicting the incongruence between the immunopeptidome derived from in silico prediction methods, in vitro mass spectrometry and tumour-specific in vivo immunopeptidomics.

found that three out of five were presented in AT2 but not bulk lung tissue (Fig. 4g). RNA expression patterns across healthy mouse tissue (Mouse Encode Project) for genes encoding peptides presented by AT2 cells would incorrectly predict potential cancer testis antigens (*Ccdc158*) or oncofetal antigens (*Gpn3* and *Znf462*). These analyses demonstrate that in vivo, cell-type-specific immunopeptidomics provides an opportunity to empirically evaluate cell-specific and tissue-specific presentation patterns and more accurately classify potential antigens

compared to in vitro or in silico approaches (Fig. 4h and Extended Data Fig. 9j,k).

## Discussion

We developed a mouse model that enabled the specific isolation of MHC-I complexes from discrete cell populations of interest in vivo. Application of this system to AT2 cells and LUAD revealed that the

immunopeptidome is highly dynamic through tumour progression and evolves with cellular states adopted by tumour cells. This raises the possibility that distinct stages of tumour development are associated with an ever-changing landscape of targets for antigen-specific T cells through tumour progression.

Notably, our results suggest that neither mRNA expression nor TE fully explain LUAD-specific antigen presentation *in vivo*. Although many neoantigen prediction pipelines include mRNA abundance as a method to triage potential neoantigens, our data indicate that many peptides are recurrently presented *in vivo* despite apparently low transcript levels, which highlights the importance of empirical mass spectrometry evidence to evaluate MHC-I presentation. Further study is needed to elucidate post-translational mechanisms that shape tumour-specific antigen presentation *in vivo*, which may ultimately improve neoantigen prediction.

TSA and TAA are known to elicit antigen-specific responses against tumours<sup>35–38</sup>. Consistent with recent work from our laboratory<sup>26,39–41</sup>, we found that vaccination was required to induce responses against antigens identified *in vivo*, which is probably due to suboptimal priming of CD8<sup>+</sup> T cells in our model. Systematic interrogation of epitopes identified with this mouse model will probably uncover additional immunogenic epitopes. These strategies will ultimately expand our understanding of the ‘altered-self’ presented by cancer cells, which can be used to design and evaluate new cancer immunotherapies in preclinical models<sup>42</sup>.

We also envision that this model system can identify peptides derived from cryptic translation events<sup>29,43–45</sup>, post-translational modifications<sup>46</sup>, transposable elements<sup>47</sup> and the microbiome<sup>48</sup>, as many of these events will be influenced by physiological cues from the *in vivo* microenvironment.

Given the prevalence of cell-specific and tissue-specific Cre-driver mouse strains, the *K<sup>d</sup>**Strep* allele presents an opportunity to map a high-resolution *in vivo* immunopeptidome atlas<sup>19</sup>, as we demonstrated with AT2 cells in this study. Cross-referencing these data to other omics-based tissue atlases can elucidate the relationship between cellular phenotype and the immunopeptidome, which cannot be accomplished using bulk tissue measurements.

In summary, we uncovered important aspects of the tumour immunopeptidome *in vivo* that should promote further investigation into context-specific antigen presentation and will improve our understanding of tumour-immune interactions. In turn, we described a versatile tool for the broader research community to interrogate mechanisms of antigen presentation at high resolution in health and disease.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-022-04839-2>.

- Jhunjhunwala, S., Hammer, C. & Delamarre, L. Antigen presentation in cancer: insights into tumour immunogenicity and immune evasion. *Nat. Rev. Cancer* **21**, 298–312 (2021).
- Donia, M. et al. Acquired immune resistance follows complete tumor regression without loss of target antigens or IFN $\gamma$  signaling. *Cancer Res.* **77**, 4562–4566 (2017).
- Yamamoto, K. et al. Autophagy promotes immune evasion of pancreatic cancer by degrading MHC-I. *Nature* **581**, 100–105 (2020).
- Schumacher, T. N. & Schreiber, R. D. Neoantigens in cancer immunotherapy. *Science* **348**, 69–74 (2015).
- Keenan, T. E., Burke, K. P. & Allen, E. M. V. Genomic correlates of response to immune checkpoint blockade. *Nat. Med.* **25**, 389–402 (2019).
- Dersh, D., Holly, J. & Yewdell, J. W. A few good peptides: MHC class I-based cancer immuno-surveillance and immuno-evasion. *Nat. Rev. Immunol.* **21**, 116–128 (2021).
- DuPage, M., Dooley, A. L. & Jacks, T. Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. *Nat. Protoc.* **4**, 1064–1072 (2009).
- Wells, D. K. et al. Key parameters of tumor epitope immunogenicity revealed through a consortium approach improve neoantigen prediction. *Cell* **183**, 818–834.e13 (2020).
- Ghorani, E. et al. The T cell differentiation landscape is shaped by tumour mutations in lung cancer. *Nat. Cancer* **1**, 546–561 (2020).
- Abelin, J. G. et al. Mass spectrometry profiling of HLA-associated peptidomes in mono-allelic cells enables more accurate epitope prediction. *Immunity* **46**, 315–326 (2017).
- Chong, C. et al. Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. *Nat. Commun.* **11**, 1293 (2020).
- Yadav, M. et al. Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature* **515**, 572–576 (2014).
- Granados, D. P. et al. Impact of genomic polymorphisms on the repertoire of human MHC class I-associated peptides. *Nat. Commun.* **5**, 3600 (2014).
- Sarkizova, S. et al. A large peptidome dataset improves HLA class I epitope prediction across most of the human population. *Nat. Biotechnol.* **38**, 199–209 (2019).
- Stopfer, L. E., Mesfin, J. M., Joughin, B. A., Lauffenburger, D. A. & White, F. M. Multiplexed relative and absolute quantitative immunopeptidomics reveals MHC I repertoire alterations induced by CDK4/6 inhibition. *Nat. Commun.* **11**, 2760 (2020).
- Khodadoust, M. S. et al. Antigen presentation profiling reveals recognition of lymphoma immunoglobulin neoantigens. *Nature* **543**, 723–727 (2017).
- Jamal-Hanjani, M. et al. Tracking the evolution of non-small-cell lung cancer. *N. Engl. J. Med.* **376**, 2109–2121 (2017).
- Chiou, S.-H. et al. Pancreatic cancer modeling using retrograde viral vector delivery and *in vivo* CRISPR/Cas9-mediated somatic genome editing. *Genes Dev.* **29**, 1576–1585 (2015).
- Schuster, H. et al. A tissue-based draft map of the murine MHC class I immunopeptidome. *Sci. Data* **5**, 180157 (2018).
- Schaum, N. et al. Single-cell transcriptomics of 20 mouse organs creates a *Tabula Muris*. *Nature* **562**, 367–372 (2018).
- LaFave, L. M. et al. Epigenomic state transitions characterize tumor progression in mouse lung adenocarcinoma. *Cancer Cell* **38**, 212–228.e13 (2020).
- Marijanovic, N. D. et al. Emergence of a high-plasticity cell state during lung cancer evolution. *Cancer Cell* **38**, 229–246.e13 (2020).
- Concepcion, C. P. et al. SMARCA4 inactivation promotes lineage-specific transformation and early metastatic features in the lung. *Cancer Discov.* **12**, 562–585 (2022).
- Rock, J. R. et al. Multiple stromal populations contribute to pulmonary fibrosis without evidence for epithelial to mesenchymal transition. *Proc. Natl Acad. Sci. USA* **108**, E1475–E1483 (2011).
- Han, X. et al. Transdifferentiation of lung adenocarcinoma in mice with Lkb1 deficiency to squamous cell carcinoma. *Nat. Commun.* **5**, 3261 (2014).
- Schenkel, J. M. et al. Conventional type I dendritic cells maintain a reservoir of proliferative tumor-antigen specific TCF-1<sup>+</sup> CD8<sup>+</sup> T cells in tumor-draining lymph nodes. *Immunity* **54**, 2338–2353.e6 (2021).
- Chuang, C.-H. et al. Molecular definition of a metastatic lung cancer state reveals a targetable CD109–Janus kinase–Stat axis. *Nat. Med.* **23**, 291–300 (2017).
- Clamer, M. et al. Active ribosome profiling with RiboLace. *Cell Rep.* **25**, 1097–1108.e5 (2018).
- Chen, J. et al. Pervasive functional translation of noncanonical human open reading frames. *Science* **367**, 1140–1146 (2020).
- Naranjo, S. et al. Modeling diverse genetic subtypes of lung adenocarcinoma with a next-generation alveolar type 2 organoid platform. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.12.07.471632> (2021).
- Perrin, J. et al. Identifying drug targets in tissues and whole blood with thermal-shift profiling. *Nat. Biotechnol.* **38**, 303–308 (2020).
- Jaeger, A. M. et al. Rebalancing protein homeostasis enhances tumor antigen presentation. *Clin. Cancer Res.* **25**, 6392–6405 (2019).
- Perez, C. R. & Palma, M. D. Engineering dendritic cell vaccines to improve cancer immunotherapy. *Nat. Commun.* **10**, 5408 (2019).
- Saini, S. K. et al. Empty peptide-receptive MHC class I molecules for efficient detection of antigen-specific T cells. *Sci. Immunol.* **4**, eaau9039 (2019).
- Zitvogel, L., Perreault, C., Finn, O. J. & Kroemer, G. Beneficial autoimmunity improves cancer prognosis. *Nat. Rev. Clin. Oncol.* **18**, 591–602 (2021).
- Sahin, U. et al. An RNA vaccine drives immunity in checkpoint-inhibitor-treated melanoma. *Nature* **585**, 107–112 (2020).
- Nelson, C. E. et al. Robust iterative stimulation with self-antigens overcomes CD8<sup>+</sup> T cell tolerance to self- and tumor antigens. *Cell Rep.* **28**, 3092–3104.e5 (2019).
- Vonderheide, R. H. CD40 agonist antibodies in cancer immunotherapy. *Annu. Rev. Med.* **71**, 47–58 (2020).
- Burger, M. L. et al. Antigen dominance hierarchies shape TCF1<sup>+</sup> progenitor CD8 T cell phenotypes in tumors. *Cell* **184**, 4996–5014.e26 (2021).
- Westcott, P. M. K. et al. Low neoantigen expression and poor T-cell priming underlie early immune escape in colorectal cancer. *Nat. Cancer* **2**, 1071–1085 (2021).
- Freed-Pastor, W. A. et al. The CD155/TIGIT axis promotes and maintains immune evasion in neoantigen-expressing pancreatic cancer. *Cancer Cell* **39**, 1342–1360.e14 (2021).
- Yarmarkovich, M. et al. Cross-HLA targeting of intracellular oncoproteins with peptide-centric CARs. *Nature* **599**, 477–484 (2021).
- Cuevas, M. V. R. et al. Most non-canonical proteins uniquely populate the proteome or immunopeptidome. *Cell Rep.* **34**, 108815 (2021).
- Erhard, F. et al. Improved Ribo-seq enables identification of cryptic translation events. *Nat. Methods* **15**, 363–366 (2018).
- Ouspenskaia, T. et al. Unannotated proteins expand the MHC-I-restricted immunopeptidome in cancer. *Nat. Biotechnol.* **40**, 209–217 (2022).
- Raposo, B. et al. T cells specific for post-translational modifications escape intrathymic tolerance induction. *Nat. Commun.* **9**, 353 (2018).
- Griffin, G. K. et al. Epigenetic silencing by SETDB1 suppresses tumour intrinsic immunogenicity. *Nature* **595**, 309–314 (2021).
- Kalaora, S. et al. Identification of bacteria-derived HLA-bound peptides in melanoma. *Nature* **592**, 138–143 (2021).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

# Article

## Methods

### Mice

All animal studies described in this study were approved by the MIT Institutional Animal Care and Use Committee. All animals were maintained on a pure C57BL/6J genetic background, except for embryonic stem cell (ESC) chimeras, which were a mix of C57BL/6J and albino C57BL/6J. Generation of *Kras<sup>LSL-G12D/+</sup>* and *Trp53<sup>rfl/rfl</sup>* (KP) mice has previously been described<sup>49,50</sup>. *Sftpc<sup>creERT2/creERT2</sup>* mice on a C57BL/6J background were purchased from the Jackson Laboratory (strain 028054)<sup>24</sup> and bred to *H2-K1<sup>Strep/Strep</sup>* mice to produce the *Sftpc<sup>creERT2/creERT2</sup>H2-K1<sup>Strep/Strep</sup>* strain.

### Mouse ESC culture and CRISPR-assisted gene targeting

Targeted insertion of the Cre-invertible StrepTag allele (*K<sup>b</sup>Strep*) into the endogenous *H2-K1* locus was performed in KP\* ESCs, which were generated by crossing a hormone-primed C57BL/6J *Trp53<sup>rfl/rfl</sup>* female with a C57BL/6J *Kras<sup>LSL-G12D/+</sup>Trp53<sup>rfl/rfl</sup>* male. At 3.5 days after coitum, blastocysts were flushed from the uterus, isolated and cultured on a mouse embryonic fibroblast (MEF) feeder layer in ESCM+LIF+2i (knockout DMEM (Gibco), 15% FBS (Hyclone), 1% NEAA (Sigma), 2 mM glutamine (Gibco), 0.1 mM β-mercaptoethanol (Sigma-Aldrich) 50 IU penicillin, 50 IU streptomycin, 1,000 U ml<sup>-1</sup> LIF (Amsbio), 3 μM CHIR99021 (AbMole) and 1 μM PD0325901 (AbMole)). After 5–7 days in culture, the outgrown inner cell mass was isolated, trypsinized and re-plated on a fresh MEF layer. ESC lines were genotyped for *Kras<sup>LSL-G12D/+</sup>Trp53<sup>rfl/rfl</sup>* and Zfy (Y-chromosome specific). Primer sequences are available upon request. ESC lines were tested for pluripotency by injection into host blastocysts from albino mice to generate chimeric mice.

DNA mixes containing 3:1 mixes of H2-K1-Strep targeting vector: U6-sgH2-K1-eCas9v1.1-T2A-BlastR were ethanol precipitated before Lipofectamine 2000 (Thermo Fisher) transfection of approximately  $3 \times 10^5$  KP\*1 mouse ESCs according to the manufacturer's instructions. Transfected mouse ESCs were plated on MEF feeder cells and selected with 4 μg ml<sup>-1</sup> blasticidin for 48 h. Cells were then trypsinized and plated at low density onto a fresh plate with MEF feeder cells. After 5–6 days, large colonies were picked using a dissecting microscope and replated into a 96-well plate containing MEF feeder cells. After 4–5 days, each well of the 96-well plate was trypsinized, and 1:2 of the material was frozen in ESCM plus 10% DMSO and the remainder of the cells were transferred to a 96-well PCR plate for PCR-based integration screening. Clones containing homozygous targeting events were then thawed and expanded into 24-well dishes, which were then subjected to genomic DNA extraction and overnight restriction digest with *SapI*. Digestions were then electrophoresed on a 0.7% agarose gel before transfer onto an Amersham Hybond XL nylon membrane (GE Healthcare). Blots were then probed with <sup>32</sup>P-labelled DNA probes comprising an internal sequence homologous to the genomic insertion containing the StrepTag exon.

Correctly targeted clones were injected into albino C57BL/6 blastocysts. Coat colour was used as a surrogate marker for chimerism. Low-degree chimeras were chosen for pancreatic organoid generation and high degree chimeras were bred to KP\* mice for germline transmission.

### Pancreatic organoid isolation

Low-degree KP;K<sup>b</sup>Strep chimeras were chosen for organoid isolation using previously described methods<sup>41</sup>. In brief, the pancreas was manually dissected, transferred to a Petri dish and thoroughly minced with a razor blade. Minced tissue was then transferred to a 1.5-ml micro-centrifuge tube with 1 ml of PBS supplemented with 125 U ml<sup>-1</sup> collagenase IV (Worthington) and incubated with rotation at 37 °C for 20–30 min. Cell suspensions were diluted with 9 ml of PBS and centrifuged at 2,000 r.p.m. for 2 min. Cell pellets were then washed with 10 ml of PBS and centrifuged at 2,000 r.p.m. for 2 min. The resulting cell pellet was then resuspended in 100% Matrigel (Corning) and plated

as 50 μl domes and solidified at 37 °C. Organoids were then cultured in organoid complete medium. Purification of cells derived from the targeted ESCs was accomplished with puromycin selection at 6 μg ml<sup>-1</sup> (a puromycin resistance gene is encoded within the LSL cassette upstream of the *Kras<sup>G12D</sup>* allele).

Medium for pancreatic organoids was formulated based on L-WRN cell conditioned medium (L-WRN CM)<sup>51</sup>. In brief, L-WRN CM was generated by collecting 8 days of supernatant from L-WRN cells grown in Advanced DMEM/F12 (Gibco) supplemented with 20% FBS (Hyclone), 2 mM GlutaMAX, 100 U ml<sup>-1</sup> of penicillin, 100 μg ml<sup>-1</sup> of streptomycin and 0.25 μg ml<sup>-1</sup> amphotericin. L-WRN CM was diluted 1:1 in Advanced DMEM/F12 (Gibco) and supplemented with additional RSPO-1 conditioned medium (10% v/v), generated using Cultrex HA-R-Spondin1-Fc 293T cells. The following molecules were also added to the growth medium: B27 (Gibco), 1 μM N-acetylcysteine (Sigma-Aldrich), 10 μM nicotinamide (Sigma-Aldrich), 50 ng ml<sup>-1</sup> EGF (Novus Biologicals), 500 nM A83-01 (Cayman Chemical), 10 μM SB202190 (Cayman Chemical) and 500 nM PGE2 (Cayman Chemical).

Pancreatic organoids were serially passaged with TrypLE Express (Life Technologies). After four passages, KP or KP;K<sup>b</sup>Strep organoids were then subjected to ex vivo transformation by dissociation, mixing with adenoviral Cre (Ad-CMV-Cre; multiplicity of infection of 500) and re-embedding in Matrigel. After 72 h, transformants were selected with Nutlin-3a (10 μM, Sigma Aldrich) to select for loss of p53.

### Orthotopic transplantation of pancreatic organoids

Orthotopic transplantation of organoids was performed as previously described<sup>41</sup>. In brief, animals were anaesthetized using isoflurane, the left subcostal region was depilated (using clippers or Nair) and the surgical area was disinfected with alternating betadine and isopropyl alcohol. A small (approximately 2 cm) skin incision was made in the left subcostal area and the spleen was visualized through the peritoneum. A small incision (about 2 cm) was made through the peritoneum overlying the spleen and the spleen and pancreas were exteriorized using ring forceps. A 30-gauge needle was inserted into the pancreatic parenchyma parallel to the main pancreatic artery and 100 μl (containing  $1.25 \times 10^5$  organoid cells in 50% PBS plus 50% Matrigel) was injected into the pancreatic parenchyma. Successful injection was visualized by the formation of a fluid-filled region within the pancreatic parenchyma without leakage. The pancreas and spleen were gently internalized, and the peritoneal and skin layers were sutured independently using 5-0 Vicryl sutures. All mice received pre-operative analgesia with sustained-release buprenorphine and were followed post-operatively for any signs of discomfort or distress. Organoid-Matrigel mixes were kept on ice throughout the entirety of the procedure to prevent solidification before injection. For orthotopic transplantation, syngeneic C57BL/6J mice (aged 4–12 weeks) were transplanted. Male pancreatic organoids were only transplanted back into male recipients.

### Retrograde pancreatic duct delivery

Retrograde pancreatic duct instillation of lentivirus has been previously described<sup>18,41</sup>. In brief, the ventral abdomen was depilated (using clippers or Nair) 1–2 days before surgery. Animals were anaesthetized with isoflurane and the surgical area was disinfected with alternating betadine and isopropyl alcohol. A small skin incision was made in the anterior abdomen (2–3 cm midline incision extending caudally from the xiphoid process). A subsequent incision was made through the linea alba and incision edges were secured in place with a Colibri retractor. The remainder of the procedure was conducted under a Nikon stereomicroscope. A moistened (with sterile 0.9% saline) sterile cotton swab was used to gently move the left lobe of the liver cranially towards the diaphragm. A second moistened sterile cotton swab was used to gently reposition the colon and small intestine into the right lower abdominal quadrant until the duodenum was visualized. The duodenum was gently repositioned (still in the abdominal cavity) using moistened cotton swabs until the

pancreas, common bile duct and sphincter of Oddi were clearly visualized. The common bile duct and cystic duct were gently separated from the portal vein and the hepatic artery using blunt dissection with Moria forceps. A microclip was placed over the common bile duct (cranial to pancreatic duct branching) to prevent influx of the viral particles into the liver or gallbladder, forcing the viral vector retrograde through the pancreatic duct. To infuse the viral vector, the common bile duct was cannulated with a 30-gauge needle at the level of the sphincter of Oddi, and 150 µl of virus was injected over the course of 30 s. Gentle pressure was applied at the sphincter of Oddi after needle exit to prevent leakage into the abdominal cavity. Subsequently, the microclip and Colibri retractor were removed. The peritoneum was closed using running 5-0 Vicryl sutures. The cutis and fascia were closed using simple interrupted 5-0 Vicryl sutures. The entire procedure was conducted on a circulating warm water heating blanket to prevent intra-operative hypothermia. All mice received pre-operative analgesia with sustained-release buprenorphine and were followed post-operatively for any signs of discomfort or distress. For retrograde pancreatic ductal installation, male mice (aged 3–6 weeks) and female mice (aged 3–8 weeks) were transduced with  $7 \times 10^8$  p.f.u. of Ad5-Pft1a-Cre (University of Iowa Viral Core) in serum-free medium (Opti-MEM, Gibco).

#### Intratracheal administration of adenovirus

Adenovirus expressing Cre recombinase from a *SPC* promoter (Ad5-SPC-Cre, University of Iowa) was prepared by diluting virus stocks to 400 p.f.u. per µl into OptiMEM (Gibco) followed by the addition of CaCl<sub>2</sub> to a final concentration of 10 µM (ref. <sup>7</sup>). Viral suspensions were then mixed and incubated at room temperature for 20 min before placing on ice. Mice were anaesthetized with isoflurane before setting on a custom wire platform to open the mouth. The trachea was then canulated with 22-gauge catheters (Exel), and 50 µl of viral suspension was added to the catheters. Once the mouse aspirated the viral suspension, they were transferred to a pre-warmed cage to recover from anaesthesia and monitored every day for 3 days to ensure recovery from the procedure. Dilutions of virus were used within 1 h of preparation.

#### Flow cytometry of cultured cells and lung tissue

For staining pancreatic organoids, cells were plated into  $4 \times 20$  µl domes per well in a 12-well plate. Cells were either left untreated or incubated with 20 ng ml<sup>-1</sup> IFNγ for 48 h. At the time of collection, organoid medium was aspirated, and the Matrigel domes were mechanically disrupted by vigorous pipetting in PBS. Cell suspensions were then transferred to 15-ml conical tubes passivated with 0.1% BSA and centrifuged at 2,000 r.p.m. for 2 min. The resulting cell pellet was then incubated with 1 ml TrypLE Express for 10–15 min. The dissociation reaction was quenched with 10 ml PBS followed by centrifugation at 2,000 r.p.m. for 2 min. Cell pellets were then resuspended in 200 µl PBS and transferred to a 96-well U-bottom plate for staining. Cells were then incubated with Zombie Aqua fixable viability stain (1:1,000, BioLegend) for 15 min before staining with primary antibodies diluted in PBS plus 2% heat-inactivated FBS (FACS buffer) for 30 min on ice (Supplementary Table 11). Cells were then washed twice with FACS buffer before analysis on a LSR II analytical flow cytometer.

For FACS analysis of tumour-bearing lung tissue, KP or KP/K<sup>b</sup>Strep mice ( $n = 3$  each) were euthanized 12 weeks after tumour initiation. Two minutes before euthanasia, an intravascular staining antibody (anti-CD45-APC-eFluor786, BioLegend) was injected retro-orbitally to stain circulating immune cells. Mice were then euthanized by cervical dislocation, and the lungs were removed and placed on ice. Approximately 100–200 mg of tumour-bearing lung tissue was then thoroughly minced with Noyes scissors before incubation with digestion buffer (HBSS supplemented with 5% HI FBS, 125 U ml<sup>-1</sup> collagenase IV (Worthington) and DNase (Roche) with rotation for 30 min at 37 °C. Cell suspensions were then macerated through 70-µm cell strainers (Corning) and centrifuged at 500g for 5 min. Cell pellets were resuspended in 1 ml

ACK lysis buffer (Gibco) and incubated at room temperature for 5 min before quenching with 10 ml RPMI and 10% HI FBS and centrifugation at 500g for 5 min. Cell pellets were then resuspended in 200 µl FACS buffer and transferred to 96-well U-bottom plates for staining with Fc block (BD Biosciences) for 20 min, Zombie fixable viability stain for 20 min and primary antibody cocktail for 30 min on ice. Cells were then washed twice with FACS buffer before fixation overnight with FOXP3/transcription factor staining buffer set (eBioscience) according to the manufacturer's instructions before analysis on a Fortessa flow cytometer. All flow cytometry data were analysed using FlowJo v.10.

#### Affinity purification of H2-K<sup>b</sup> with StrepTactin

Whenever possible, great care was taken to keep samples ice cold at all times to maintain MHC-I complex stability. For cultured cells, about  $6 \times 10^7$  cells were used for each replicate (4 × 15 cm dishes). In the culture dish, cells were washed twice with PBS before lysis with 2 ml MHC extraction buffer (MEB; 20 mM Tris pH 8.0, 100 mM NaCl, 1 mM EDTA, 1% Triton X-100, 60 mM octyl-glucopyranoside, 6 mM MgCl<sub>2</sub> and 1× HALT protease inhibitors (Pierce))<sup>10</sup>. Cell lysates were then transferred to 2-ml microcentrifuge tubes and supplemented with 20 U benzonase and 10 U avidin (to block endogenously biotinylated protein) before incubating with rotation at 4 °C. Lysates were then cleared by centrifugation at 16,000g for 15 min before incubation with MagStrep Type 3 StrepTactin XT beads (IBA Biosciences).

For pancreatic and lung tissues, tumour-bearing tissue was dissected and immediately lysed or flash frozen in liquid nitrogen for later processing. For lysis, fresh or frozen tissue was quickly minced with Noyes scissors and transferred to a 7-ml glass Dounce homogenizer (Sigma), precooled on ice. MEB (4 ml) was then added, and the tissue was thoroughly homogenized with 10–20 passes of a loose-fitting pestle followed by 5–10 passes of a tight-fitting pestle. Tissue homogenates were transferred to 5-ml centrifuge tubes (Eppendorf) and supplemented with 20 U benzonase and 10 U avidin before incubating with rotation at 4 °C for 20 min and subsequent removal of debris by centrifugation at 16,000g for 15 min.

Before incubating with cleared lysate, StrepTactin beads were equilibrated by magnetizing and washing once with MEB. For each cell culture sample or tissue sample, 1 ml of bead suspension (50 µl bed volume) was used. Equilibrated beads were then added to cleared lysates and incubated with rotation at 4 °C for 1–3 h. After incubation, beads were washed twice with MEB, twice with TBS and twice with 20 mM Tris. On the last wash, suspended beads were transferred to a new Lo-Bind microcentrifuge tube before elution. Strep-tagged H2-K<sup>b</sup> was then eluted from the StrepTactin resin by adding 400 µl of 0.5× buffer BXT (IBA Biosciences) and incubating on ice for 20–30 min with occasional flicking of the tube to maintain the beads in suspension. Beads were then magnetized, and the supernatant was transferred to a new Lo-Bind microcentrifuge tube. Biotin, H2-K<sup>b</sup> heavy chain and B2M light chain were then precipitated by adding 1% trifluoroacetic acid slowly while gently vortexing the elution. The fluffy white precipitate was then pelleted by centrifugation at 20,000g for 10 min. Supernatants containing liberated MHC-I peptides were then directly aspirated into 8 µg binding capacity C18 solid-phase extraction pipette tips (Pierce). Tips were preequilibrated with 50% acetonitrile (ACN) and 0.1% formic acid (FA) according to the manufacturer's instructions. The 400 µl of eluted material was loaded, 20 µl at a time, until the entire volume of the elution was passed over the C18 sorbent. Tips were then washed twice with 5% ACN and 0.1% FA before elution in 10 µl of 30% ACN and 0.1% FA. Desalting peptides were then dried down before reconstitution in 3% ACN and 0.1% FA in an autosampler vial for liquid chromatography and tandem mass spectrometry (LC–MS/MS) analysis.

#### Antibody immunoprecipitation of MHC-I

Peptide MHC isolation was performed as previously described<sup>33</sup>. Healthy lung tissue or tumour-bearing lung tissue was homogenized

# Article

and cleared as described for 'Affinity purification of H2-K<sup>b</sup> with Strep-Tactin'. Per sample, 1 mg of anti-H2-K<sup>b</sup> (clone Y3, BioXCell) was bound to 20 µl (bed volume) FastFlow Protein A sepharose beads (GE Healthcare) by incubating for 1 h at 4 °C. Beads were then washed with lysis buffer and samples were incubated for 2–4 h with rotation at 4 °C. Beads were then centrifuged at 1,000 rpm, washed twice with MEB, twice with 1× TBS and eluted with 10% acetic acid at room temperature. Eluate was then filtered using 10 kDa MWCO spin filters (PALL Life Science), which were passivated with 0.1% BSA and acidified with 10% acetic acid before filtration. Filtered peptides were then further purified with 8 µg binding capacity C18 tips (Pierce) before LC–MS/MS analysis. For multiplexing, lyophilized pMHCs were resuspended in 33 µl of labelling buffer (50% ethanol, 150 mM TEAB) and mixed with 40 µg of pre-aliquoted TMT 6plex (Thermo Scientific) resuspended in 10 µl of anhydrous ACN. The labelling reaction was carried out on a shaker for 1 h at room temperature and quenched with 0.3% hydroxylamine. Samples were combined and dried in SpeedVac before being cleaned up using the SP3 protocol as previously described<sup>15</sup>.

## LC–MS/MS

pMHC samples were analysed using an Exploris 480 Hybrid Quadrupole–Orbitrap mass spectrometer (Thermo Scientific) coupled with an UltiMate 3000 RSLC Nano LC system (Dionex), Nanospray Flex ion source (Thermo Scientific) and column oven heater (SonaT). Samples were resuspended in 0.1% FA and directly loaded onto an analytical capillary chromatography column with an integrated electrospray tip (about 1 µm orifice), prepared and packed in-house (50 µm ID × 15 cm and 1.9 µm C18 beads, ReproSil-Pur). All or half the pMHC elutions were analysed in a given LC–MS/MS analysis. Peptides were eluted using a linear gradient with 6–25% buffer B (70% ACN, 0.1% FA) for 53 min, 25–45% for 12 min, 45–97% for 3 min, hold for 1 min and 97% to 3% for 1 min.

Standard MS parameters were as follows: spray voltage, 2.0–2.5 kV; no sheath or auxiliary gas flow; and heated capillary temperature, 275 °C. The Exploris was operated in data-dependent acquisition (DDA) mode. Full scan mass spectra (350–1,200 *m/z*, 60,000 resolution) were detected in the orbitrap analyser after accumulation of  $3 \times 10^6$  ions (normalized automatic gain control (AGC) target of 300%), automatic maximum injection time (IT). For every full scan, MS<sup>2</sup> spectra were collected during a 3 s cycle time. Ions were isolated (0.4 *m/z* isolation width) for a maximum IT of 150/250 ms or  $1 \times 10^5$  to  $7.5 \times 10^4$  ions (100%/75% normalized AGC target) and fragmented by higher energy disassociation collision (HCD) with 30% collision energy (CI) at a resolution of 60,000. Charge states <2 and >4 were excluded, and precursors were excluded from selection for 30 s if fragmented  $n = 2$  times within 20 s window.

Synthetic peptides were ordered from Genscript at ≥85% purity, with no trifluoracetate removal, and stock solutions were dissolved in DMSO at a peptide concentration of 10 mM. Synthetic peptides analysis was also performed on the Exploris 480 coupled to an Agilent 1260 LC system using a custom analytical chromatography column, prepared as described above. Peptides were eluted using a 105-min gradient (3.4% for 10 min, 3.4–6.9% for 2 min, 6.9–29% for 53 min, 51–100% for 3 min, hold for 1 min, 100–3.4% for 1 min) with 70% ACN in 0.2 M acetic acid at the flow rate of 0.2 ml min<sup>-1</sup> and a pre-column split of 2,000:1. The Exploris was operated with a precursor scan range of 440–544 *m/z* and a targeted inclusion list for synthetic peptides (Supplementary Table 10). Additional acquisition parameters are as follows: full scan spectra were collected at 60,000 resolution, 300% normalized AGC, automatic IT; MS<sup>2</sup> spectra were collected at 60,000 resolution, isolation width 0.4 *m/z*, maximum IT of 250 ms, 100% normalized AGC target fragmented by HCD with 30% CI with a 3 s cycle time.

TMT-labelled samples were analysed on a QExactive Plus mass spectrometer coupled to Agilent 1260 LC. Samples were resuspended in 3% ACN, 0.1% FA and one-third of the labelled mixture was loaded onto a

pre-column (100 µm ID × 10 cm) packed in-house with 10 µm C18 beads (YMC gel, ODS-A, AA12S11) connected in tandem to an in-house packed analytical column (50 µm ID × 15 cm and 5 µm C18 beads). Peptides were separated by a 105-min LC gradient with a flow rate of 0.2 ml min<sup>-1</sup> and a similar pre-column split ratio. MS<sup>1</sup> scans were performed with the following settings: 350–1,200 *m/z* range, resolution of 70,000, AGC target of  $3 \times 10^6$  and maximum IT of 50 ms. The top 15 abundance ions were isolated and fragmented with HCD (31% CI) with the following parameters: resolution of 70,000, AGC target of  $1 \times 10^5$ , maximum IT of 350 ms and isolation window of 0.4 *m/z*. Charge states <2 and >4 were excluded. TMT channels were as follows: 126, KP; 127, KP 128; KP/KbStrep; 129, KP/KbStrep.

## MS data analysis

All mass spectra were analysed with Proteome Discoverer (v.2.5) and searched using Mascot (v.2.4) against the mouse SwissProt database (2021\_03, 2021\_02 for label-free quantification). Peptides were searched with no enzyme and variable methionine oxidation. Peptide spectrum matches were further filtered according to the following criteria: ions score ≥ 15, search engine rank = 1 and results from technical replicates of each sample analysis were combined. Median retention time was calculated using the retention time values of filtered peptide spectrum matches from all replicates of a given sample. Label-free quantitation was done using the Minora Feature Detector (precursor abundance values measured based on area under the curve) in Proteome Discoverer with match between runs enabled and filtered for peptides with ions score ≥ 15 and search engine rank = 1. Abundances were averaged across technical and biological replicates. All data were processed in R studio (R v.4.1.0) and Microsoft Excel (v.16.57).

Endogenous and synthetic peptide spectra were compared by head-to-tail plots using MSnbase R package<sup>52</sup>. Pearson's correlations and dot products were calculated using all the detected fragment ions in both spectra. Spectra that had the highest ions score in tumour samples (endogenous) or synthetic peptide sample were used (details listed in Supplementary Table 10) for head-to-tail plots and visualization using Interactive Peptide Spectral Annotator<sup>53</sup>. For Interactive Peptide Spectral Annotator, all fragment ions were plotted.

## Immunofluorescence of lung and pancreas tissue

Tumour-bearing lung or pancreas tissue was manually dissected and embedded in optimal cutting temperature (OCT) compound and slowly frozen on dry ice. Frozen tissue sections were stored at -80 °C until sectioning. On the day of sectioning, frozen tissue was allowed to equilibrate to -20 °C in the cryostat for at least 1 h. Sections (8 µm) were then cut and transferred to microscope slides (Fisher) before fixation in 100% acetone at -20 °C for 10 min and dried for 15 min at room temperature. Slides were then stored at -20 °C until staining.

Tissue sections were circled with a hydrophobic pen and then rehydrated with PBS for 5–10 min and then blocked in PBS supplemented with 5% BSA for 45 min. Primary antibodies were then added at the indicated dilutions (Supplementary Table 11) for 1 h at 25 °C. Slides were then washed four times with PBS and BSA, and incubated with fluorescently labelled secondary antibodies where indicated for 1 h at 25 °C. Stained sections were then washed three times with PBS, incubated with DAPI for 5 min, washed once with PBS and then mounted with Prolong Diamond AntiFade Mountant (Thermo). Slides were scanned on a VERSA 8 slide scanner (Leica) before analysis in ImageScope v64 and ImageJ.

## Immunoblotting

For organoids, cells were dissociated with TrypLE, washed once with PBS, and then lysed in cell lysis buffer (CLB; 50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% SDS and 1× HALT protease inhibitors). For monolayer cultures, medium was aspirated and cells were washed twice with PBS before lysis with CLB. Protein concentration

was quantified with BCA (Pierce) and samples were loaded onto 4–12% Bis-Tris SDS-PAGE gels (Invitrogen) and electrophoresed at 150 V until the loading dye reached the bottom of the gel. Gels were then transferred onto 0.45 µm nitrocellulose membranes overnight at 20 V in a cold room. Blots were blocked with PBST (PBS plus 0.5% Tween20) and 5% milk for 30 min at room temperature, incubated with primary antibody for 1 h at room temperature, washed four times in PBST, incubated with HRP-conjugated secondary antibody for 1 h at room temperature, and then washed four times in PBST. Blots were developed with Clarity or Clarity-Max ECL substrate (Bio-Rad) and imaged on a ChemiDoc Gel Imaging System (Bio-Rad).

### PCR with reverse transcription

RNA was isolated from organoid cultures by aspirating medium and then resuspending the Matrigel domes in 1 ml TRIzol (Ambion). RNA was then extracted from the homogenate using a PureLink RNA Mini kit (Ambion) according to the manufacturer's instructions with the final elution step in 40 µl of nuclease-free water. Isolated RNA was then quantified on a NanoDrop 2000 (Thermo). Total RNA (1 µg) was then added to a cDNA synthesis reaction using a High Capacity cDNA Reverse Transcription kit (Applied Biosciences) according to the manufacturer's instructions, with random hexamers as primers and including RNase inhibitor in the reaction. cDNA was then diluted 1:10 with nuclease-free water before adding 1 µl to a PCR reaction containing 500 µM forward and reverse primer and 1× Q5 master mix (NEB). Reactions were cycled 35 times with annealing temperatures calculated by NEB  $T_m$  calculator before loading on 2% agarose gels. DNA bands were imaged by staining with ethidium bromide and ultraviolet transillumination.

### Ribosome-protected fragment isolation with RiboLace

Ribosome profiling was performed using a previously described method using puromycin-conjugated magnetic beads to isolate actively translating ribosomes<sup>28</sup>. To obtain cellular material for Ribo-seq, organoid cultures of normal AT2 cells or organoids derived from 12-week KP tumours were transiently adapted to monolayer culture by washing tissue cultures dishes with a 10% Matrigel solution before plating cells in either complete lung organoid medium for AT2 cells or lung organoid base medium for tumour organoids<sup>30</sup>. After 48 h, cells were treated with 100 µM cycloheximide for 5 min and immediately transferred to ice. Plates were then washed once with 10 ml PBS and 100 µM cycloheximide followed by direct lysis and thorough mechanical disruption in the plate according to the RiboLace Protocol (Immagina Biotech). Lysates were then cleared by centrifugation at 20,000g for 15 min. Nucleic acid content was quantified with a NanoDrop and each lysate was normalized to 0.3 a.u.<sub>260 nm</sub> in 150 µl of W-buffer. SS solution (0.3 µl) and diluted RNase Nux solution (5 µl of 1:66.7 dilution) were then added, and RNase digestion was stopped with 0.5 µl SUPERaseIn for 10 min on ice.

RiboLace beads were prepared according to RiboLace kit instructions. In brief, per sample, 90 µl of beads was magnetized, washed once with buffer OH, once with nuclease-free water, twice with B buffer and then resuspended in 30 µl RiboLace probe followed by incubation shaking for 1 h at room temperature. Beads were then passivated with 3 µl PEG for 15 min at room temperature and washed twice with 500 µl of buffer W.

Prepared beads were then resuspended in RNase digested lysates and incubated with slow rotation for 70 min at 4 °C. Beads were then washed twice with 500 µl of buffer W before protease digestion by addition of 20 µl of SDS and 5 µl of proteinase K and incubating for 75 min at 37 °C. RNA was then extracted with phenol-chloroform, precipitated overnight with isopropanol and then electrophoresed on denaturing 15% TBE-urea gels. Gel pieces were excised from the region of 20–40 nucleotides to enrich for ribosome-protected fragments (RPFs). Gel pieces were then crushed by centrifugation through a punctured 0.5-ml tube placed in a 1.5 ml microcentrifuge tube. Gel debris was then

incubated overnight rotating in 400 µl of gel elution buffer (20 mM Tris pH 7.5, 250 mM sodium acetate, 1 mM EDTA and 0.25% SDS). RNA was precipitated from the elutions by adding 700 µl of isopropanol and 1.5 µl of GlycoBlue (Thermo) followed by overnight incubation at –80 °C. RNA was then pelleted by centrifugation at 20,000g for 30 min at 4 °C. The RNA pellet was then washed with 70% ethanol, air dried for 2–3 min and resuspended in 10 µl of nuclease-free water.

Dephosphorylation of RPFs was performed by adding 23 µl nuclease-free water, 5 µl T4 PNK buffer (NEB), 1 µl SUPER In (Invitrogen) and 1 µl T4 PNK (NEB) followed by incubation for 60 min at 37 °C, then 10 min at 70 °C, and cooled to room temperature. RNA was then precipitated by adding 39 µl nuclease-free water, 10 µl 3 M sodium acetate, 150 µl isopropanol and 1 µl GlycoBlue and incubated at –80 °C overnight. RNA was pelleted by centrifugation at 20,000g for 30 min, washed with 70% ethanol and then resuspended in 7 µl of nuclease-free water.

### RPF sequencing library preparation

Illumina sequencing libraries of isolated RPFs was performed using a LACEseq kit (Immagina Biotechnology). Library preparation was performed according to the manufacturer's protocol with the following parameters. RPFs (5 ng) were 5' phosphorylated in a 50 µl reaction containing 5 µl buffer BPK, 5 µl ATP, 1 µl PK and 34 µl nuclease-free water. The reaction was incubated for 1 h at 37 °C. RNA was then purified with an RNA Clean and Concentrator - 5 kit (Zymo) and eluted in 6 µl. A linker was then ligated with 6 µl of RNA from the previous step, 1 µl buffer BA, 0.5 µl GTP, 0.6 µl MnCl<sub>2</sub>, 1 µl enzyme mix A, 0.25 µl linker MC and 0.75 µl nuclease-free water. The reaction was then incubated for 1 h at 37 °C and purified with an RNA Clean and Concentrator kit with 8 µl elution in nuclease-free water. Circularization was then performed with 8 µl of the eluted RNA, 2 µl buffer BLB, 1 µl ATP, 8 µl PEG8000 and 1 µl enzyme mix B. The resulting reaction was then incubated for 2 h at 25 °C. The reaction was then purified with a RNA Clean and Concentrator kit with 10 µl elution in nuclease-free water. The 10 µl of circularized RNAs were then primed by adding 1 µl of dNTPs, 1 µl of RTT primer and 2 µl of nuclease-free water and incubating at 70 °C for 5 min. Next, a master mix containing 4 µl buffer BRT, 1 µl DTT and 1 µl RT enzyme was added followed by an incubation at 50 °C for 40 min and heat inactivation for 5 min at 80 °C. cDNA was then amplified by adding 50 µl amplification mix, 0.8 µl forward primer, 0.8 µl reverse primer and 28.4 µl nuclease-free water. The reactions were then cycled seven times according to the manufacturer's PCR parameters. PCRs were cleaned up using AMPURE XP beads at a 1.6× ratio and elution in 40 µl of nuclease-free water. A second PCR amplification was performed using 20 µl from the previous PCR, 50 µl amplification mix, 1 µl LACEseq UDIs and 29 µl nuclease-free water. Reactions were cycles six times according for manufacturer's parameters. The final libraries were size selected using PAGE purification and cleaned up using a DNA Clean and Concentrator (Zymo) before Illumina sequencing.

### RNA-seq library preparation

RNA was extracted from the remaining AT2 or tumour lysates that were used for Ribo-seq using TRIzol according to the manufacturer's instructions. The aqueous phase containing RNA was precipitated with 2 volumes of isopropanol and washed once with 70% ethanol and air dried. Pellets were resuspended in nuclease-free water before polyA+ selective mRNA library preparation at the MIT BioMicroCenter.

### Ribo-seq data analysis

Raw fastq files from each Ribo-seq library were trimmed according to recommendations in the LACE-seq protocol (Immagina Biosciences). Fastq files from each sample type were then merged before submission to the RiboToolKit server (<http://rnabioinfor.tch.harvard.edu/RiboToolKit/>) using a 26–38 nt size filter and a predicted open reading frame cut-off of  $P < 0.05$ . Data were then downloaded and visualized in

# Article

R. For analysis of TE, raw RNA-seq and Ribo-seq data were aligned to the transcriptome with kallisto 0.46.0. Aligned reads were then imported into R for DEseq2 analysis using the tximport package. Differential expression at the RNA and RPF levels were calculated using standard DEseq2 workflows. For TE analysis, the DEseq2 results object was generated incorporating both sequence type (Ribo versus RNA) and sample (LUAD versus AT2) as independent variables to identify genes that were differentially translated in LUAD versus AT2.

## HSP90i treatment

The HSP90i (NVP-HSP990, Selleck Chemical) was administered in the drinking water<sup>32</sup>. The average water consumption of the mice was calculated by measuring water bottle weight before and after 72 h of housing to determine the average consumption per mouse per day. Across all experiments, C57Cl/6 mice consumed approximately 4 ml every day. Using these water consumption values and mouse weight, a 4 mg ml<sup>-1</sup> stock solution of NVP-HSP990 (in 100% PEG400) was diluted directly into the drinking water to achieve a target dose of 0.5 mg kg<sup>-1</sup> day<sup>-1</sup>. HSP90i treatment began 8 weeks after tumour induction, and water was replaced twice per week for 4 weeks before euthanasia.

## CD8 depletion and CD40/FLT3L treatment

KP/K<sup>b</sup>Strep tumour-bearing mice at 2 weeks (chronic depletion) or 14 weeks (acute depletion) after tumour initiation were treated with 100 µg of anti-CD8a-depleting antibody (BioXCell, BE0061) every 3 or 4 days until animals reached 16 weeks after tumour initiation. At necropsy, a small portion of tumour-bearing lung tissue (about 100 mg) was taken for flow cytometry confirmation of CD8a<sup>+</sup> immune cell depletion. The remaining tissue was flash frozen and processed for pMHC isolation.

KP/K<sup>b</sup>Strep tumour-bearing mice at 15 weeks after tumour initiation were treated for 7 days with 10 µg FLT3L-Ig (BioXCell, BE0342) with intraperitoneal injection<sup>26</sup>. On day three of FLT3L-Ig treatment, mice were also dosed with 100 µg of agonistic CD40 antibody (BioXCell, BP0016-2) by intraperitoneal injection. On day 7 of treatment (16 weeks after tumour initiation), mice were euthanized and a small portion of tumour-bearing lung tissue (about 100 mg) was taken for flow cytometry analysis of tumour-infiltrating CD8<sup>+</sup> T cells, and the remaining tissue was flash frozen for pMHC isolation.

## OT-I killing assay

OT-I transgenic mice on a *Rag2*<sup>-/-</sup> background were euthanized by CO<sub>2</sub> inhalation and spleens were dissected and macerated through a 70-µm cell strainer. Splenocytes were then pelleted by centrifugation at 500g for 5 min, followed by red blood cell lysis using ACK lysis buffer for 2 min at room temperature. ACK was quenched with RPMI 1640 and 10% HI FBS, and splenocytes were pelleted by centrifugation at 500g for 5 min. The resulting cell pellet was then resuspended in 10 ml of T cell medium (RPMI 1640 supplemented with 10% HI FBS, 1× penicillin-streptomycin (Gibco), 1× non-essential amino acids (Gibco), 1× L-glutamine (Gibco), 1× HEPES (Gibco) and recombinant IL-2 (Peprotech)). At the time of plating, 1 µM SIINFEKL peptide was added to stimulate T cell activation and differentiation. At 18 h after plating cells, the SIINFEKL peptide was removed and cells were replated in T cell medium. At 72 h after addition of SIINFEKL, cells were prepped for target killing assays.

Target cells were KP or KP/K<sup>b</sup>Strep cell lines transduced with either lentivirus encoding mScarlet or mScarlet-SIINFEKL. On the day of the assay, target cells were trypsinized, washed and counted. KP-mScarlet cells ( $1 \times 10^6$ ) were incubated with 100 nM CFSE (KP-CFSE<sup>lo</sup>) and  $1 \times 10^6$  KP-mScarlet-SIINFEKL cells were incubated with 5 µM CFSE (KP-CFSE<sup>hi</sup>) for 20 min at room temperature in PBS<sup>54</sup>. In parallel,  $1 \times 10^6$  KP/K<sup>b</sup>Strep-mScarlet cells were incubated with 100 nM CFSE (KP/K<sup>b</sup>Strep CFSE<sup>lo</sup>), and  $1 \times 10^6$  KP/K<sup>b</sup>Strep mScarlet-SIINFEKL cells were incubated with 5 µM CFSE (KP/K<sup>b</sup>Strep CFSE<sup>hi</sup>) for 20 min at room temperature in PBS. Labelling reactions were quenched with DMEM and 10% FBS and

cells pelleted at 500g for 5 min. KP CFSE<sup>hi</sup> or CFSE<sup>lo</sup> and KP/K<sup>b</sup>Strep CFSE<sup>hi</sup> or CFSE<sup>lo</sup> were mixed in a 1:1 ratio and plated into round-bottom 96-well plates at 20,000 cells per well (10,000 CFSE<sup>hi</sup> and 10,000 CFSE<sup>lo</sup>). Transgenic OT-I cells were then counted, washed and plated into wells containing target cells at 0:1, 5:1 and 1:1 effector:target ratios. After 4 h, cells were processed for flow cytometry. Cells were gated on live/dead viability stain, and CD45<sup>+</sup> cells were excluded. Specific killing was calculated with the following formula:  $100 - ((CFSE^{hi}/CFSE^{lo} - \text{with effector cells})/(CFSE^{hi}/CFSE^{lo} \text{ without effector cells})) \times 100$ .

## Bone-marrow-derived dendritic cell vaccination

Peptides were prioritized for vaccination as described in Fig. 4a. Candidates were then further prioritized through manual inspection of RNA expression patterns across healthy mouse tissues. We selected peptides that have very low RNA expression in healthy tissue (*Slc26a4* and *Prdm15*), peptides that resemble cancer testis antigens (*Ift74*, *Ccd158* and *Arrdc5*), oncofetal antigens (*Gpn3* and *Znf462*), and widespread expression despite tumour-specific presentation (*Csf2ra*). This effort was not comprehensive of all potential antigens, and continued interrogation of peptides from our model is warranted.

Isolation of bone-marrow-derived dendritic cells (BMDCs) was carried out by euthanizing mice and dissecting the femur and tibia into PBS<sup>33</sup>. The ends of the bones were then clipped, and the open end of each bone was placed facing down into a 0.5-ml tube with a hole in the bottom and placed into a 1.5 ml tube. The tube assembly was then quickly centrifuged on a tabletop microcentrifuge. Red blood cells were removed from the pelleted cell mass with ACK lysis buffer (Gibco), and the resulting cell pellet was resuspended at  $1.5 \times 10^6$  cells ml<sup>-1</sup> in BMDC medium (RPMI 1640 supplemented with 10% FBS, penicillin-streptomycin, 50 µM β-mercaptoethanol, 600 ng ml<sup>-1</sup> FLT3L (BioXCell) and 5 ng ml<sup>-1</sup> GM-CSF (R&D Systems)) and cultured in non-tissue culture treated dishes. After 5 days, cells were washed, recounted and replated in BMDC medium. On day 10, cells were activated with 10 µg CpG ODN 1826 (Abeomics) for 18–24 h.

On day 11, activated BMDCs were gently washed with BMDC medium and pulsed with 10 µg of each peptide in a 15-ml conical tube with the lid not fully tightened for 2 h in a 37 °C tissue culture incubator. Pulsed cells were then washed once with BMDC medium, and two times with PBS. For injection,  $1.5 \times 10^6$  peptide-loaded BMDCs were subcutaneously injected into the flank. As a negative control, mice received BMDCs that were not pulsed with peptide.

Ten days after BMDC injection, mice received a booster of 6 µg of each peptide with 25 µg of c-di-GMP as adjuvant. At 14 days after the first boost (24 days after BMDC injection), mice received a booster of 6 µg of each peptide with 1 nmol of LipoCpG as adjuvant (provided from the Irvine Laboratory at MIT). Seven days after the second boost, splenocytes were collected for ELISPOT analysis or tetramer staining for tumour bearing animals.

## Custom pMHC tetramer generation

H2-K<sup>b</sup> tetramers were generated in-house using previously described protocols<sup>34</sup>. Specifically, we used disulfide bridged H2-K<sup>b</sup> (Y84C, A139C). In brief, recombinant MHC-I heavy chain (DS-K<sup>b</sup>) and light chain (*Homo sapiens* B2M) were expressed in *Escherichia coli* using the lac-operon regulated pET-15b plasmid. Inclusion bodies were prepped as previously described<sup>55</sup>. Inclusion bodies were solubilized in 8 M urea, 100 mM Tris-HCl pH 8.0 and added to a refolding buffer (100 mM Tris-HCl pH 8.0, 400 mM L-arginine (Sigma), 5 mM reduced glutathione (Sigma), 0.5 mM oxidized glutathione (Sigma), 2 mM EDTA (Gibco), 1× HALT protease inhibitor (Roche)). At the time of refolding, 1 mM Gly-Leu dipeptide was added to refolding buffer, followed by 2 µM B2M, 1 µM DS-K<sup>b</sup> and 1 mM PMSF for three consecutive days. After 72 h, the reaction was concentrated to about 2 ml by applying the reaction to a 10,000 MWCO filter by nitrogen flow. The resulting concentrated protein solution was then purified by gel-filtration on an s200 Sephadex column

in HBS (100 mM HEPES pH 8.0, 150 mM NaCl). Relevant fractions were collected and concentrated to approximately 8 ml. Purified monomers were then biotinylated using BirA Ligase according to manufacturer's instruction (Avidity). Biotinylated monomers were then further purified with another round of gel-filtration on an s200 Sephadex in HBS, and the relevant fractions were collected and concentrated to 2 mg ml<sup>-1</sup> and 50 µl aliquots were flash frozen in liquid nitrogen until use.

To generate tetramers, 15.9 µl of Streptavidin-PE (SA-PE, Molecular Probes) was added to 50 µl DS-K<sup>b</sup> monomers and incubated for 10 min at room temperature, protected from light. Addition of SA-PE was repeated 9 times for a total of 10 times resulting in a final volume of 209 µl tetramerized DS-K<sup>b</sup>. At the time of staining, 20 µl of tetramer was incubated with 50 µM peptide of interest for 1 h on ice. The resulting solution was used at a 1:200 dilution in flow cytometry staining solutions.

### ELISPOT assay

Before plating cells, wells of a precoated ELISPOT plate were loaded with 100 litre of DMEM and 10% FBS and penicillin-streptomycin containing 10 µM of individual peptides, medium only or PMA/ionomycin as a positive control. Spleens from control or vaccinated mice were dissected and macerated through a 70-µm filter. Cells were then centrifuged for 5 min at 1,500 r.p.m. and pellets were then lysed with ACK lysis buffer and the remaining splenocytes were counted with a haemocytometer. Cells were then plated across control and peptide stimulation conditions at a concentration of 750,000 cells per well and gently mixed before incubation for 20 h. After overnight incubation, cells were decanted and the plates were washed and IFN $\gamma$  spots detected using the procedure outlined in the Mouse IFN $\gamma$  ELISPOT kit (ImmunoSpot). After development, plates were dried overnight before automated spot counting on an ImmunoSpot scanner.

### Data analysis, statistics and visualization

All quantitative data were processed and visualized in R (R v.4.0.2, RStudio v.1.3.959). Peptides from PDAC or LUAD samples were concatenated and organized in Microsoft Excel with predicted affinity values calculated using NetMHCpan 4.1. Following import into R, peptide lists were filtered for length (8–11 amino acids) and predicted affinity (<1,000 nM) before further analysis. Non-metric multidimensional scaling (NMDS) analysis was performed using a previously published R script<sup>14</sup> and applied to 8-mer and 9-mer peptides separately. Peptide clusters were empirically chosen through hierarchical clustering of the calculated NMDS matrix. Peptide motifs were generated with the ggseqlogo package in R. All boxplots are in the Tukey style with the horizontal mark at the median, the box extending from the first quartile to the third quartile, and whiskers extending 1.5 $\times$  the interquartile range from the box.

Bulk RNA-seq was adapted from ref.<sup>27</sup> by taking the mean expression of each gene across all normal, early, non-metastatic and metastatic primary tumour samples before cross-comparison with peptide data.

scRNA-seq data were analysed using Seurat (v.4.0.1). Cells by genes matrices were downloaded from the Tabula Muris (healthy lung and pancreas) or from the Broad Institute's Single Cell Portal (KP scRNA-seq data). Data were filtered with a gene count cut-off of 1,100 before normalization, scaling and dimensionality reduction with uniform manifold approximation and projection according to standard Seurat pipelines (<https://satijalab.org/seurat/>). Custom modules were added with the AddModuleScore function according to gene lists from peptide data or published reports. Volcano plots for comparing signatures across cell types in the healthy lung data were calculated with two-tailed Student's *t*-test for all pairwise comparisons of Normal, Ab and Strep signatures within each cell type with a Bonferroni multiple comparisons adjustment. The median score for each signature was then used to calculate the log<sub>2</sub>(fold change) and *P* values were transformed to -log<sub>10</sub>(*P* values) for plotting. Pearson's correlations between gene

signatures were calculated using AddModuleScore in Seurat, extraction of metadata, conversion of the data to a matrix containing cells  $\times$  signatures and calculation of Pearson's correlation for signatures across cells. For correlation of peptide signatures to all genes, the scaled data from Seurat was extracted to a matrix, and the late-stage or Tumour-unique peptide signature was appended onto the matrix. Pearson's correlation was then calculated between the peptide signature versus each individual gene. Genes were then ranked by correlation and analysed with pre-ranked gene set enrichment in GSEA (Broad Institute).

Gene ontology analysis was performed on Gene Lists in StringDb (<http://www.stringdb.org>) with high confidence interaction cut-offs. Gene set enrichment analysis (Fig. 2) was performed with the GSEA software package using a pre-ranked gene list according to correlation to the late-stage or Tumour-unique peptide signature.

Subcellular compartment analysis was performed by extracting compartment locations from UniProt for source proteins that had available localization data. Comparison of subcellular distributions was performed with Fisher's exact test with Monte Carlo simulation. Protein length information was also extracted from UniProt for each source protein. Thermal stability and protein half-life were obtained from published datasets<sup>30</sup>.

Statistical analyses were performed in R. To compare median fluorescence intensities from flow cytometry data, two-tailed Student's *t*-tests were performed for all pairwise comparisons. For comparisons of gene expression or signature scores, unpaired two-sample Wilcoxon test (Mann-Whitney *U*) was performed. For analysis of peptide abundance distributions in vehicle-treated and HSP90i-treated samples, Kolmogorov-Smirnov tests were used.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

### Data availability

All MS data have been deposited to the Proteomics Identifications Database (PRIDE) repository with the dataset identifier PXD033232. Raw RNA-seq and Ribo-seq data have been submitted to the Gene Expression Omnibus with the dataset identifier GEO178944. Source data are provided with this paper.

49. Marino, S., Vooijs, M., Gulden, H. V. D., Jonkers, J. & Berns, A. Induction of medulloblastomas in p53-null mutant mice by somatic inactivation of *Rb* in the external granular layer cells of the cerebellum. *Genes Dev.* **14**, 994–1004 (2000).
50. Jackson, E. L. et al. Analysis of lung tumor initiation and progression using conditional expression of oncogenic *K-ras*. *Genes Dev.* **15**, 3243–3248 (2001).
51. VanDussen, K. L., Sonnek, N. M. & Stappenbeck, T. S. L-WRN conditioned medium for gastrointestinal epithelial stem cell culture shows replicable batch-to-batch activity levels across multiple research teams. *Stem Cell Res.* **37**, 101430 (2019).
52. Gatto, L., Gibb, S. & Rainer, J. MSnbase, efficient and elegant R-based processing and visualization of raw mass spectrometry data. *J. Proteome Res.* **20**, 1063–1069 (2021).
53. Brademan, D. R., Riley, N. M., Kwiecien, N. W. & Coon, J. J. Interactive peptide spectral annotator: a versatile web-based tool for proteomic applications. *Mol. Cell. Proteomics* **18**, S193–S201 (2019).
54. Noto, A., Ngauv, P. & Trautmann, L. Cell-based flow cytometry assay to measure cytotoxic activity. *J. Vis. Exp.* <https://doi.org/10.3791/51105> (2013).
55. Grotenbreg, G. M. et al. Discovery of CD8<sup>+</sup> T cell epitopes in *Chlamydia trachomatis* infection through use of caged class I MHC tetramers. *Proc. Natl Acad. Sci. USA* **105**, 3831–3836 (2008).

**Acknowledgements** We thank members of the Jacks Lab for critical comments on the work presented here, especially Z. Ely and N. Pattada for conceptual and technical support; G. Eng for support with organoid work; T. Koller and R. Schiavoni of the Koch Institute Proteomics core for technical insight on proteomic workflows; and staff at the Koch Institute Swanson Biotechnology Center for technical support, specifically the Flow Cytometry, Histology, Preclinical Modeling, Imaging and Testing, and Integrative Genomics and Bioinformatics core facilities. Data accessibility from the Tabula Muris was essential for multiple analyses in this manuscript. This work was supported by the Damon Runyon Cancer Research Foundation (A.M.J.), an NCI K99 Pathway to Independence Award (A.M.J.), the Howard Hughes Medical Institute, the Johnson & Johnson Lung Cancer Initiative, NCI Cancer Center Support Grant P30-CA1405, the Lustgarten Foundation Pancreatic Cancer Research Laboratory at MIT, the Stand Up To Cancer-Lustgarten Foundation Pancreatic Cancer Interception Translational

# Article

Cancer Research Grant (grant number SU2C-AACR-DT25-17, W.A.F.-P. and T.J.), the MIT Center for Precision Cancer Medicine (L.E.S., R.A. and F.M.W.), the Margaret A. Cunningham Immune Mechanisms of Cancer Research Fellowship (L.E.S.), the Melanoma Research Alliance (L.E.S.), NIH Training Grant (T32-ES007020), the Ludwig Center at MIT (R.A.), the Pew-Stewart Scholars Program for Cancer Research (A.K.S.), and the Cancer Research Institute (T.F. and S.S.).

**Author contributions** A.M.J. and T.J. conceived, designed and directed the study. L.E.S., R.A. and F.M.W. directed all MS analyses. A.M.J., L.E.S., R.A., E.A.S., D.A.S. and R.E.K. performed all experiments. W.A.F.-P. conducted all orthotopic and autochthonous pancreatic surgeries. S.N. provided guidance and reagents for AT2 organoid culture. W.M.R. conducted mouse ESC targeting and chimera generation. A.M.J., L.E.S., R.A. and T.F. performed data analyses. P.M.K.W. assisted with custom pMHC tetramer generation. K.B.N. and S.S. provided assistance with vaccination and ELISPOT analysis. P.S.W. and A.K.S. provided guidance for analysis of the scRNA-seq data. J.S. and S.-L.S. provided technical and conceptual support of the study. A.M.J., L.E.S., R.A., F.M.W. and T.J. wrote the manuscript with input from all authors.

**Competing interests** T.J. is a member of the Board of Directors of Amgen and Thermo Fisher Scientific, and a co-Founder of Dragonfly Therapeutics and T2 Biosystems. T.J. serves on the Scientific Advisory Board of Dragonfly Therapeutics, SQZ Biotech and Skyhawk Therapeutics.

T.J. is also President of Break Through Cancer. His laboratory currently receives funding from Johnson & Johnson and The Lustgarten Foundation, and funds from the Lustgarten Foundation supported the research described in this manuscript. S.S. is a co-founder of Danger Bio and serves on the Scientific Advisory Board of Related Sciences/DanderBio, Ankyra Therapeutics, Arcus Biosciences, Takeda and Ribon, and serves as an advisor for Dragonfly Therapeutics and Merck. The S.S. Lab currently receives funding from Leap Therapeutics. A.K.S. reports compensation for consulting and/or SAB membership from Merck, Honeycomb Biotechnologies, Cellarity, Repertoire Immune Medicines, Ochre Bio, Third Rock Ventures, Hovione, Relation Therapeutics, FL82, Empress Therapeutics and Dahlia Biosciences. None of these affiliations influenced the work conducted or analysis of data presented in this manuscript.

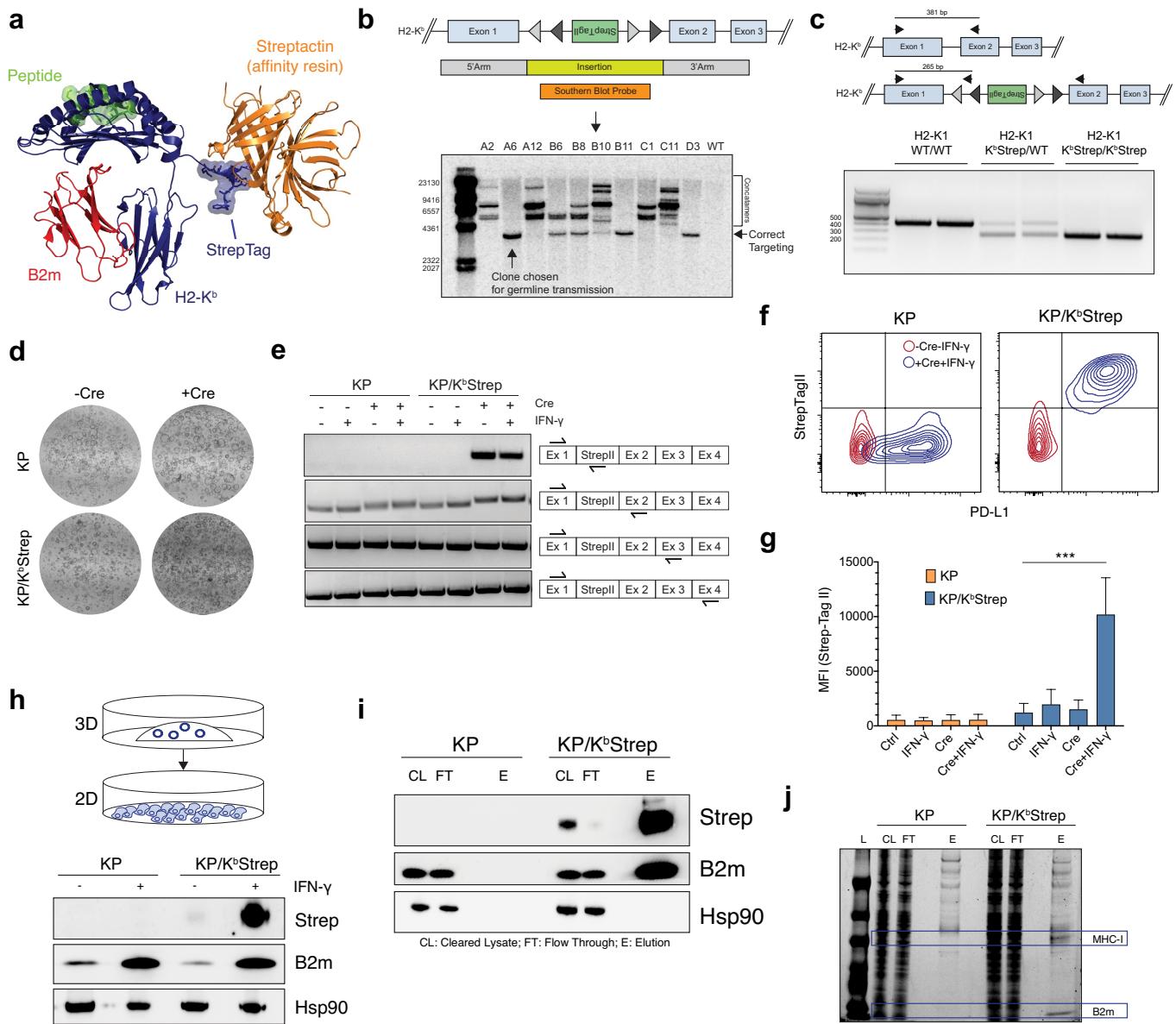
## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-022-04839-2>.

**Correspondence and requests for materials** should be addressed to Tyler Jacks.

**Peer review information** Peer review statement: *Nature* thanks Lélia Delamarre and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

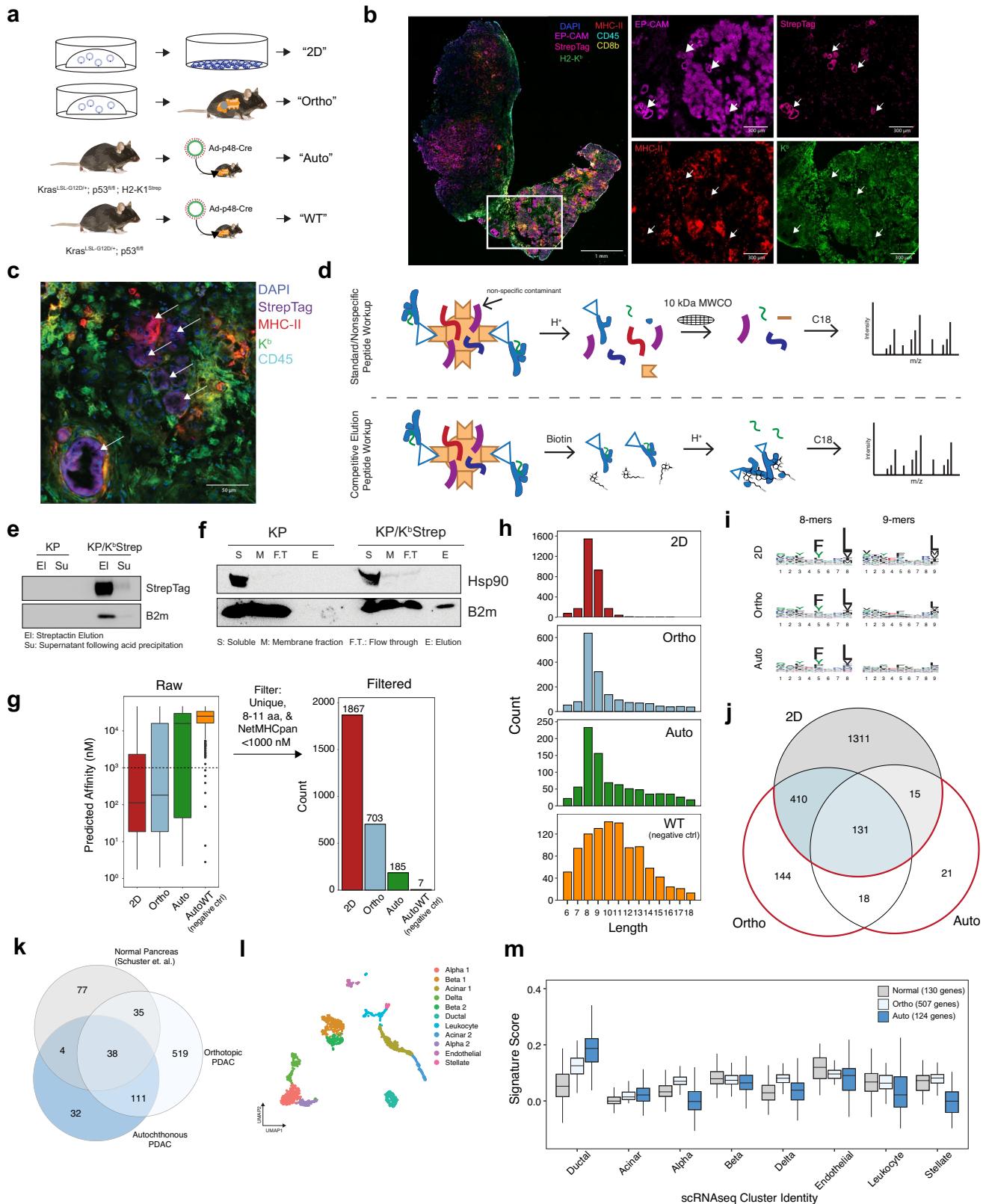
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**Extended Data Fig. 1 | (Related to Fig. 1). In vitro validation of the K<sup>b</sup>Strep allele.** **a)** Structural model depicting the topology of the K<sup>b</sup>Strep protein during affinity purification. The StrepTagII engagement with Streptactin affinity resin does not interfere with peptide or B2m binding. **b)** Southern blot analysis of K<sup>b</sup>Strep targeted KP\* ES cells. **c)** Representative genotyping for WT/WT, K<sup>b</sup>Strep/WT heterozygotes, and K<sup>b</sup>Strep/K<sup>b</sup>Strep homozygotes. **d)** Brightfield images of KP and KP/K<sup>b</sup>Strep pancreatic organoids pre- and post-Ad-CMV-Cre mediated transformation *ex vivo*. **e)** RT-PCR analysis of KP or KP/K<sup>b</sup>Strep pancreatic organoids with or without Cre recombination and with or without IFN-γ treatment. Each row represents a distinct primer set showing no discernable alterations in mRNA splicing with or without StrepTagII activation. **f)** Representative flow cytometry plots detecting cell surface

expression of PD-L1 and StrepTagII at baseline (red) and following Cre activation and IFN-γ treatment (blue). **g)** Quantification of the median fluorescence intensity (MFI) of StrepTagII staining in KP (orange) and KP/K<sup>b</sup>Strep (blue) organoids in control, IFN-γ treated, Cre transformed, and Cre+IFN-γ treated samples. Data are mean ± sem ( $n = 3$ ). Two-sided Student's *t*-test. **h)** Immunoblot analysis of whole cell lysate from KP or KP/K<sup>b</sup>Strep PDAC cells after adaptation to 2D following treatment with IFN-γ. **i)** Immunoblot depicting affinity purification of intact MHC-I with Streptactin resin as evidenced by the co-precipitation of B2m. **j)** Coomassie staining of samples taken from KP or KP/K<sup>b</sup>Strep lysates at various stages of purification. In this experiment, the elution was taken by incubating washed Streptactin resin with SDS-PAGE loading buffer.

# Article

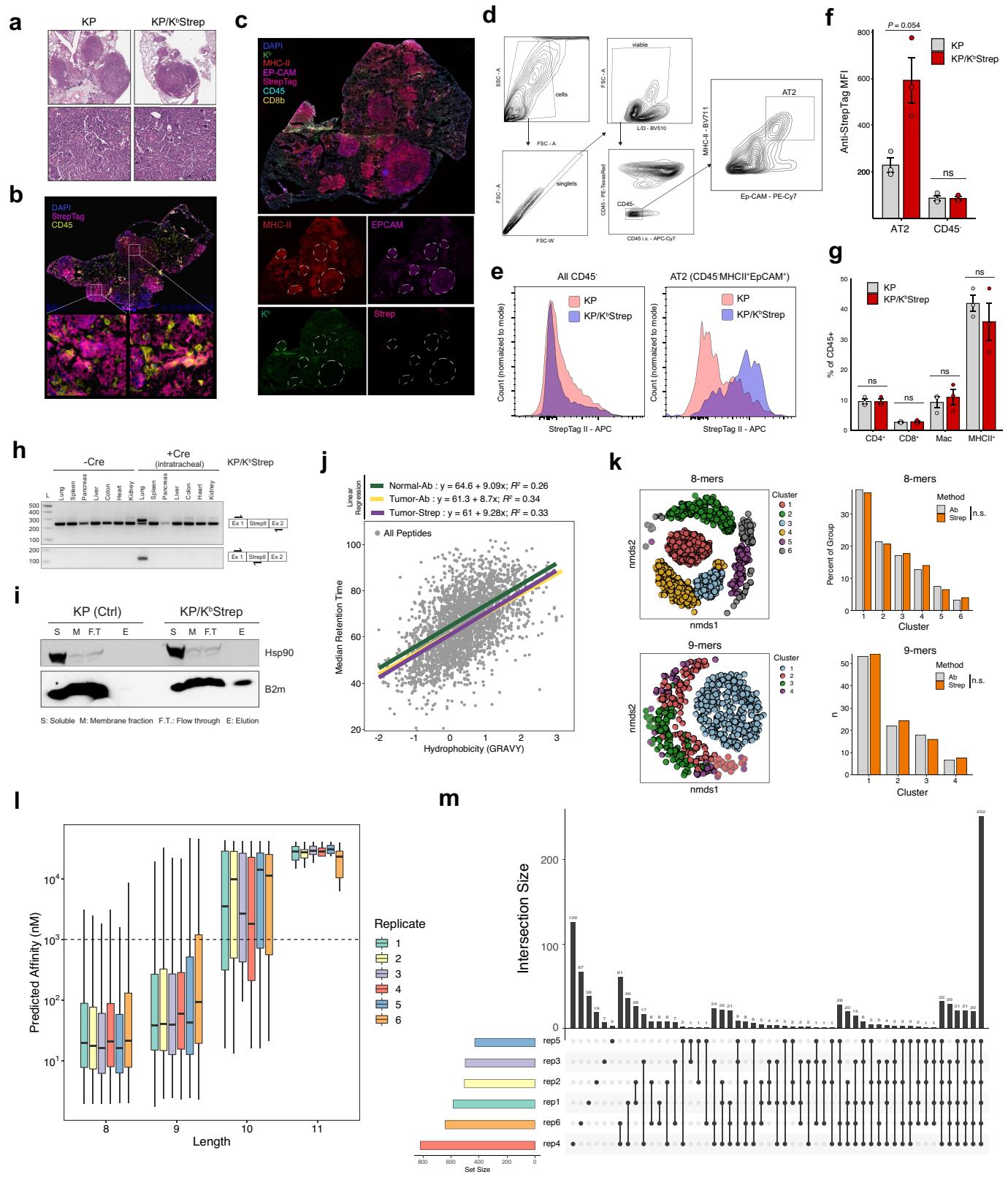


**Extended Data Fig. 2** | See next page for caption.

**Extended Data Fig. 2 | (related to Fig. 1). Isolation of MHC-I complexes from PDAC *in vivo*.** **a**) Experimental illustration of samples used for immunopeptidome comparison in PDAC. **b**) Multiplexed immunofluorescence of a representative autochthonous PDAC tumor. White arrows indicate cancer cell nests. **c**) High magnification multiplexed immunofluorescence image depicting the specificity of StrepTagII staining on cancer cells. **d**) Schematic illustration of the traditional method of peptide extraction (top) and the method used for competitive elution of MHC-I complexes with biotin used in this study (bottom). **e**) Immunoblot demonstrating efficient precipitation of the MHC-I heavy and light chain with acid (1% TFA) prior to peptide clean-up with C18 ziptip. **f**) Representative immunoblot demonstrating purification of MHC-I specifically from KP/K<sup>b</sup>Strep tumor bearing mice. **g**) Number of unique peptides identified in each sample type after filtering for length (8-11 amino acids) and NetMHCpan predicted affinity (<1000 nM). **h**) Amino acid length

distribution of all peptides identified from 2D, Ortho, Auto, or WT samples. **i**) Peptide motifs of 8- and 9-mers isolated from 2D, orthotopic, and autochthonous tumors. **j**) Venn diagram comparison of peptides found in 2D, Ortho, or Auto samples. Peptides only found *in vivo* are outlined in red. **k**) Venn diagram comparing MHC-I peptides derived from normal pancreas in Schuster et. al. (gray) versus orthotopic transplant (light blue) or autochthonous PDAC (dark blue) in this study. **l**) UMAP embedding of reanalyzed pancreatic scRNAseq data from the Tabula Muris. For clarity of presentation in Extended Data Fig. 2m, cells from a specific lineage were collapsed into a single cluster if they originally separated into multiple clusters (i.e. Alpha1 and Alpha 2 → Alpha). **m**) Expression of gene signatures derived from genes encoding for normal pancreas peptides (gray), orthotopic PDAC peptides (light blue), and autochthonous PDAC peptides (dark blue) in all cell types of the normal pancreas as measured by scRNAseq from Tabula Muris.

# Article

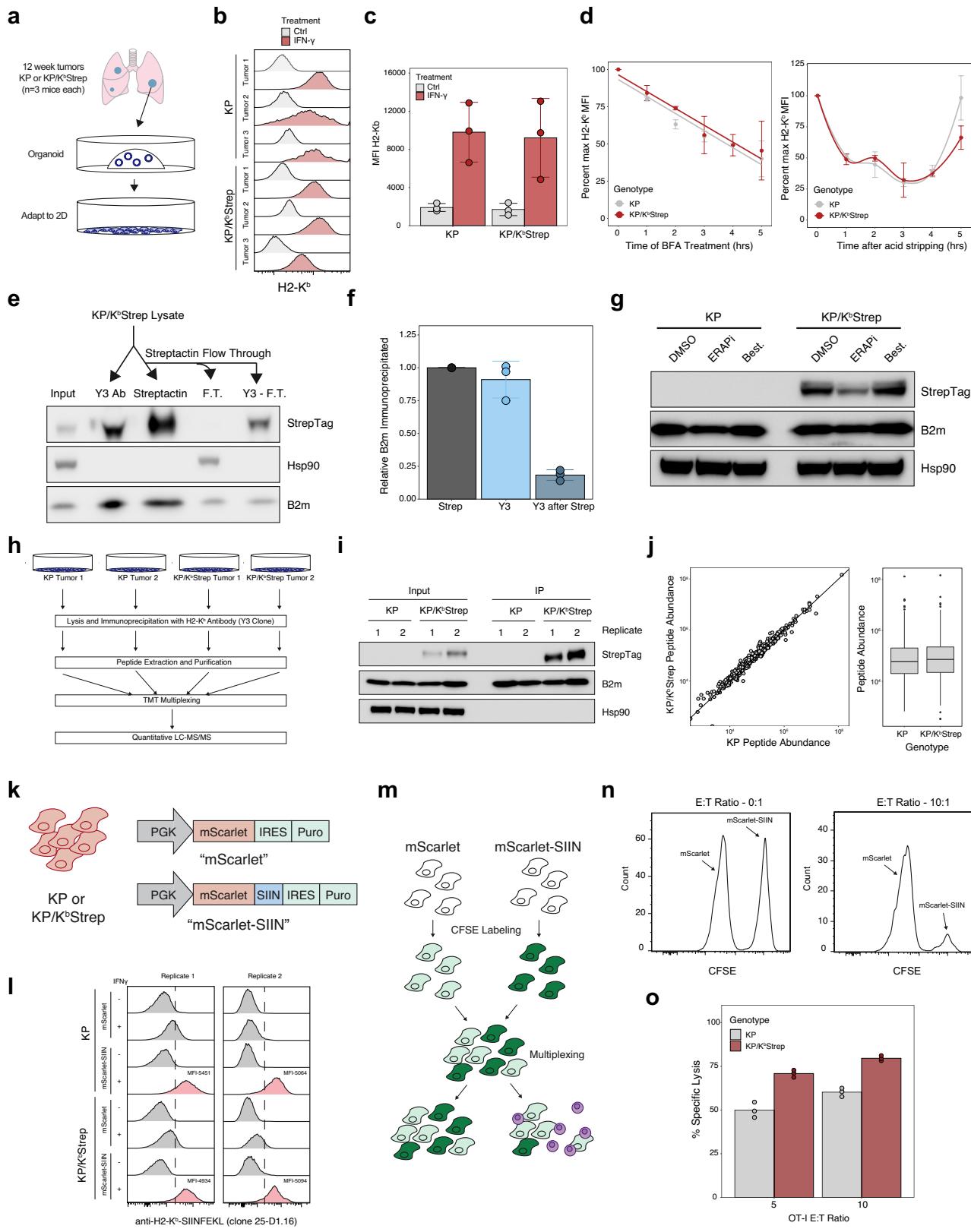


**Extended Data Fig. 3** | See next page for caption.

**Extended Data Fig. 3 | (related to Fig. 1). *In vivo* validation of K<sup>b</sup>Strep allele in KPLUAD.** **a)** Representative H&E images demonstrating adenocarcinoma in KP and KP/K<sup>b</sup>Strep tumors. **b)** Multiplexed immunofluorescence of 16-week KP/K<sup>b</sup>Strep tumors demonstrating specific StrepTagII detection on tumor cells within the tumor microenvironment. **c)** Multiplexed immunofluorescence of a WT KP tumor demonstrating no detection of the StrepTagII in tumors outlined in white dotted lines. **d)** Gating strategy for isolating cells positive for an alveolar type 2 (AT2) phenotype from KP tumor-bearing lung tissue. **e)** Histograms depicting StrepTagII staining intensity across all CD45<sup>+</sup> cells (left) or after gating for AT2 cells (right) in KP (pink) or KP; K<sup>b</sup>Strep (purple) tumors. **f)** Quantification of StrepTagII MFI on AT2 or CD45 cells in the tumor microenvironment from KP or KP/K<sup>b</sup>Strep tumors. Data are mean ± sem. Two-sided Student's *t*-test. **g)** Relative abundance of CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, Macrophages, and CD45<sup>+</sup>MHCII<sup>+</sup> immune cells in the tumor microenvironment of KP and KP/K<sup>b</sup>Strep tumors. Data are mean ± sem. Two-sided Student's *t*-test. **h)** RT-PCR analysis of diverse tissues in KP/K<sup>b</sup>Strep mice with and without intratracheal Adeno-SPC-Cre administration. Expression of the Strep tagged

K<sup>b</sup> allele is only present in the lung after Cre induction. **i)** Immunoblot depicting isolation of intact MHC-I complexes specifically from KP/K<sup>b</sup>Strep tissue as evidenced by co-purification of B2m. **j)** Comparison of predicted peptide hydrophobicity (GRAVY) versus median peptide retention time in MS analysis for all peptides presented in Fig. 1i. Linear regression is shown for peptides identified in Normal-Ab, Tumor-Ab, and Tumor-Strep datasets indicating no detectable differences in the biochemical features of identified peptides across methods. **k)** (left) Non-metric multidimensional scaling (nmds) plots depicting clusters of 8- and 9-mer peptides identified from Normal-Ab, Tumor-Ab, and Tumor-Strep samples. (right) Histogram showing the distribution of unique peptides from antibody (Ab) and Streptactin affinity purification (Strep) methods across peptide clusters identified with nmds analysis (n.s. – not significant, Fisher's exact test). **l)** Distribution of predicted peptide affinity for MHC-I peptides identified in 6 KP/K<sup>b</sup>Strep replicates. **m)** Upset plot depicting the peptide identification overlap between 6 KP/K<sup>b</sup>Strep replicates after length and affinity filtering. >77% of all identified peptides were found in at least 2/6 replicates.

# Article



**Extended Data Fig. 4** | See next page for caption.

**Extended Data Fig. 4 | (related to Fig. 1). Biochemical comparison of WT and Strep tagged H2-K<sup>b</sup>.** **a**) Isolation strategy for KP and KP/K<sup>b</sup>Strep cell lines.

**b**) Histograms depicting fluorescence staining intensity of H2-K<sup>b</sup> across KP and KP/K<sup>b</sup>Strep cell lines at baseline (gray) or following treatment with IFN-γ (red).

**c**) Median fluorescence intensity quantification of H2-K<sup>b</sup> staining across KP or KP/K<sup>b</sup>Strep cell lines at baseline (gray) or following treatment with IFN-γ (red).

**d**) Relative H2-K<sup>b</sup> staining intensity on KP (gray) or KP/K<sup>b</sup>Strep (red) cell lines following incubation with brefeldin A (BFA, left) or acute stripping with acid (300 mM glycine, pH 3.0, right) for the indicated times. **e**) Representative immunoblot depicting relative amounts of H2-K<sup>b</sup> immunoprecipitation with antibody (Y3-Ab), Streptactin, or antibody following Streptactin (Y3-F.T.).

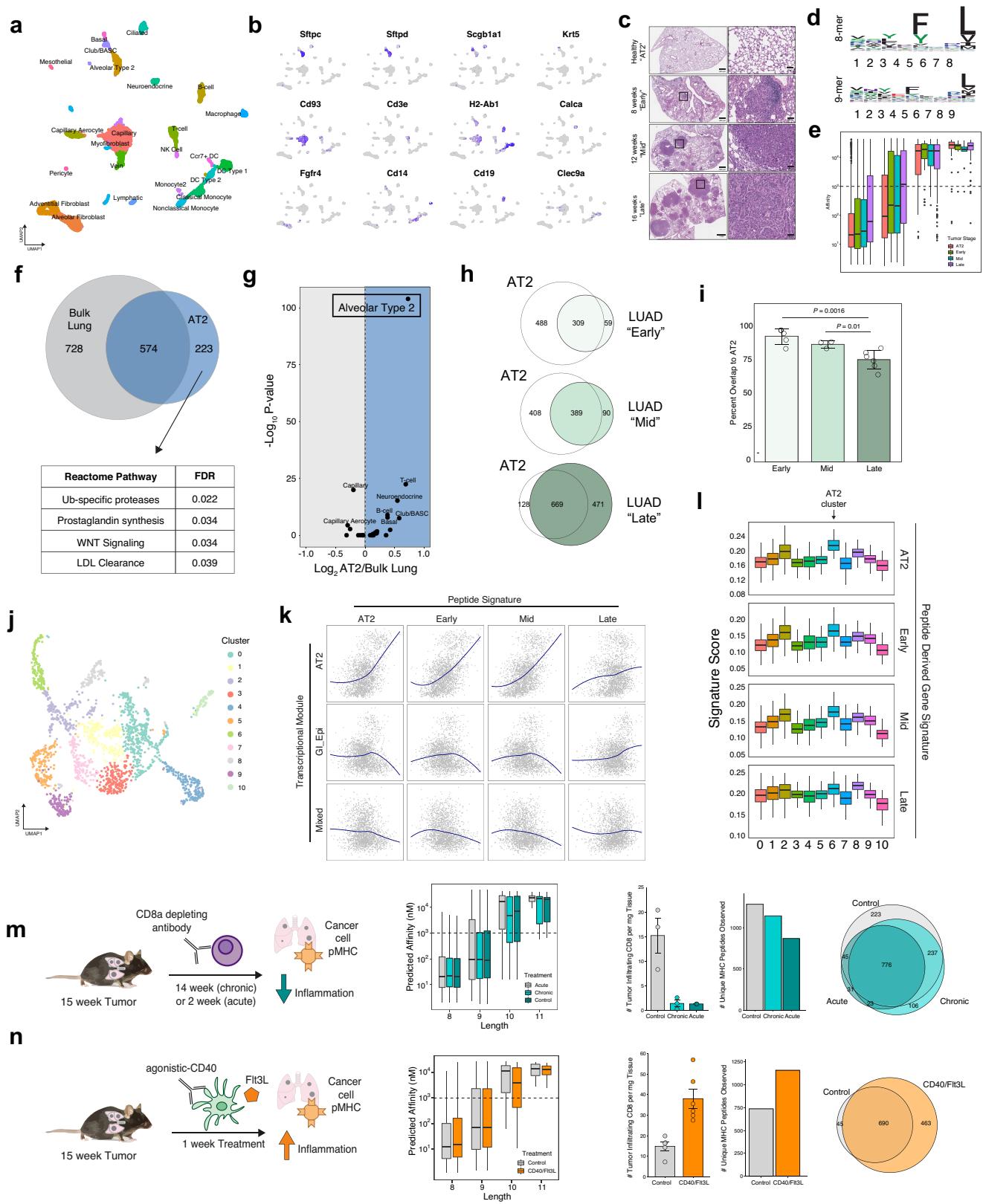
**f**) Densitometric quantification of immunoprecipitated B2m intensity following antibody (Y3), streptactin, or antibody following streptactin (Y3 after Strep) purification schemes. **g**) Representative immunoblot depicting

Strep-tagged H2-K<sup>b</sup> expression in KP or KP/K<sup>b</sup>Strep cell lines following incubation with the aminopeptidase inhibitors ERAP1-in-1 or Bestatin.

**h**) Experimental schematic for comparison of KP and KP/K<sup>b</sup>Strep immunopeptidomes from cultured cells using a quantitative, tandem mass tag (TMT) mass spectrometry strategy. **i**) Immunoblot depicting the abundance of immunoprecipitated MHC-I from samples described in h).

**j**) Quantitative abundance comparison between all peptides identified in KP and KP/K<sup>b</sup>Strep samples. **k**) Illustration of the lentiviral constructs used for stable expression of SIINFEKL in KP and KP/K<sup>b</sup>Strep cell lines. **l**) Flow cytometric analysis of SIINKFEKL-H2-K<sup>b</sup> complex surface expression using 25-D1.16 antibody. **m**) Experimental schematic used to evaluate specific T cell killing mediated by OT-ITCR transgenic T cells. **n**) Representative flow cytometry histograms depicting raw data used for calculating % specific lysis. **o**) Quantification of OT-IT cell killing in KP and KP/K<sup>b</sup>Strep cells.

# Article

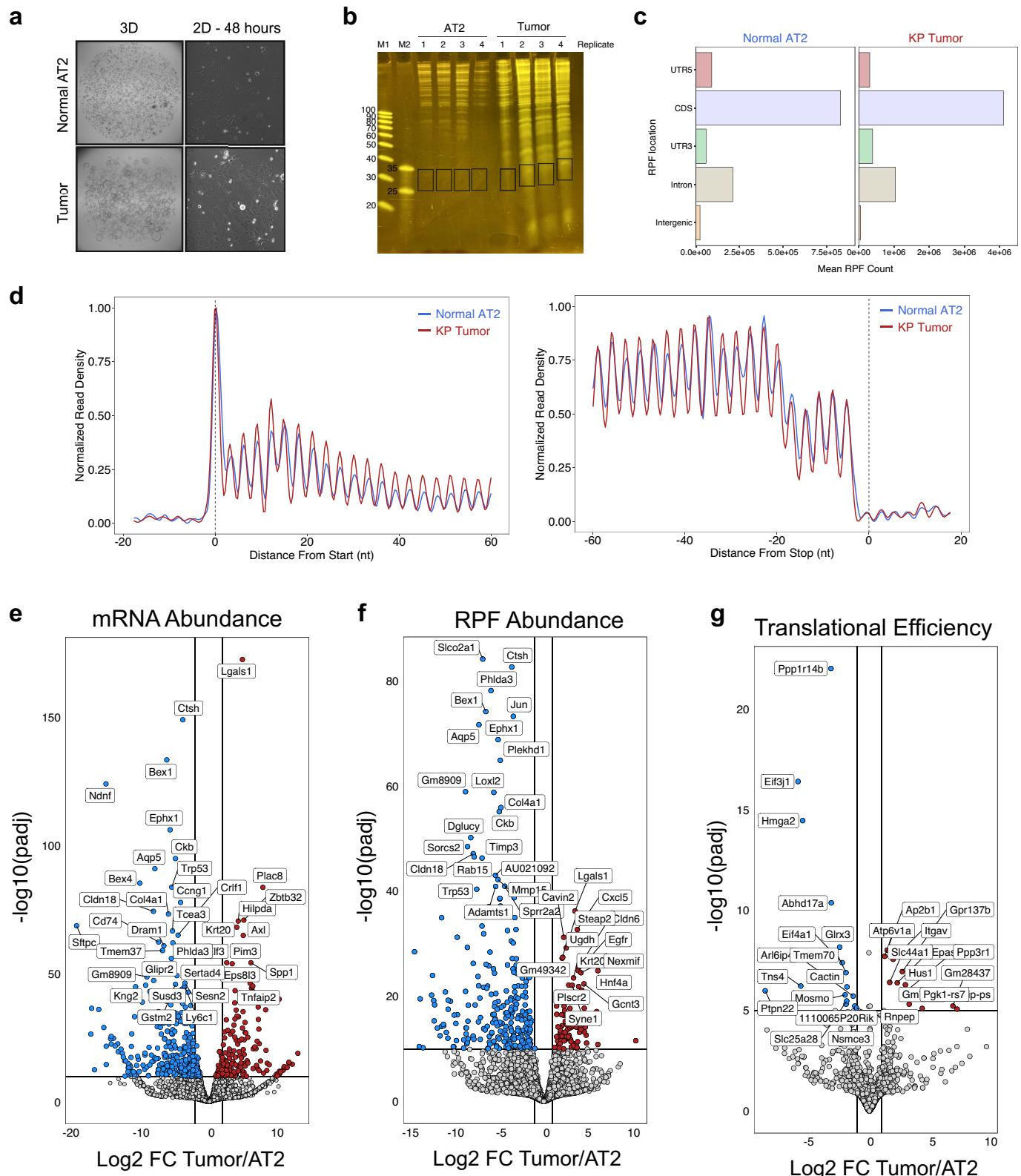


**Extended Data Fig. 5** | See next page for caption.

**Extended Data Fig. 5 | (Related to Fig. 2). Analysis of the LUAD immunopeptidome throughout tumor evolution.** **a**) UMAP embedding of clusters used for signature expression analysis in Fig. 2a. **b**) Gene expression profiles of cell type marker genes indicating robust clustering of known cell types in the healthy lung. **c**) Representative H&E stains for healthy lung (AT2), Early-, Mid- and Late-stage tumor samples. **d**) Peptide motifs of 8- and 9-mer peptides identified from healthy AT2 cells. **e**) Length and affinity characteristics of peptides identified in AT2, Early-, Mid-, and Late-stage tumors. **f**) Venn diagram comparison of peptides identified in bulk, healthy lung versus those identified specifically presented by normal AT2 cells. Pathways enriched by gene ontology depicted on bottom. **g**) Comparative analysis of gene signatures derived from peptides detected on AT2 cells versus bulk lung applied to the Tabula Muris data. Volcano plot shows the strong enrichment for an AT2 phenotype in the AT2 immunopeptidome versus bulk

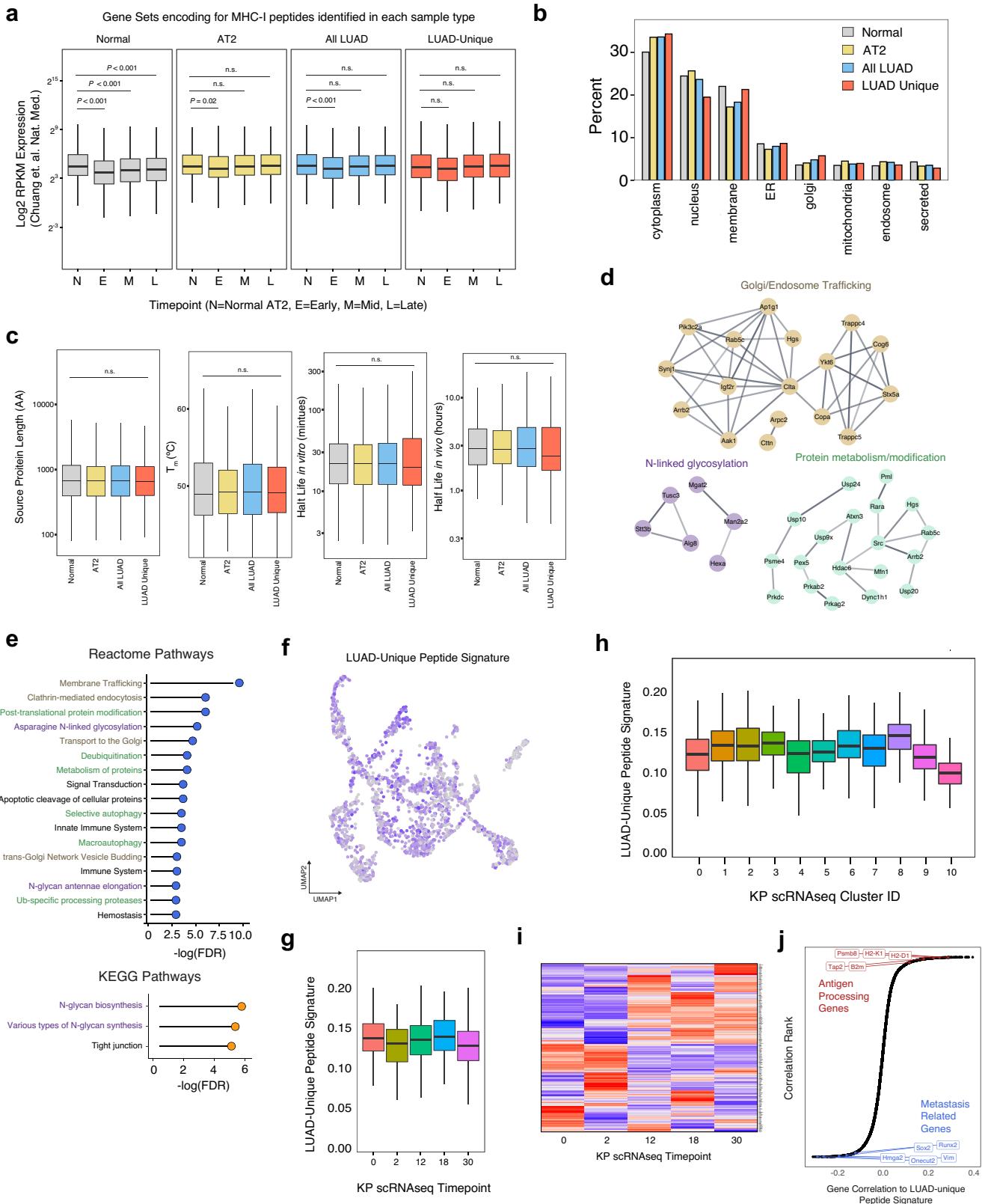
lung immunopeptidome. **h**) Comparison of AT2 and Early-, Mid-, and Late-Stage immunopeptidomes. **i**) Quantification of percent overlap from h. **j**) UMAP embedding of clusters from reanalyzing scRNAseq data from Marjanovic et. al. and used for analysis of Fig. 2j. **k**) Loess regression analysis across all cells scored for AT2, Early, Mid, and Late peptide signatures versus the AT2, GI-Epi, and Mixed transcriptional modules. **l**) Signature distribution for AT2-, Early-, Mid-, and Late-Stage peptide signatures across all KP scRNAseq clusters. **m**) Comparison of the immunopeptidome in control tumors (gray), tumors chronically depleted of CD8 T-cells (light teal), and tumors acutely depleted of CD8 T-cells (dark teal). **n**) Comparison of the immunopeptidome from control tumors (gray) or tumors treated with agonistic-CD40 antibody and Flt3-L (orange). CD8 depletion experiments and CD40/Flt3L experiments were carried out independently and analyzed on different MS runs separated by ~2.5 months.

# Article



**Extended Data Fig. 6 | (Related to Fig. 3). Identification of differentially translated genes in LUAD versus AT2 cells.** **a**) Representative images of normal AT2 and tumor organoid cultures in 3D organotypic culture or after transient adaptation to 2D monolayer culture prior to RiboSeq and RNaseq processing. **b**) Denaturing PAGE gels of RNA purified from RiboLace purification. Excised bands used for RiboSeq are indicated on the right and were selected for RNA that was ~30 bp in length. **c**) Localization of RPF alignments within transcripts depicting enrichment for annotated coding sequences (CDS). **d**) Normalized metagene density profiles of reads from normal and tumor cells at translation initiation (left) and termination (right). Both normal and tumor metaprofiles exhibit 3 nucleotide periodicity, indicative of active translation. **e**) Volcano plot depicting differential mRNA expression in tumor versus normal AT2 cells. **f**) Volcano plot depicting differential RPF abundance in tumor versus normal AT2 cells. **g**) Volcano plot depicting differential translation efficiency in tumor versus normal AT2 cells.

sequences (CDS). **d**) Normalized metagene density profiles of reads from normal and tumor cells at translation initiation (left) and termination (right). Both normal and tumor metaprofiles exhibit 3 nucleotide periodicity, indicative of active translation. **e**) Volcano plot depicting differential mRNA expression in tumor versus normal AT2 cells. **f**) Volcano plot depicting differential RPF abundance in tumor versus normal AT2 cells. **g**) Volcano plot depicting differential translation efficiency in tumor versus normal AT2 cells.

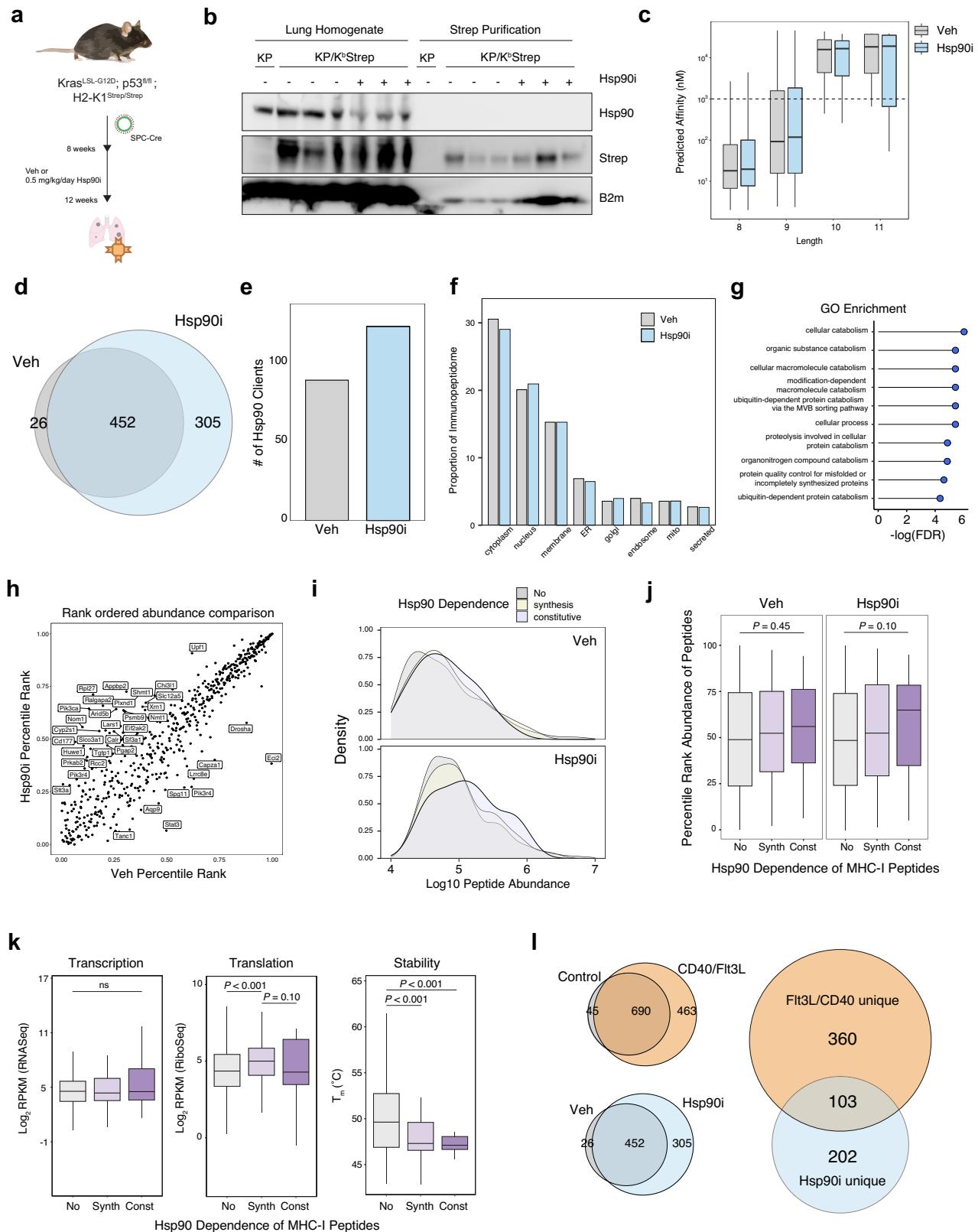


**Extended Data Fig. 7** | See next page for caption.

# Article

**Extended Data Fig. 7 | (Related to Fig. 3). Transcriptomic and proteomic data associated with LUAD-unique peptides.** **a)** mRNA expression of genes encoding for Normal (bulk-lung), AT2, All-LUAD, or LUAD-unique peptides across normal AT2 cells, early-, mid-, and late-stage sorted tumor cells. (Adapted from Chuang et. al.). **b)** Subcellular compartment distribution of source proteins for peptides found in bulk Normal Lung, AT2 cells, all tumor peptides, and LUAD-unique peptides. **c)** Distribution of Protein length, thermal stability, or protein half-life for source proteins of peptides found in Normal Tissue, All Tumor peptides, or LUAD-unique peptides. **d)** StringDb analysis of source proteins for LUAD-unique peptides indicated in Fig. 3a. Clusters of

enriched protein families are depicted. **e)** Gene ontology analysis of LUAD-unique peptides from KEGG and Reactome databases. **f)** Expression of the LUAD-unique signature across all cells in the KP scRNASeq dataset (Marjonovic et. al.) **g)** Expression of the LUAD-unique signature across tumor progression in KP scRNASeq. **h)** Expression of LUAD-unique peptide signature across clusters in KP scRNASeq. **i)** Expression of individual genes encoding LUAD-unique peptides across KP timepoints. **j)** Correlation of the LUAD-unique peptide signature to all genes detected in scRNASeq. Genes related to antigen presentation are highlighted in red and genes related to metastasis are highlighted in blue.



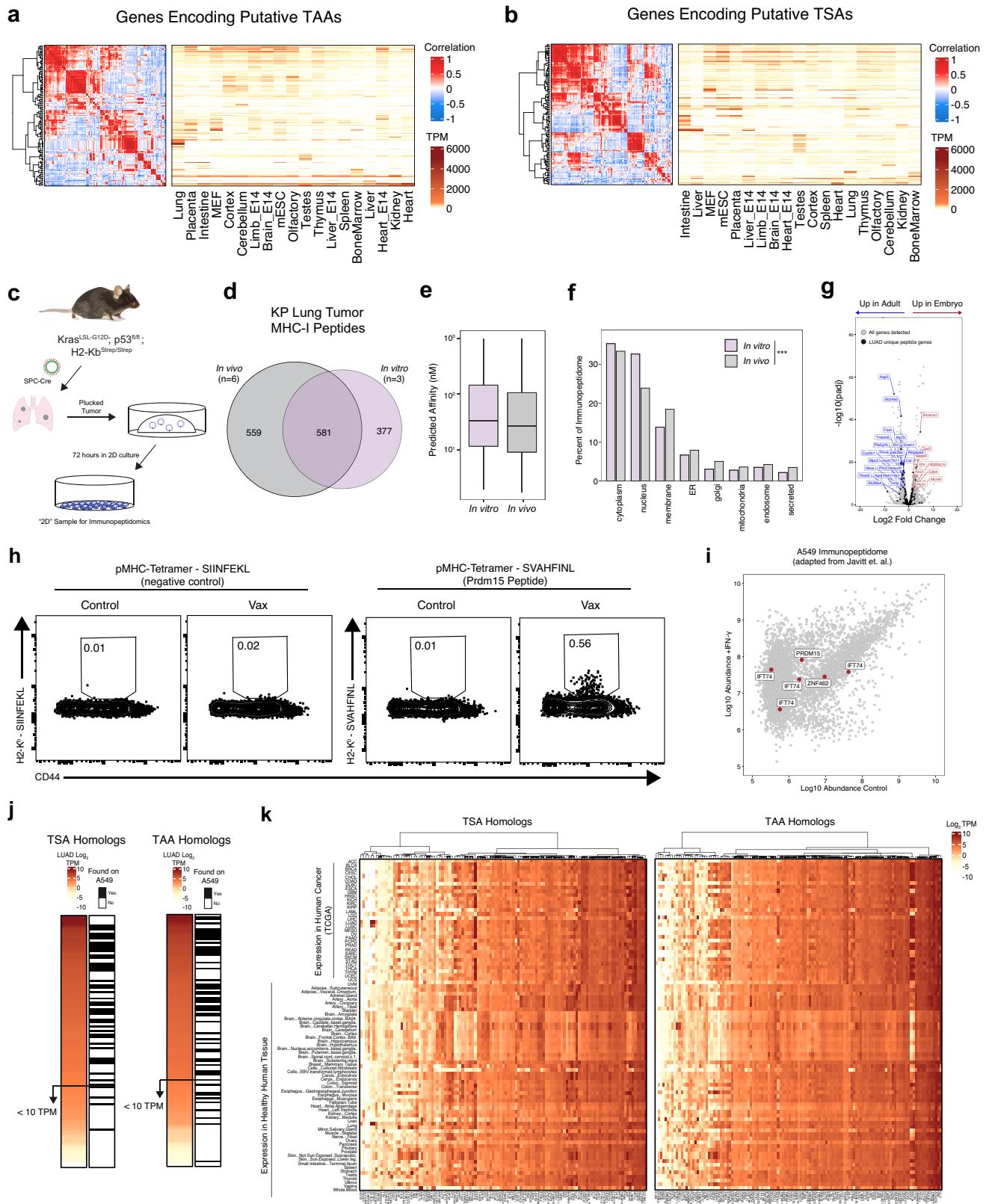
**Extended Data Fig. 8** | See next page for caption.

# Article

## Extended Data Fig. 8 | (Related to Fig. 3). Modulation of protein folding through Hsp90 inhibition alters the immunopeptidome *in vivo*.

**a)** Experimental schematic of KP/K<sup>b</sup>Strep tumor treatment with either vehicle control or 0.5 mg/kg/day NVP-HSP990 prior to tumor specific MHC-I isolation. **b)** Immunoblot analysis of purified MHC-I from Vehicle (Veh) and Hsp90 inhibitor (Hsp90i) treated tumor samples or KP control tumors. **c)** Length and affinity distribution of peptides found in Veh (grey) and Hsp90i (blue) treated samples. **d)** Venn diagram of peptides found in 12-week control tumors or Hsp90i treated samples. **e)** Number of Hsp90 clients giving rise to peptide in either Veh (grey) or Hsp90i (blue) samples. **f)** Distribution of peptides identified in Veh or Hsp90i treated samples across subcellular compartments. **g)** Gene ontology analysis of source proteins for peptides that were only found

in Hsp90i treated samples ranked according to FDR enrichment significance. **h)** Comparing the rank ordered abundance of all common peptides between Hsp90i treatment and control. **i)** Density plots of raw peptide abundance for non-clients, synthesis clients or constitutive clients in Veh (top) or Hsp90i (bottom) treated samples. **j)** Rank ordered abundance of peptides derived from Non-clients (No, gray), synthesis clients (Synth, light purple) or constitutive clients (Const, dark purple) in Veh and Hsp90i treated samples. *P* calculated with the Kologorov-Smirnov Test. **k)** RNA abundance, translation rate, and melting temperatures across non-clients (No), synthesis clients (Synth) and constitutive clients (Const) that are source proteins for MHC-I presentation. *P* calculated with Mann-Whitney *U* Test. **l)** Comparison of peptides unique to Flt3L/aCD40 treatment and those unique to Hsp90i.



**Extended Data Fig. 9** | See next page for caption.

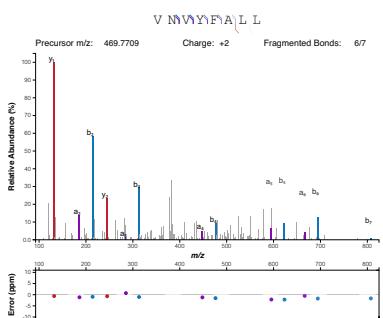
# Article

**Extended Data Fig. 9 | (Related to Fig. 4). Expression and presentation of putative tumor specific and tumor associated antigens.** **a)** Correlogram and heatmap depicting transcript abundance (transcripts per million, TPM) of putative TSA genes across mouse tissues. **b)** Correlogram and heatmap depicting TPM abundance of putative TAA genes across mouse tissues. **c)** Experimental schematic showing the derivation of samples for 2D immunopeptidomics. **d)** Venn diagram depicting the relationship between peptides identified by KP tumors *in vivo* and those identified *in vitro*. **e)** Boxplot showing the predicted affinity distributions of peptides isolated *in vivo* and *in vitro*. **f)** Distribution of source protein subcellular compartments for peptides identified *in vivo* (gray) and *in vitro* (green). **g)** Volcano plot indicating differentially expressed genes between EPCAM+ cells from embryonic day 16.5 and post-natal day 28 mouse lung (Adapted from Lung Map Project). Data

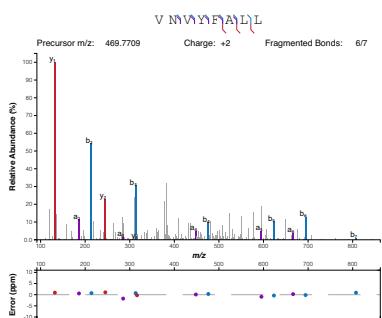
analyzed and  $P$  calculated with DEseq2. All genes detected are shown in grey and genes encoding for LUAD-unique peptides are indicated with black dots. **h)** Flow cytometry analysis of tumor-bearing lung tissue from naïve and vaccinated mice stained with control pMHC-I tetramer (SIINFEKL) or TAA tetramer (SVAHFINL). **i)** Peptides identified in A549 cells (Javitt et. al.) with and without treatment of IFN- $\gamma$ /TNF $\alpha$ . Peptides derived from source proteins homologous to those using in the pooled vaccine are indicated in red. **j)** Heatmap depicting expression of the human homologs of putative TSAs and TAAs from this study and whether or not peptides derived from those genes were found to be presented on A549 cells from Javitt et. al. **k)** Heatmap depicting the RNA Expression of homologs of potential TSA and TAA genes as found in Fig. 4a across all individual human tissues and 33 cancer types within TCGA.

**a**

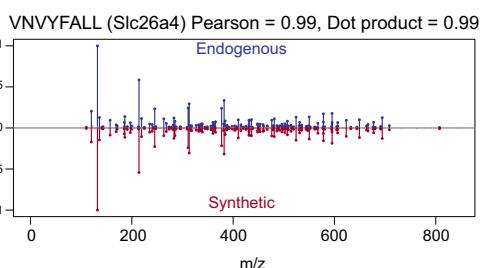
## Endogenous



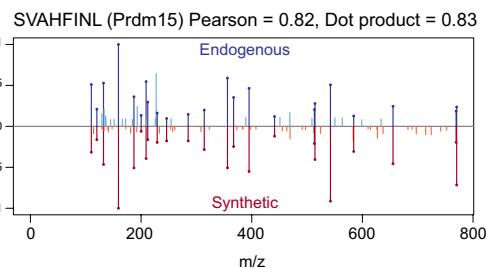
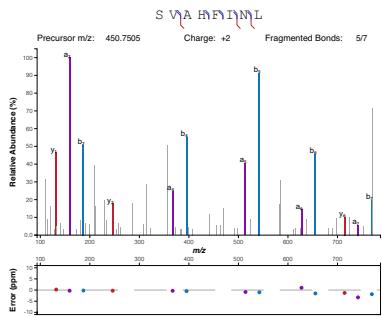
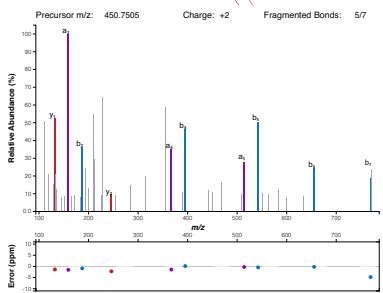
## Synthetic



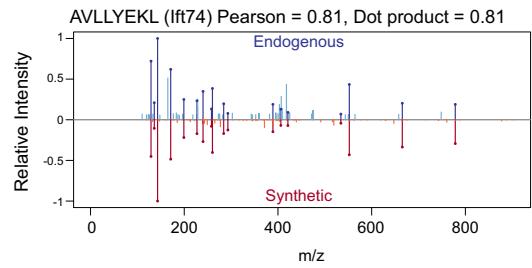
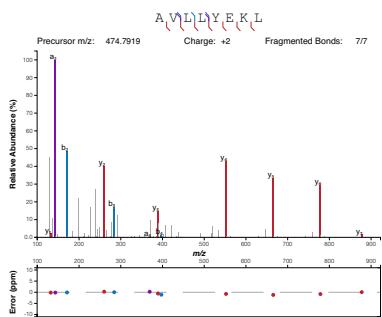
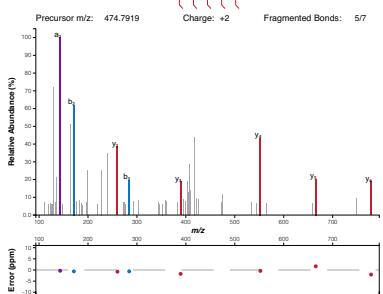
## Correlation

**b**

## S V A H F I N L

**c**

## A V L L Y E K L



**Extended Data Fig. 10 | (Related to Fig. 4). Mass spectrometry validation of immunogenic epitopes with synthetic peptides.** **a)** Mass spectrometry comparison of spectra from endogenously identified VNVYFALL peptide (Slc26a4) and a synthetic standard. **b)** Mass spectrometry comparison of spectra from endogenously identified SVAHFNL peptide (Prdm15) and a

synthetic standard. **c)** Mass spectrometry comparison of spectra from endogenously identified AVLLYEKL peptide (Ift74) and a synthetic standard. In the left panel, y-, b-, and a-ions are colored in bold. In the right panel, common peaks are drawn in darker colour.

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection	no software used for data collection
Data analysis	NetMHCPan (v4.1), FlowJo (v10), Microsoft Excel (v16), Seurat (v4.0), R (v4.0), RStudio (v1.3), RiboToolKit, QuPath (v0.2.3), GSEA (v4.1), StringDb (v11), kallisto (0.46.0),

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Raw mass spectrometry data files and raw ribosome sequencing data files were submitted to publicly available databases as described in the methods (PRIDE and GEO, respectively)

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was not chosen based on a predetermined statistical power calculation. For these studies, our goal was an unbiased description of peptides presented by MHC-I and quantitative comparisons between animals was not performed.
Data exclusions	The only data processing that occurred was to ensure high confidence in mass spectrometry analysis (i.e. filtering peptide spectra for low false discovery rate or manual inspection of spectra for quality).
Replication	In vivo experiments were performed with a minimum of three individual mice per sample group. Mass spectrometry analyses of the data obtained from these mice were performed in technical duplicates where possible to ensure reproducibility and confidence in the reported findings.
Randomization	Randomization was not required for this study given that each animal/sample was used for a separate discovery analysis.
Blinding	Blinding was not required for this study as there were no subjective analyses performed that would be biased by knowledge of subjects to be analyzed.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

- |                                     |   |
|-------------------------------------|---|
| n/a                                 | Involved in the study   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Antibodies                  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Eukaryotic cell lines       |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology          |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                          |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern           |

### Methods

- |                                     |  |
|-------------------------------------|--|
| n/a                                 | Involved in the study                              |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq                  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging    |

## Antibodies

### Antibodies used

- StrepTag Biolegend 688202 Western Blot/FACS
- StrepTag Abcam ab76949 Western Blot/FACS/IF
- Hsp90 Cell Signaling 4877 Western Blot
- B2m Abcam ab218230 Western Blot
- CD45-eFluor780 eBioscience 47-0451-82 FACS
- CD45-PE-CF594 BD Biosciences 562420 FACS
- CD45-PE eBioscience 12-0451-83 IF
- CD45-BV570 Biolegend 103136 IF
- CD44-BV785 Biolegend 103041 FACS
- Ep-CAM-PE-Cy7 Biolegend 118216 FACS
- Ep-CAM-BV650 Biolegend 118241 IF
- CD8-BUV805 BD Biosciences 612898 FACS
- CD4-BV785 Biolegend 100552 FACS
- MHC-II-BV711 Biolegend 107643 FACS
- MHC-II-BV785 Biolegend 107645 IF
- H2-Kb-FITC eBioscience 11-5958-82 IF
- PDL1-PE-Cy7 Biolegend 124314 FACS
- CD8b-PE eBioscience 12-0083-82 IF
- H2-Kb-SIINFEKL - PE Biolegend 141603 FACS
- H2-Kb Purified BioXCell BE0172 Immunoprecipitation

CD8a Depleting antibody BioXCell BE0061 In vivo T cell depletion  
 agonistic CD40 BioXCell BP0016-2 In vivo DC stimulation  
 Flt3L BioXCell BE0098 In vivo DC stimulation

## Validation

Validation information for each antibody can be found on the vendor's website. No antibodies were empirically validated in house but appropriate controls were included for each experiment to ensure accurate conclusions were drawn.

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	All cell lines used were generated and described for the first time in this study.
Authentication	All de novo generated cell lines were verified with PCR genotyping and/or FACS verification.
Mycoplasma contamination	Not tested.
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	None used.

## Animals and other organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Female and Male C57BL/6J C57BL/6J Kras-LSL-G12D; p53 fl/fl C57BL/6J Kras-LSL-G12D; p53 fl/fl; KbStrep fl/fl C57BL/6J SftpcCreERT2; KbStrep fl/fl
Wild animals	None.
Field-collected samples	None.
Ethics oversight	Massachusetts Institute of Technology Committee on Animal Care

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Flow Cytometry

### Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

Sample preparation	Sample preparation from organoids and mouse lung tissue is described in detail in the methods section.
Instrument	Data was obtained on an LSR-II or Fortessa
Software	Data was collected with BD Facs Diva and analyzed with FlowJo v10
Cell population abundance	Conclusions regarding cell population abundance were only used in the evaluation of Tetramer+ T cells in Figure 4f. For these comparisons, a customized gate was drawn based on the negative control samples. This gate was then applied to all other samples to obtain a population abundance for each sample.
Gating strategy	For analysis of cells with an AT2 phenotype, single cells were gated on FSC/SSC, followed by gating on negative staining for Zombie Viability Dye (BV510). CD45- cells were identified by negative staining for both intravascular (APC-Cy7) and ex vivo CD45 staining (PE-Texas Red). AT2 cells were then enriched for a population that was double positive for EPCAM (PE-Cy7) and MHC-II (BV711). The MFI of rabbit StrepTagII (anti-rabbit 647 secondary; APC) was then determined for each sample.

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.