

Supplementary Information for

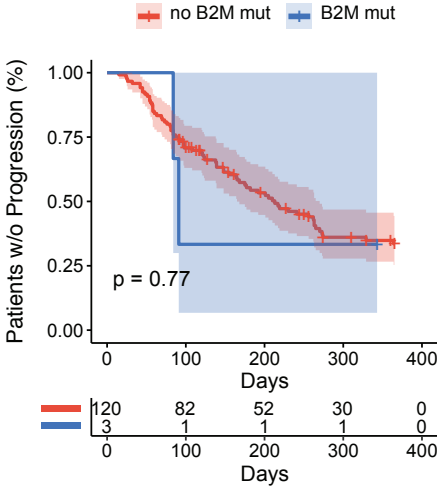
Integration of tumor extrinsic and intrinsic features associates with immunotherapy response in non-small cell lung cancer

Denise Lau, Sonal Khare, Michelle M. Stein, Prerna Jain, Yinjie Gao, Aicha Ben-Taib, Tim A. Rand, Ameen Salahudeen, Aly A. Khan

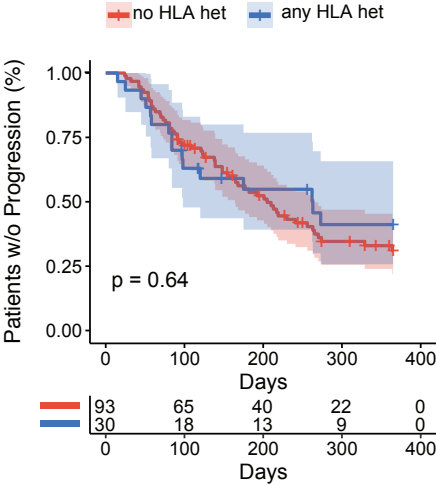
Includes:

Supplementary Fig. 1-9

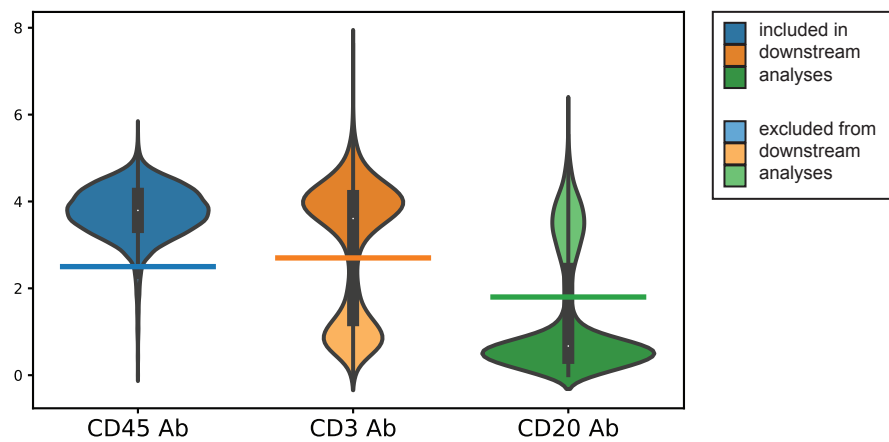
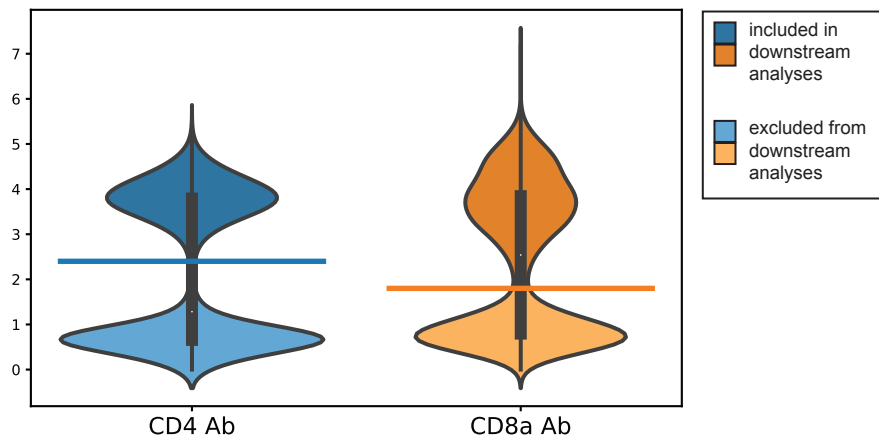
a



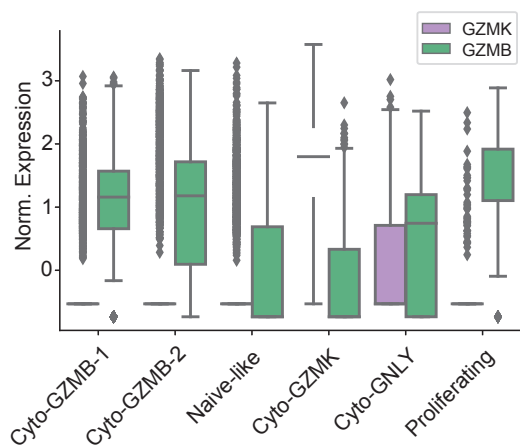
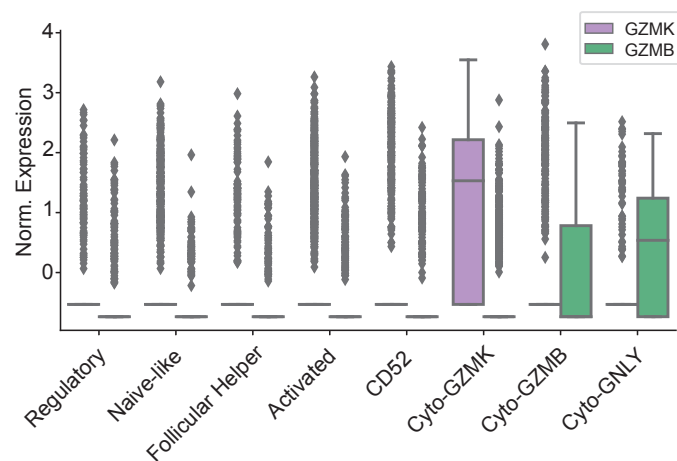
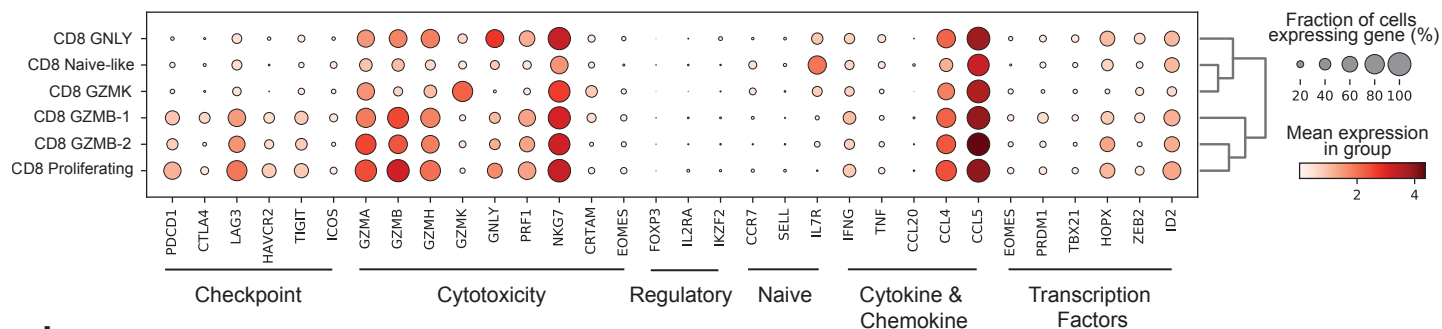
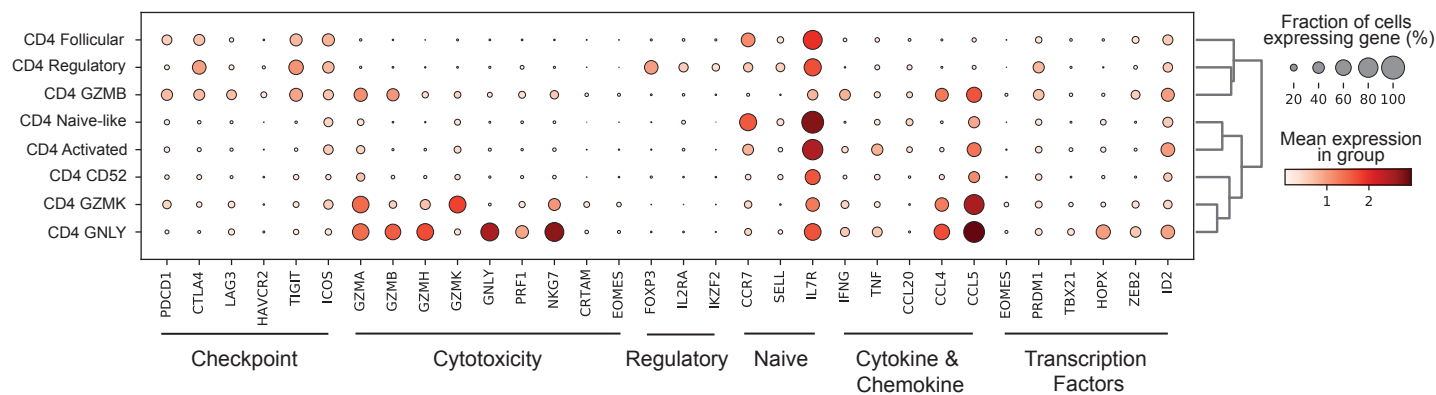
b



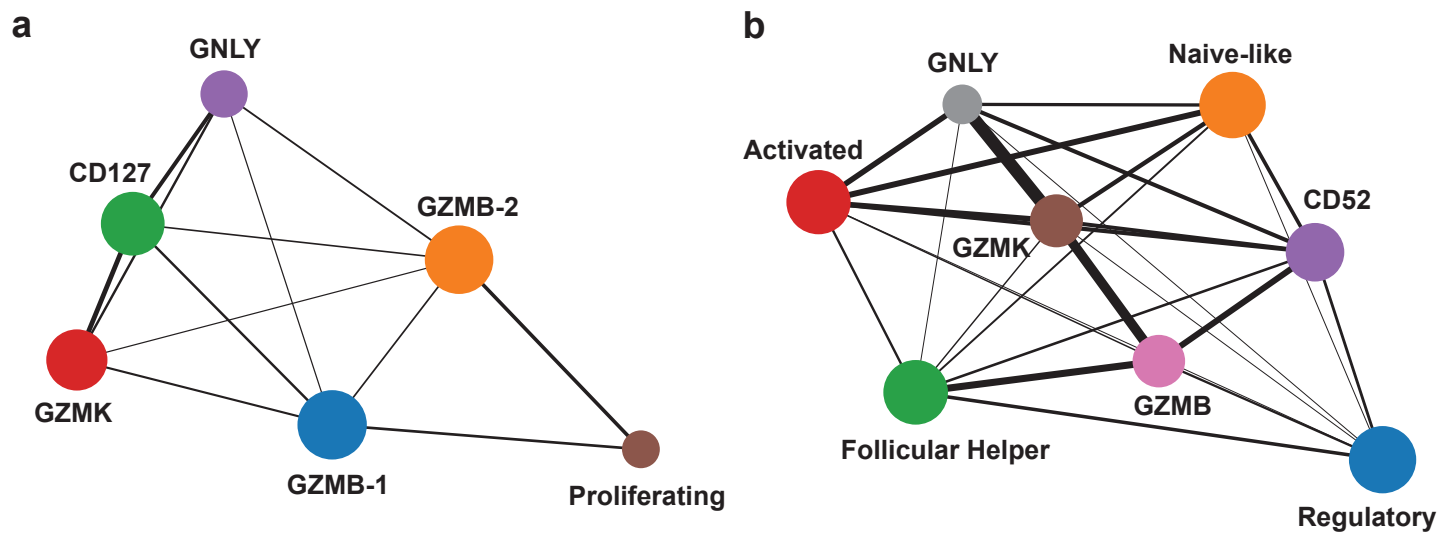
Supplementary Fig. 1: **HLA class I deficiencies and TTP in the NSCLC ICB cohort.** **a**, Kaplan-Meier plots showing time to progression on ICB therapy, stratified by *B2M* mutation (n=123, HR=1.24, p=0.77, log rank), and **b**, HLA homozygosity status (n=123, HR=0.88, p=0.64, log rank). Source data are provided as a Source Data file.

a**b**

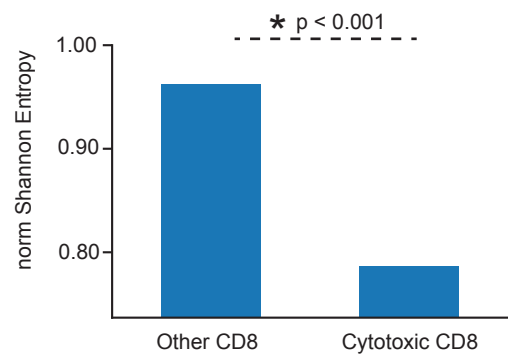
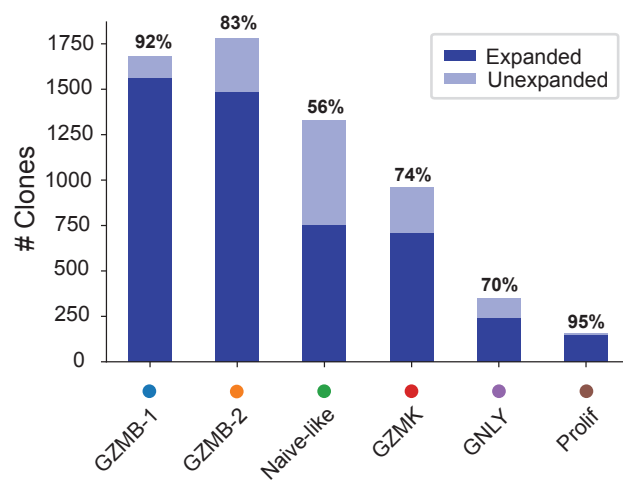
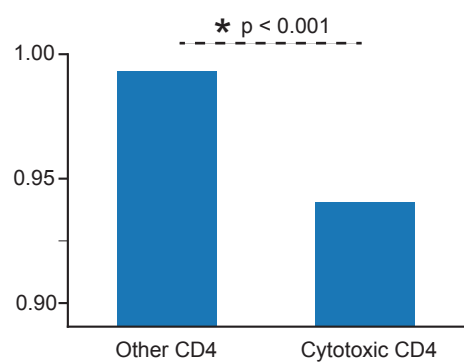
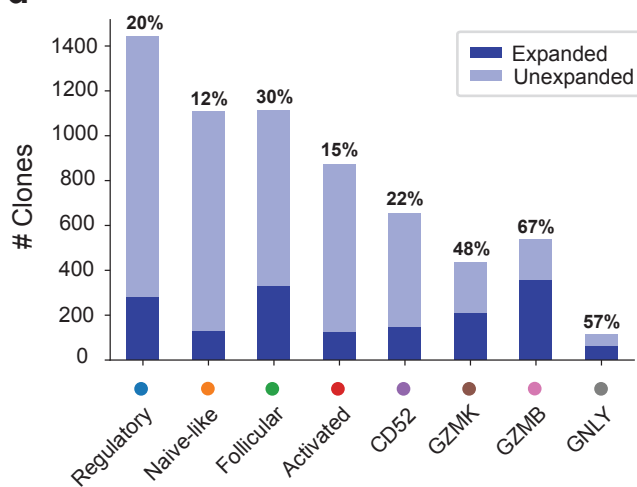
Supplementary Fig. 2: **Protein expression distribution of lineage markers.** **a**, Violin plots of normalized protein expression distribution of CD45 (blue), CD3 (orange) and CD20 (green) in 47,991 cells. Box inside each violin represents the interquartile range (IQR), whiskers extend up to 1.5 times the interquartile range (IQR). The point in white represents median expression. Each line indicates the threshold used to computationally isolate T cells for downstream analysis, where area of violin in bold color was included, and lighter region was excluded. **b**, Violin plots of normalized protein expression distribution (n=29,391 cells) of CD4 (blue) and CD8a (orange) following isolation of T cells using the thresholds in panel a. Each line indicates the threshold used to computationally isolate cells. Area of violin in bold color was considered CD4+ (using CD4) or CD8+ (CD8a). Box inside each violin represents the interquartile range (IQR), whiskers extend up to 1.5 times the interquartile range (IQR). The point in white represents median expression. Source data are provided as a Source Data file.

a**CD8⁺ T cells****b****CD4⁺ T cells****c****d**

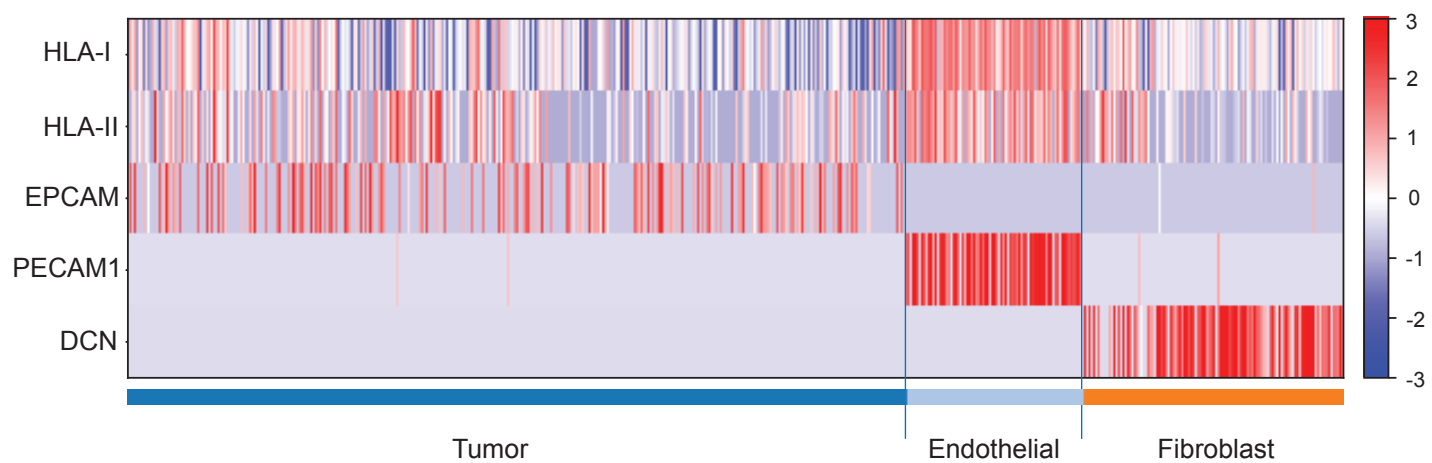
Supplementary Fig. 3: **GZMB** and **GZMK** expression in CD4⁺ and CD8⁺ T cells. **a**, Normalized RNA expression of *GZMK* and *GZMB* for each Leiden cluster in the CD8⁺ (Cyto-GZMB-1: 3933 cells; Cyto-GZMB-2: 3743 cells; Naive-like: 2779 cells; Cyto-GZMK: 2350 cells; Cyto-GNLY: 811 cells; Proliferating: 319 cells) and **b**, CD4⁺ T cell populations (Regulatory: 2524 cells; Naive-like: 2398 cells; Follicular Helper: 2138 cells; Activated: 2085 cells; CD52: 1426 cells; Cyto-GZMK: 944 cells; Cyto-GZMB: 911 cells; Cyto-GNLY: 285 cells). The box in the boxplot represents the interquartile range, with the center line at the median. The whiskers extend up to 1.5 times the interquartile range (IQR). **c**, Bubble plot showing the expression of key immune function genes, including cytotoxic genes, in the CD8⁺ T cell and **d**, the CD4⁺ T cell Leiden clusters. Bubble size reflects the fraction of cells that have non-zero expression of the gene and the color represents the mean expression of the population. Source data are provided as a Source Data file.



Supplementary Fig. 4: **Trajectory analysis of CD4⁺ and CD8⁺ T cell populations.** **a**, PAGA graph of the CD8⁺ and **b**, CD4⁺ T cell populations. The nodes represent the Leiden clusters and the edges represent a statistical measure of connectivity based on RNA expression. The node size reflects the number of cells belonging to the cluster.

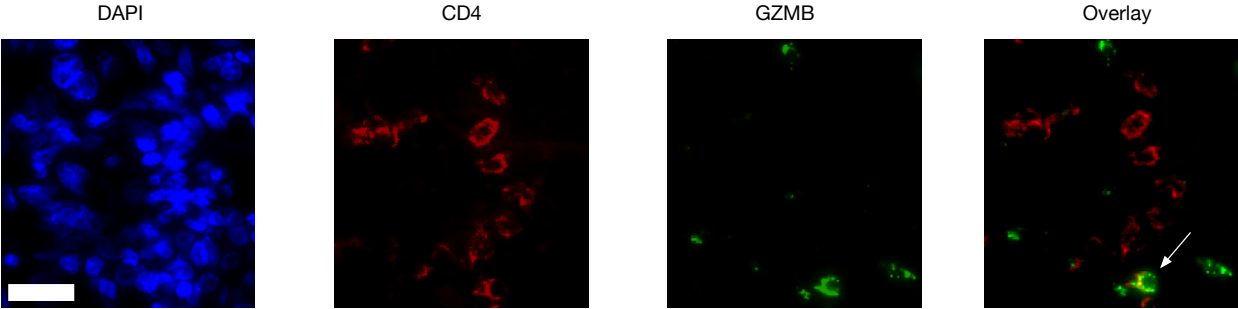
a**b****c****d**

Supplementary Fig. 5: **Clonal expansion in CD4⁺ and CD8⁺ T cells.** **a**, Normalized Shannon entropy for the CD8⁺ cells and **c**, CD4⁺ cells in cytotoxic clusters compared to those in the non-cytotoxic clusters (CD8⁺: $p=0$, CD4⁺: $p=0$, two-sided Hutcheson's t-test). Proportion of clones that were expanded (more than one cell) for each Leiden cluster in **b**, CD8⁺ and **d**, CD4⁺ T cells. The percentage of expanded clones is shown above each bar. Source data are provided as a Source Data file.



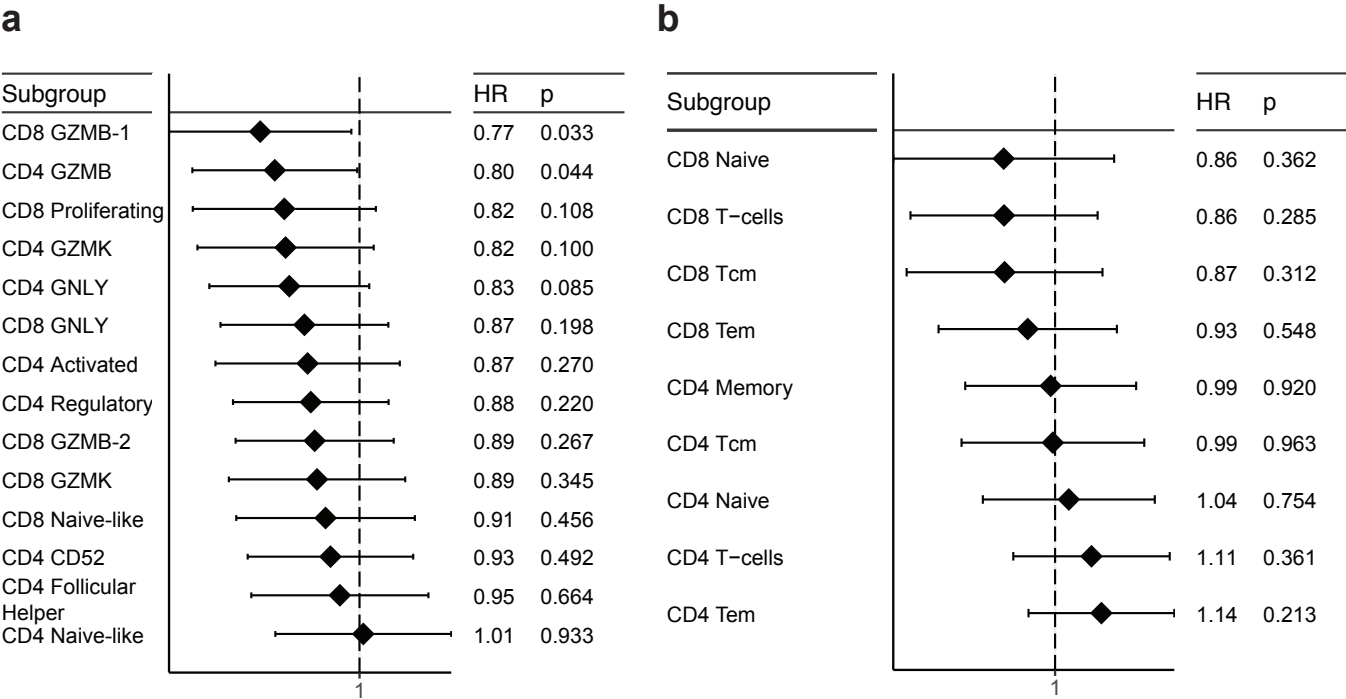
Supplementary Fig. 6: **HLA and lineage marker expression in the CD45⁺ population.** Heatmap showing the expression of HLA genes, as well as the lineage markers used to identify the tumor, endothelial and fibroblast populations in the CD45⁺ fraction. Each column represents one cell. RNA expression is mean centered and scaled. Source data are provided as a Source Data file.

Patient 2

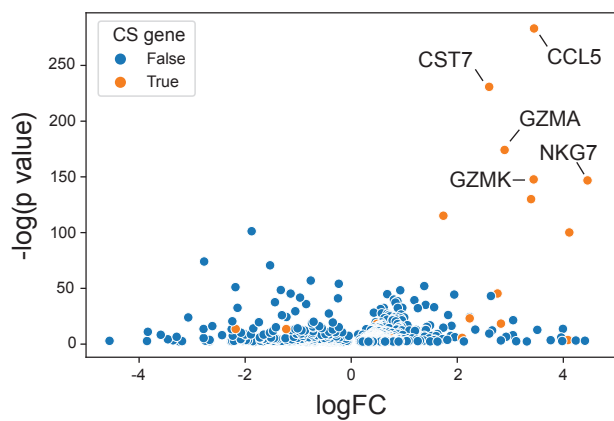
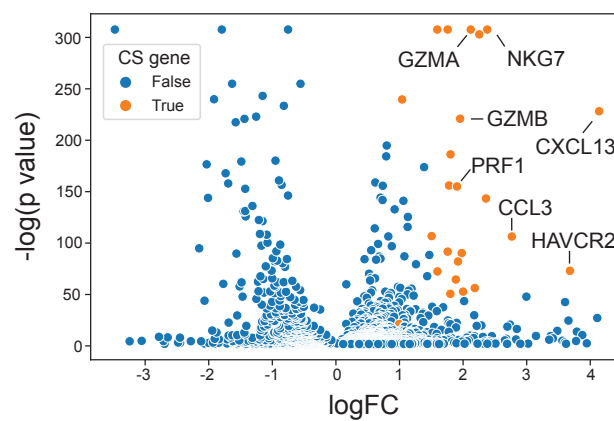
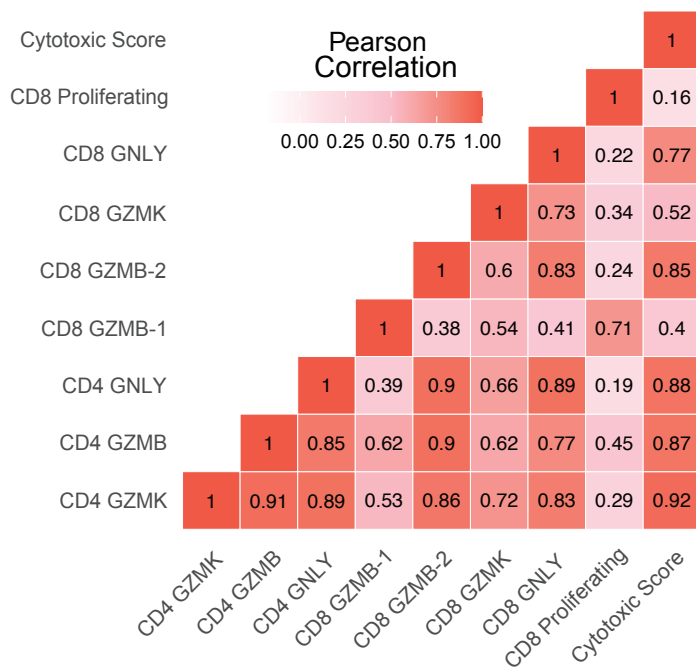
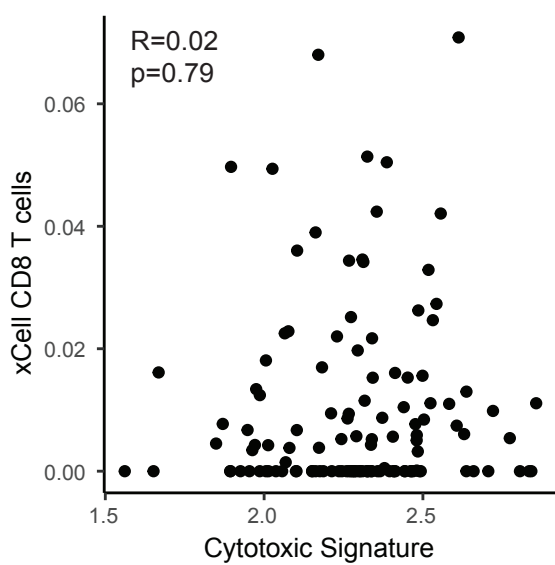
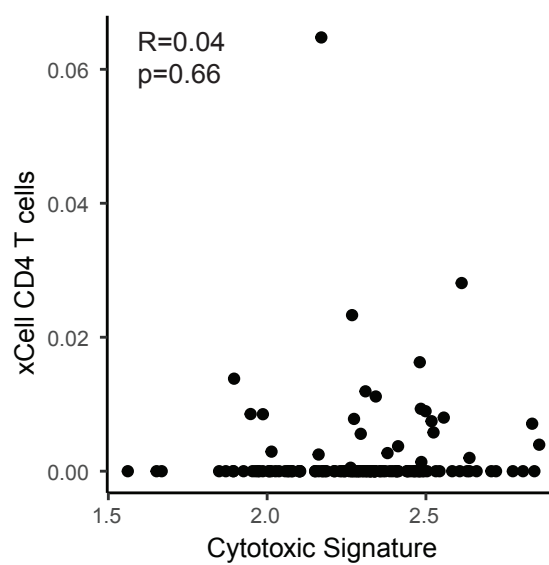


Supplementary Fig. 7. **Additional example of GZMB protein expression in CD4⁺ T cells.**

Representative multiplex immunofluorescent staining of DAPI (blue), CD4 (red), GZMB (green) in an NSCLC tumor (n = 2). Scale bar, 25 μ m.



Supplementary Fig. 8: **Association between T cell sub-populations and ICB outcomes.** Forest plots showing the univariate Cox proportional hazards models of TTP and expression of **a**, gene signatures for Leiden clusters, and **b**, xCell estimates of T cell infiltration in the Tempus NSCLC ICB cohort (n=123). Forest plots reflect results for mean-centered and scaled gene signatures and infiltration estimates and are not adjusted for multiple testing. The diamond represents the hazard ratio (HR) while the bars show the 95% confidence intervals. Source data are provided as a Source Data file.

a**b****c****d****e**

Supplementary Fig. 9: **Cytotoxic score (CS) is positively correlated with the presence of CD4⁺ and CD8⁺ cytotoxic T cell populations.** Volcano plot showing the log fold change and Benjamini-Hochberg adjusted $-\log_{10}$ p-values of genes in cytotoxic T cell clusters compared to non-cytotoxic T cell clusters in the **a**, CD8⁺ and **b**, CD4⁺ T cell compartments. Expression differences were assessed by Wilcoxon rank sum test. Genes in the CS gene list are colored orange and key genes are labelled. **c**, Heatmap showing the correlation between CS and the gene signatures for cytotoxic T cell populations identified from Leiden clustering in the Tempus NSCLC ICB cohort (n=123). Each box is labelled with the Pearson correlation coefficient. **d, e** Scatter plots showing the correlation between xCell estimates of CD8⁺ T cell and CD4⁺ T cell infiltration (CD8: R=0.02, p=0.79, CD4: R=0.04, p=0.66, Pearson correlation test, no adjustments for multiple comparisons) in the NSCLC ICB cohort (n=123). Source data are provided as a Source Data file.