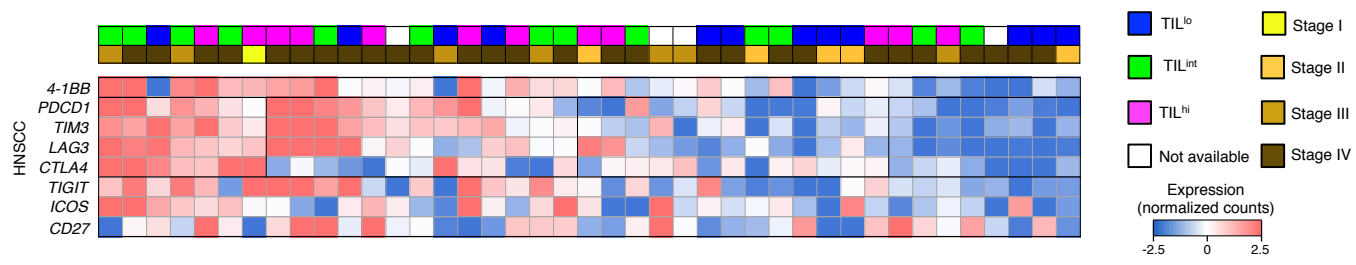


Supplementary Figure 1

RNA-Seq analysis of CD8⁺ TILs and N-TILs.

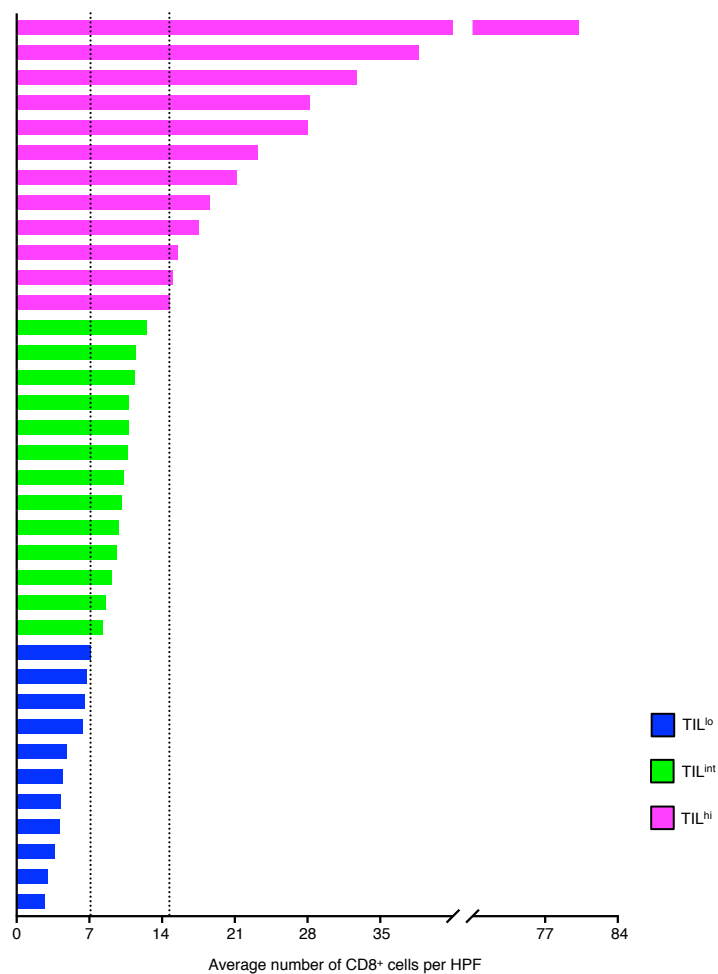
(a) Schematic representation of the tumor and cell types used for the study. HNSCC, head and neck squamous cell cancer; NSCLC, non-small cell lung cancer; TIL, tumor-infiltrating lymphocyte; N-TIL, non-tumor-infiltrating lymphocyte. (b) 'Minus-average' (MA) plots illustrate differentially expressed genes (red dots) for pairwise comparison of lung CD8⁺ N-TILs *versus* NSCLC CD8⁺ TILs (left), comparison of CD8⁺ TILs from NSCLC adenocarcinoma *versus* NSCLC squamous carcinoma (center), comparison of CD8⁺ TILs from HNSCC HPV-positive *versus* HPV-negative tumors (right) (DESeq2 analysis, Benjamini-Hochberg adjusted $P < 0.05$) (Online Methods and **Supplementary Table 3**). (c) Unsupervised hierarchical clustering of the transcriptomes of all CD8⁺ N-TILs (grey) and CD8⁺ TILs from NSCLC adenocarcinoma (red) and squamous carcinoma (pink), HPV-negative (light blue) and HPV-positive (dark blue) HNSCC, based on the expression of genes with the highest variance ($n = 1000$); each line represents an independent sample. (d) t-SNE plots of CD8⁺ T cell core transcriptome (symbol) from N-TILs and TILs of indicated cancer subtypes. Color scheme is identical to (c).



Supplementary Figure 2

Heterogeneity in the expression of immunotherapy target molecules.

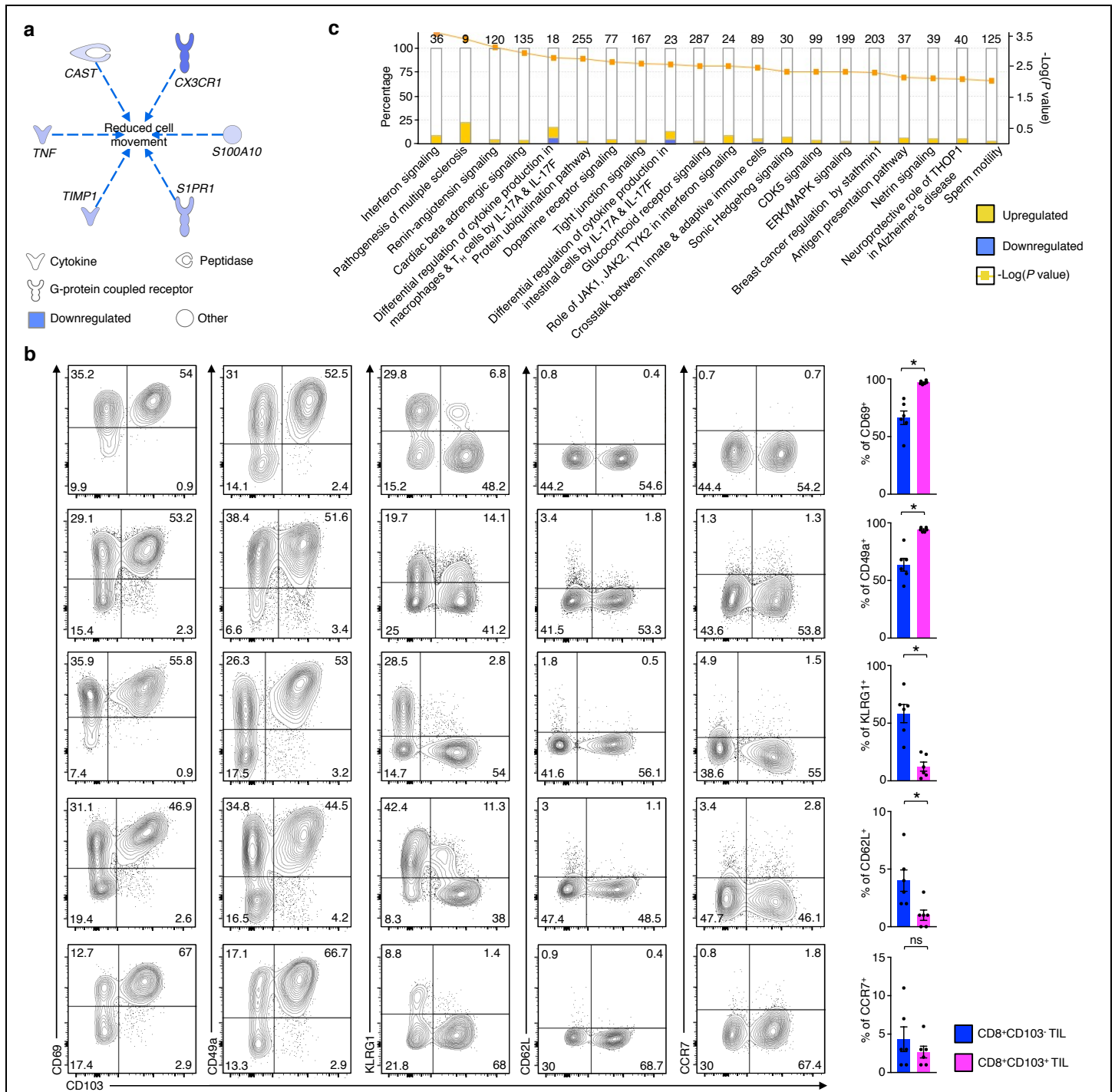
RNA-Seq analysis of HNSCC CD8⁺ TILs showing the row-wise normalized expression of indicated transcripts (rows) for each patient (column). Panel above heat map shows the CD8⁺ TIL density (top) and tumor stage (bottom) of each patient.



Supplementary Figure 4

TIL status of NSCLC patients.

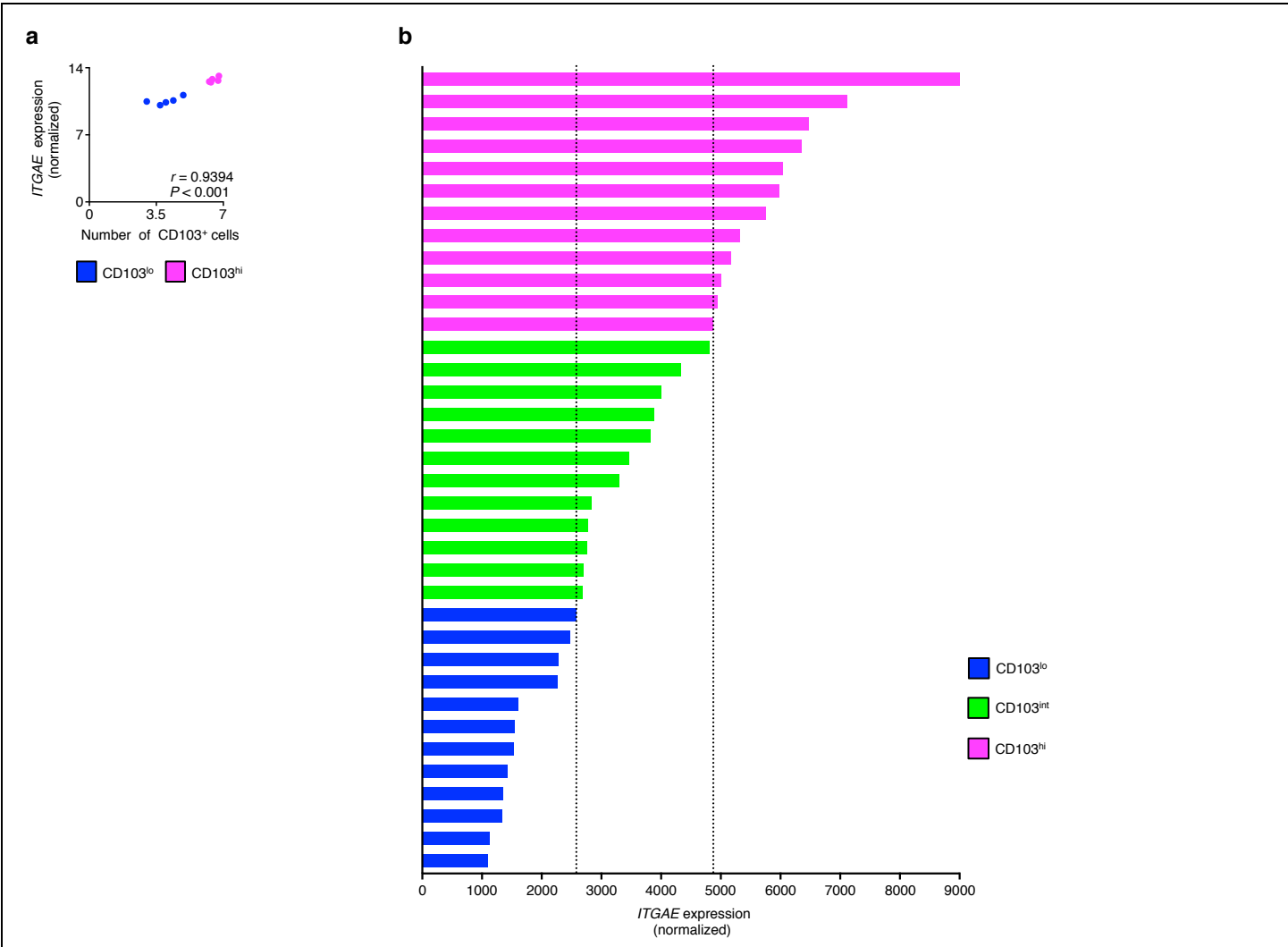
Graph shows the average number of CD8 α ⁺ cells per high power field (HPF) in tumor samples from each NSCLC patient (Online Methods).



Supplementary Figure 5

Pathways enriched in CD8⁺ TILs from NSCLC TIL^{hi} tumors and phenotype of CD8⁺CD103⁺ TILs.

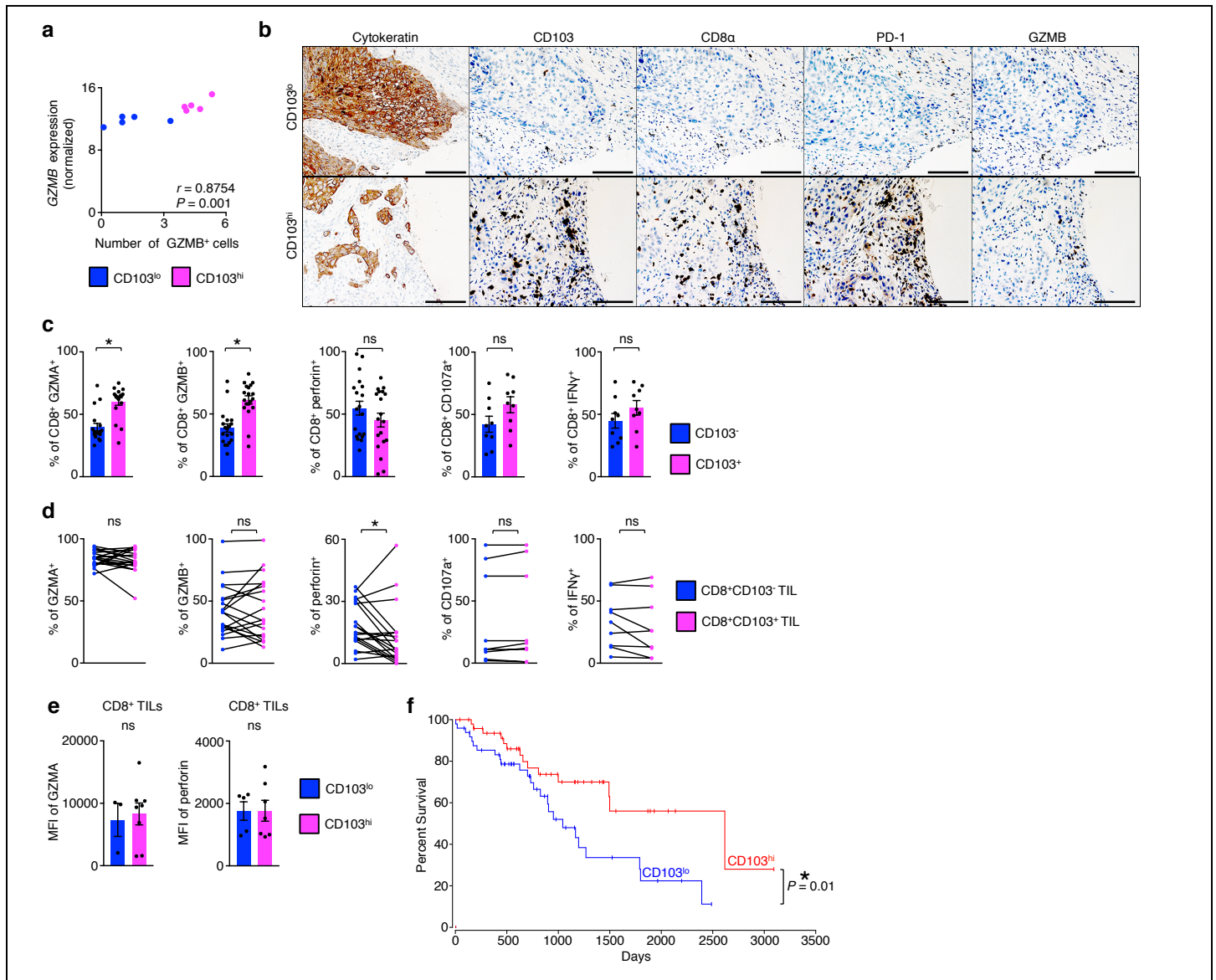
(a) Ingenuity Pathway Analysis shows downregulated transcripts that are associated with tissue retention in NSCLC TIL^{hi} tumors. (b) Contour plots show expression of CD103, CD69, CD49a, KLRG1, CD62L and CCR7 in live and singlet-gated CD45⁺CD3⁺CD8⁺ T cells obtained from NSCLC TILs ($n = 6$); graphs show the percentage of CD103⁺CD8⁺ and CD103⁻CD8⁺ TILs that express the indicated surface marker (error bars are mean \pm SEM); * $P = 0.0025$ (CD69), $P = 0.0025$ (CD49a), $P = 0.0016$ (KLRG1), $P = 0.0021$ (CD62L) (paired Student's two-tailed t -test); each dot represents data from a patient. (c) Ingenuity Pathway Analysis of canonical pathways enriched in NSCLC TIL^{hi} tumors, P values calculated by Fisher Exact Test (Online Methods).



Supplementary Figure 6

CD103 (*ITGAE*) status of CD8⁺ TILs.

(a) Correlation of *ITGAE* transcript expression (log₂ normalized counts) in NSCLC CD8⁺ TIL with the average number of tumor-infiltrating CD103⁺ cells quantified by immunohistochemistry ($n = 10$). r value indicates the Spearman correlation coefficient. (b) Graph shows classification of tumors based on *ITGAE* (*CD103*) transcript expression (normalized counts) in CD8⁺ TILs from each NSCLC patient (Online Methods).



Supplementary Figure 7

CD103 density predicts survival in lung cancer.

a) Correlation of *GZMB* transcript expression (log₂ normalized counts) in NSCLC CD8⁺ TIL with the average number of tumor-infiltrating GZMB⁺ cells quantified by immunohistochemistry ($n = 10$). r value indicates the Spearman correlation coefficient. **(b)** Immunohistochemistry staining for cytokeratin, CD103, CD8α, PD-1 and GZMB in CD103^{lo} (top panel) and CD103^{hi} NSCLC tumors (bottom panel); scale bar is 100 μm. **(c)** Graphs show the percentage of CD8⁺TILs expressing the indicated molecules that are either CD103⁺ or CD103⁻ ($n = 9-19$, error bars are mean \pm SEM); * $P = 0.003$ for GZMA and $P = 0.0029$ for GZMB (paired Student's two-tailed t -test). **(d)** Graphs shows the percentage of CD103⁺CD8⁺ and CD103⁻CD8⁺ TILs ($n = 9-19$) that express the indicated molecules; * $P = 0.047$ (paired Student's two-tailed t -test). **(e)** Bar graphs show mean expression (error bars are mean \pm SEM of geometric MFI values) of granzyme A and perforin in CD8⁺ TILs in CD103^{lo} ($n = 3-5$) versus CD103^{hi} ($n = 7-8$) tumors; P value insignificant by Mann-Whitney test. **(f)** Kaplan–Meier curves for lung cancer mortality in patients from the TCGA lung adenocarcinoma dataset classified based on expression of *ITGAE* transcripts into CD103^{hi} (upper 10th percentile, $n = 49$) and CD103^{lo} (lower 10th percentile, $n = 49$); (log-rank test, $P = 0.0148$; Online Methods).