**Extended Data Fig. 1 | Technical reproducibility and quality metrics of Xenium profiling.**
**a**, Pairwise Pearson correlation of gene expression between 12 normal lung reference regions processed across 12 independent runs. Diagonal cells are masked. **b**, Pearson correlation between Normal_Lung_01 and Normal_Lung_02, computed from log1p-transformed per-gene mean counts (r = 0.97). **c**, Xenium QC metrics across samples shown as differences from the across-sample mean, including high-quality transcripts per μm², detected cells per μm², median genes per cell, and median transcripts per cell. **d**, High-quality transcripts per 100 μm², **e**, cells detected per 100 μm² and **f**, median transcripts per cell across all samples, grouped by tissue type (n = 365 samples). Densities were calculated relative to the analysed region area. Tissues are ordered by median transcripts per cell, with normal lung shown beneath lung tumour samples. Boxes show the median and interquartile range, and whiskers extend to 1.5× the interquartile range; each dot represents one sample.

**Extended Data Fig. 2 | Benchmarking ASTRA2K against v1 and Xenium 5K panels.**
**a–c**, Comparison of the Xenium v1, ASTRA2K and Xenium 5K panels in six samples profiled with each panel, computed on the 183 genes shared by the three panels. **a**, Kernel density estimates of transcripts per cell. **b**, Kernel density estimates of detected genes per cell. **c**, Median transcripts per cell for each sample (n = 6 per panel). P values, two-sided paired Wilcoxon signed-rank tests with Benjamini–Hochberg correction across the three pairwise comparisons. **d**, Gene-wise Pearson correlations across platforms, computed between ASTRA2K and v1/5K for matched samples (range across samples: r = 0.79–0.93). **e**, Distribution of per-cell *YAP1* counts with the 95th percentile indicated (jichihgnec27). **f**, Relative *YAP1* detection in Xenium v1 and ASTRA2K normalised to Xenium 5K (= 1), quantified as fold change in per-sample mean per-cell counts relative to the matched 5K sample. **g**, Relative detection of the 183 genes shared across the three panels in Xenium v1 and ASTRA2K, normalised to Xenium 5K (= 1). **h**, Representative Xenium Explorer snapshots of *PECAM1* across platforms; *PECAM1* is probe-boosted in ASTRA2K (more than six probes). **i**, Distribution of per-cell *PECAM1* counts with the 95th percentile indicated. **j**, Relative *PECAM1* detection in Xenium v1 and ASTRA2K normalised to Xenium 5K (= 1), quantified as fold change in per-sample mean per-cell counts relative to the matched 5K sample. **k**, Relative detection of the probe-boosted genes in ASTRA2K, normalised to Xenium 5K (= 1).

**Extended Data Fig. 3 | Cell-type composition and marker-defined cell states across subtypes.**
**a**, Proportions of major cell types in each subtype, ordered by cell-type frequency. Bars show the mean ± s.e.m. across samples. **b–f,** Percentage of marker-positive cells within each cell type: **b,** *MKI67* in epithelial cells; **c,** *POSTN* in fibroblasts; **d,** *PLVAP* in endothelial cells; **e,** *CD8A* in T cells; **f,** *CD163* in macrophages. Cells were considered positive if they contained ≥1 transcript of the marker. Each dot represents one sample. P values, two-sided Kruskal–Wallis tests across subtypes, Benjamini–Hochberg-adjusted.

**Extended Data Fig. 4 | ASTRA-CYTE annotation hierarchy, marker-based validation and hepatocyte contamination filtering.**

**a,** Schematic of the hierarchical ASTRA-CYTE model. Cells are first assigned to a broad class (epithelial, fibroblast, endothelial, lymphoid or myeloid) or left unassigned. Lymphoid cells are then resolved into T cells, NK cells, B cells and plasma cells, and myeloid cells into monocytes, macrophages and neutrophils. **b,** Expression of hepatocyte markers (*ALB*, *APOA1*, *CYP2E1* and *TF*) and lineage markers for fibroblasts, macrophages, endothelial cells and T cells in the reference training data. Cells retained after quality control ('kept') are compared with cells excluded for hepatocyte transcript contamination ('removed'). Removed cells show high hepatocyte transcript levels, whereas kept cells retain lineage-specific profiles with little hepatocyte signal. In **b**, **c** and **e**, dot size shows the fraction of cells expressing each gene, and dot colour shows mean expression, scaled from 0 to 1 per gene. **c,** Expression of canonical marker genes (columns, grouped by the lineage they define) across annotated cell types (rows). **d**, Internal cross-validation of ASTRA-CYTE on the training data. The training set comprised one sample from each of ten tissue types, and each fold of five-fold cross-validation held out two tissue types, so no held-out tissue type contributed cells to training. Bars show macro-averaged F1 scores for each held-out tissue type for the broad-class (level 1), lymphoid (level 2a) and myeloid (level 2b) classifiers. Level 2 scores were computed on cells with a correct level 1 assignment. The y axis starts at 0.5. **e,** Comparison of manual annotations (M) with ASTRA-CYTE predictions (P) in liver. Rows labelled 'agree' show cells where the two annotations match. Rows labelled 'X (M) → Epithelial (P)' show cells manually annotated as cell type X but predicted as epithelial. These discordant cells express hepatocyte and epithelial markers (*ALB*, *APOA1*, *CYP2E1*, *TF*, *GPC3* and *AFP*) and show little expression of lineage markers matching their manual label, suggesting that the manual labels reflect hepatocyte transcript contamination.

**Extended Data Fig. 5 | Cancer hallmark profiles by tissue and spatial illustration of invasion-associated transcripts.**

**a,** Radar charts showing the scores of 13 cancer hallmarks for each of 11 tissue types. Lymph node and kidney samples were excluded. Hallmark scores were calculated for each epithelial cell, averaged within each sample, then within each subtype, z-scored across subtypes for each hallmark and centred per subtype. Subtypes from the same tissue type were then averaged, weighted by sample number, and z-scored per hallmark to give tissue-level profiles. The dashed circle indicates $z = 0$. **b,** Representative images of the 'activating invasion and metastasis' hallmark, showing *MMP7* (purple) and *SPP1* (orange) transcripts overlaid on epithelial cells in pancreatic cancer (P24_2604_PK) and bladder cancer (H19_01410). Scale bar, 50 μm.

**Extended Data Fig. 6 | Robustness of cellular neighbourhood identification.**

**a**, Clustering stability (mean pairwise ARI across 10 initialisations; top) and silhouette width (bottom) for 4–14 clusters. Mean ± s.d.; arrow, eight clusters used throughout. **b,** Correspondence between the 10-cluster and 8-cluster solutions. Each row shows the fraction of cells in a 10-cluster

CN. The eight original CNs are largely preserved when the cluster number increases, and the additional clusters ('mixed') draw cells from several CNs rather than resolving new biological CNs. **c,** Sensitivity of CN identification to neighbourhood size. Composition vectors were recomputed using 25, 50 and 200 nearest neighbours and re-clustered with k = 8, and the resulting clusters were matched to the reference CNs. Top, ARI between the resulting labels and the reference labels (100 nearest neighbours). Bottom, fraction of cells in each CN assigned to the matched CN. **d,** Overview of HCC samples from Australia (n = 63) and Japan (n = 19). **e,** Reproducibility of CN composition across cohorts. Each point shows the mean proportion of one cell type within one CN, compared between the Australian and Japanese cohorts. Pearson r = 0.984; the dashed line indicates y = x. **f,** Proportion of each CN in individual HCC samples, by cohort. Box plots show the median and interquartile range, with whiskers extending to 1.5× the interquartile range. Points represent individual samples.

**Extended Data Fig. 7 | Conservation of cellular-neighbourhood-induced expression shifts across cell types and tissue types.**
**a,** Schematic of the CN-shift calculation and leave-one-tissue-type-out CN prediction. For each cell type in each tissue type, the CN-induced expression shift was calculated as the difference between the pseudobulk profile of that cell type in a given CN and the mean of its profiles across all CNs in the same sample. To test whether shifts generalise across tissue types, a classifier was trained on shift profiles from all but one tissue type and used to predict the CN of origin in the held-out tissue type. **b,** Pairwise Pearson correlations of CN-induced expression shifts between CN–tissue-type pairs, shown for each cell type. Cell types are ordered by conservation score, defined as the median correlation between the same CN in different tissue types minus the median correlation between different CNs. **c,** Median correlation of expression shifts between different tissue types for the same CN (filled) and for different CNs (open), by cell type. $q < 0.01$ for all cell types; one-sided permutation test (200 permutations of CN labels within samples) with Benjamini–Hochberg correction. **d,** CN-induced expression shifts in macrophages within the T-cell CN (CN3) and the stromal CN (CN8) across tissue types. The heatmap shows the union of the top 20 genes per CN, ranked by absolute median shift across tissue types. **e,** Confusion matrix of CN predictions in held-out tissue types, pooled across cell types and row-normalised to the true CN. Balanced accuracy = 66%; chance = 12.5%.

**Extended Data Fig. 8 | Selection, robustness and characterisation of Spatial Theratopes.**
**a**, Inertia, silhouette width and subsampling stability (mean adjusted Rand index (ARI) across 50 random 80% subsamples; shading, ± s.d.) for k-means clustering of target-pair scores from 340 tumours with k = 2–10. Red points mark the selected solution (k = 4). **b**, Correspondence between Spatial Theratopes (STs) and the k = 3 and k = 6 solutions. Numbers are tumour counts, and shading shows the fraction of each ST assigned to each cluster. **c**, Stability of each ST, measured as the mean Jaccard index across 100 bootstrap resamples (± s.d.). Dashed lines mark the conventional thresholds for stable (0.75) and unstable (0.5) clusters. **d**, Agreement (ARI) between

k-means STs and clusters from Ward hierarchical clustering or a Gaussian mixture model, each with four clusters. **e**, Difference in the percentage of tumours with a high target-pair score (> 0.5) between each ST and all other tumours. Black outlines mark pairs with Benjamini–Hochberg-adjusted FDR < 0.05 (two-sided Mann–Whitney U test) and a difference of at least 10 percentage points. **f**, One-vs-rest area under the ROC curve (AUC) of random forest classifiers predicting each ST from CN proportions, tissue, histological subtype or their combinations (n = 340 tumours; mean ± s.d. across 10 repeats of five-fold cross-validation). The dashed line indicates chance (AUC = 0.5). **g**,**h**, Principal component analysis of target-pair scores, min–max scaled within each tumour, coloured by **g**, ST and **h**, tissue.

**Extended Data Fig. 9 | Characterisation of the HCC cohort treated with atezolizumab plus bevacizumab.**

**a,** Cell-type composition and **b**, cellular neighbourhood (CN) composition of each tumour in the Gifu HCC cohort (n = 22 patients). **c,** Cell-type composition of each CN in the Gifu cohort, after label transfer, compared with the discovery cohort (n = 365 samples). Each point represents the mean proportion of one cell type within one CN. Grey line, y = x; r, Pearson correlation coefficient. **d**, Projection of Gifu HCC samples (diamonds) onto the principal component analysis (PCA) space of target-pair scores from the discovery cohort (dots). Large circles mark the centroids of ST1–ST4 and lines connect each Gifu sample to its assigned centroid (n = 25; 23 assigned to ST3, one to ST2 and one to ST4). **e**, Target-pair scores for each Gifu HCC sample (n = 25) with the left bar showing ST assignment. **f**, Univariable Cox proportional hazards models for progression-free survival (PFS) by CN, with each CN modelled as present (>0)/absent or split at the median. Points, hazard ratios (HRs); error bars, 95% confidence intervals (CIs); *P*, two-sided Wald test; *q*, Benjamini–Hochberg-adjusted *P* across evaluable CNs. CNs with fewer than five patients or five events in either group were not tested. **g,** VEGF-A–PD-L1 pair score and **h**, CN8 proportion by best response according to mRECIST (PD, progressive disease; SD, stable disease; PR, partial response; CR, complete response). ρ and *P*, two-sided Spearman's rank correlation, with response treated as ordinal (PD < SD < PR < CR; positive ρ indicates higher values with better response). Box plots show the median and interquartile range, with whiskers extending to 1.5× the interquartile range; each point represents one patient.