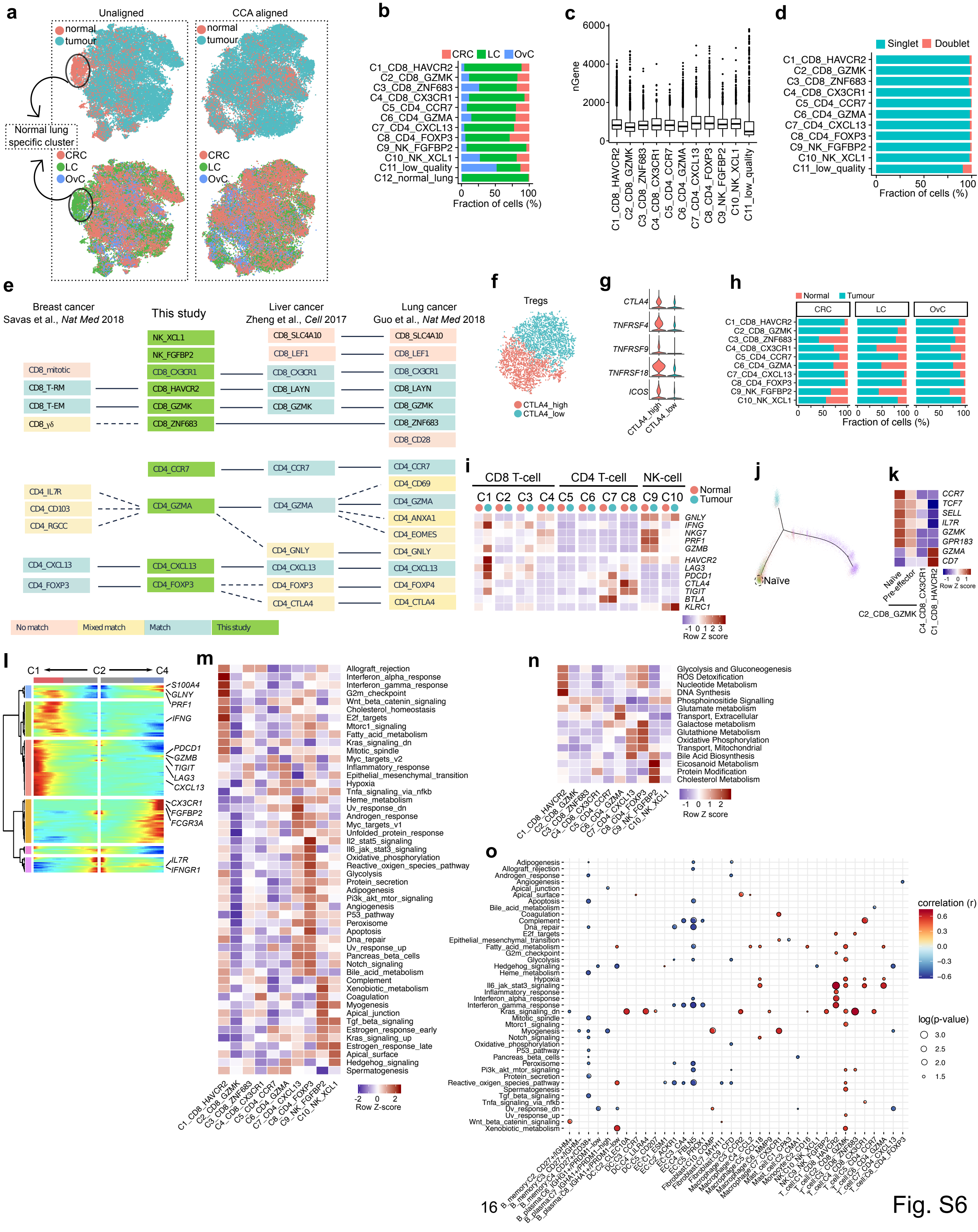
a
Unaligned
CCA aligned
normal
tumour
Normal lung specific cluster
CRC
LC
OvC
b
CRC LC OvC
C1_CD8_HAVCR2
C2_CD8_GZMK
C3_CD8_ZNF683
C4_CD8_CX3CR1
C5_CD4_CCR7
C6_CD4_GZMA
C7_CD4_CXCL13
C8_CD4_FOXP3
C9_NK_FGFBP2
C10_NK_XCL1
C11_low_quality
C12_normal_lung
Fraction of cells (%)
c
nGene
6000
4000
2000
0
d
Singlet
Doublet
e
Breast cancer Savas et al., Nat Med 2018
This study
Liver cancer Zheng et al., Cell 2017
Lung cancer Guo et al., Nat Med 2018
CD8_mitotic
CD8_T-RM
CD8_T-EM
CD8_γδ
NK_XCL1
NK_FGFBP2
CD8_CX3CR1
CD8_HAVCR2
CD8_GZMK
CD8_ZNF683
CD8_SLC4A10
CD8_LEF1
CD8_LAYN
CD8_CD28
CD4_IL7R
CD4_CD103
CD4_RGCC
CD4_CXCL13
CD4_FOXP3
CD4_CCR7
CD4_GZMA
CD4_GNLY
CD4_CTLA4
CD4_CD69
CD4_ANXA1
CD4_EOMES
CD4_FOXP4
No match
Mixed match
Match
This study
f
Tregs
CTLA4_high
CTLA4_low
g
CTLA4
TNFRSF4
TNFRSF9
TNFRSF18
ICOS
h
Normal
Tumour
CRC
LC
OvC
Fraction of cells (%)
i
CD8 T-cell
CD4 T-cell
NK-cell
C1 C2 C3 C4 C5 C6 C7 C8 C9 C10
Normal
Tumour
GNLY
IFNG
NKG7
PRF1
GZMB
HAVCR2
LAG3
PDCD1
CTLA4
TIGIT
BTLA
KLRC1
Row Z score
j
Naïve
k
CCR7
TCF7
SELL
IL7R
GZMK
GPR183
GZMA
CD7
Naïve
Pre-effector
C2_CD8_GZMK
C4_CD8_CX3CR1
C1_CD8_HAVCR2
Row Z score
l
C1
C2
C4
S100A4
GLNY
PRF1
IFNG
PDCD1
GZMB
TIGIT
LAG3
CXCL13
CX3CR1
FGFBP2
FCGR3A
IL7R
IFNGR1
m
Allograft_rejection
Interferon_alpha_response
Interferon_gamma_response
G2m_checkpoint
Wnt_beta_catenin_signaling
Cholesterol_homeostasis
E2f_targets
Mtorc1_signaling
Fatty_acid_metabolism
Kras_signaling_dn
Mitotic_spindle
Myc_targets_v2
Inflammatory_response
Epithelial_mesenchymal_transition
Hypoxia
Tnfa_signaling_via_nfkb
Heme_metabolism
Uv_response_dn
Androgen_response
Myc_targets_v1
Unfolded_protein_response
Il2_stat5_signaling
Il6_jak_stat3_signaling
Oxidative_phosphorylation
Reactive_oxigen_species_pathway
Glycolysis
Protein_secretion
Adipogenesis
Pi3k_akt_mtor_signaling
Angiogenesis
P53_pathway
Peroxisome
Apoptosis
Dna_repair
Uv_response_up
Pancreas_beta_cells
Notch_signaling
Bile_acid_metabolism
Complement
Xenobiotic_metabolism
Coagulation
Myogenesis
Apical_junction
Tgf_beta_signaling
Estrogen_response_early
Kras_signaling_up
Estrogen_response_late
Apical_surface
Hedgehog_signaling
Spermatogenesis
Row Z-score
n
Glycolysis and Gluconeogenesis
ROS Detoxification
Nucleotide Metabolism
DNA Synthesis
Phosphoinositide Signalling
Glutamate metabolism
Transport, Extracellular
Galactose metabolism
Glutathione Metabolism
Oxidative Phosphorylation
Transport, Mitochondrial
Bile Acid Biosynthesis
Eicosanoid Metabolism
Protein Modification
Cholesterol Metabolism
Row Z score
o
correlation (r)
log(p-value)
B_memory:C2_CD27+/IGHM+
B_memory:C3_CD27+/IGHM-
B_memory:C4_CD27-/CD38+
B_plasma:C6_IGHG1+/PRDM1-low
B_plasma:C7_IGHA1+/PRDM1-high
B_plasma:C8_IGHA1+/PRDM1-low
DC:C2_CLEC10A
DC:C3_CCR7
DC:C4_LILRA4
DC:C5_CD207
EC:C1_ESM1
EC:C2_ACKR1
EC:C3_CA4
EC:C4_FBLN5
EC:C5_PROX1
Fibroblast:C10_COMP
Fibroblast:C9_MYH11
Fibroblast:C9_CFD
Macrophage:C2_CCR2
Macrophage:C4_CCL2
Macrophage:C5_CCL18
Macrophage:C6_MMP9
Macrophage:C7_CX3CR1
Mast_cell:C1_CMA1
Mast_cell:C2_CPA3
Monocyte:C2_CD16
NK:C10_NK_XCL1
NK:C9_NK_FGFBP2
T_cell:C1_CD8_HAVCR2
T_cell:C2_CD8_GZMK
T_cell:C3_CD8_ZNF683
T_cell:C4_CD8_CX3CR1
T_cell:C5_CD4_CCR7
T_cell:C6_CD4_GZMA
T_cell:C7_CD4_CXCL13
T_cell:C8_CD4_FOXP3

Fig. S6

Fig. S6. Profiling of T- and NK-cells

a t-SNE plots showing unaligned (left panels) or CCA-aligned (right panels) clusters colour-coded by sample origin (upper panels), or cancer type (lower panels). Normal lung specific cluster are indicated by circles and arrows. **b** Fraction of cells for T-/NK-cell subclusters (unaligned) from different cancer types. **c** Number of genes detected from different T-cell subclusters (CCA aligned). **d** Fraction of singlet and doublet cells for T-/NK- cell subclusters predicted by DoubletFinder. **e** Comparison of T-/NK-cell subclusters defined in this pan-cancer study and previous breast, liver and lung cancer studies [52,54,55]. **f** t-SNE plot showing subclusters within CD4_FOXP3 Tregs after further subclustering. **g** Violin plot showing marker gene expressions between two Tregs subclusters. **h** Fraction of cells derived from normal (red) or tumour tissue (green) for each T-/NK-cell subcluster in each cancer type. **i** Heatmap of differentially expressed genes for T-/NK-cell subclusters with different sample origins. **j** Trajectories of CD8$^+$ T-cell differentiation. Naïve cells are located at the root of the C2_CD8_GMZK branch (dash line circled). **k** Heatmap of naïve marker gene expression. The naïve subcluster within the C2_CD8_GZMK cluster, as circled in **j**, expresses a number of naïve markers, including *CCR7* and *TCF7*. **l** Heatmap of expression dynamics of differentially expressed genes along the pseudotime of the 2 CD8$^+$ lineages (from C2_CD8_GZMK to C1_CD8_HAVCR1 or C4_CD8_CX3CR1). **m** Differences in cancer hallmark pathway activities scored per cell by AUCell for T-/NK-cell subclusters. **n** Metabolic activity of each T-/NK-cell pheenotype scored by AUCell. **o** Heatmap showing positive (red) and negative (blue) correlations between cancer hallmark pathway activation in cancer cells (Y-axis) and the prevalence of the stromal cell phenotypes.