Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A pan-cancer spatial atlas from Asia Pacific reveals recurrent spatial theratopes

Ankur Sharma

ankur.sharma@garvan.org.au

Garvan Institute of Medical Research https://orcid.org/0000-0002-6862-136X

Lucy Chhuo

Garvan Institute of Medical Research

Yoshiki Otsuka

The University of Tokyo

Hitomi Kawai

University of Tsukuba

Merrin Eapen

Garvan Institute of Medical Research

Ryutaro Sekimoto

University of Tsukuba

Kei Asayama

University of Tsukuba

Ryota Matsuoka

University of Tsukuba

Abhishek Singh

Garvan Institute of Medical Research

Toshiro Niki

Jichi Medical University

Noriyoshi Fukushima

Jichi Medical University

Kentaro Inamura

Jichi Medical University

Skye van Esch

Garvan Institute of Medical Research https://orcid.org/0009-0008-7999-4218

Elgene Lim

The Kinghorn Cancer Centre and Cancer Research Theme, Garvan Institute of Medical Research, St Vincent's Clinical School, Faculty of Medicine, UNSW Sydney, St Vincent's Hospital
https://orcid.org/0000-0001-8065-8838

Egor Revkov

Garvan Institute of Medical Research
Dana Idais
Garvan Institute of Medical Research
Shiro Oka
Hiroshima University
Atsushi Ono
Hiroshima University
Makoto Yamagishi
The University of Tokyo https://orcid.org/0000-0002-8142-4686
Thang Le
National Taiwan University
Wei-Chou Lin
National Taiwan University
Lin-Hung Huang
National Taiwan University
Yi-Chia Chiu
National Taiwan University
Atsushi Kaneda
Chiba University Graduate School of Medicine https://orcid.org/0000-0002-6980-5515
toyoyuki Hanazawa
Chiba University
Takuya Nakagawa
Chiba University
Annabel Faulkner
The University of Sydney
Jordan Conway
The University of Sydney
Georgina Long
Melanoma Institute Australia, The University of Sydney, and Royal North Shore and Mater Hospitals
Ines Pires de Silva
Melanoma Institute Australia, The University of Sydney https://orcid.org/0000-0003-3540-8906
Helen McGuire
The University of Sydney https://orcid.org/0000-0003-2047-6543
Hidewaki Nakagawa
RIKEN Center for Integrative Medical Sciences https://orcid.org/0000-0003-1807-772X
Ruby Huang
School of Medicine, College of Medicine, National Taiwan University https://orcid.org/0000-0001-6376-3185
Takayuki Ueno

Japanese Foundation for Cancer Research
Junki Koike
Japanese Foundation for Cancer Research
Rie Horii
Japanese Foundation for Cancer Research
Satoi Nagasawa
Japanese Foundation for Cancer Research
Masaki Shiota
Kyushu University https://orcid.org/0000-0002-3306-4858
Sotaro Chikamatsu
Kyushu University
Shota Sasagawa
RIKEN Center for Integrative Medical Sciences
Yutaka Hatanaka
Hokkaido University
Masakazu Fujii
Hokkaido University Hospital
Masaya Onishi
Gifu University
Masahito Shimizu
Gifu University
Liang Qiao
The University of Sydney
Ayako Suzuki
University of Tokyo https://orcid.org/0000-0001-9917-2280
Daisuke Matsubara
University of Tsukuba
Jacob George
Westmead Hospital https://orcid.org/0000-0002-8421-5476
ASTRA Consortium
Garvan Institute of Medical Research
Yutaka Suzuki
The University of Tokyo https://orcid.org/0000-0003-4852-1879
Yukinori Ozaki
Japanese Foundation for Cancer Research

Biological Sciences - Article

Keywords:

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11641830/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

A pan-cancer spatial atlas from Asia Pacific reveals recurrent Spatial Theratopes

Lucy Chhuo[1,#], Yoshiki Otsuka[2,#], Hitomi Kawai[3], Merrin Mary Eapen[1,4], Ryutaro Sekimoto[3], Kei Asayama[3], Ryota Matsuoka[3], Abhishek Singh[1], Toshiro Niki[5], Noriyoshi Fukushima[5], Kentaro Inamura[6,7], Skye van Esch[1], Elgene Lim[4,8], Egor Revkov[1,4], Dana Idas[1,4], Shiro Oka[9], Atsushi Ono[9], Makoto Yamagishi[2], Thang Truong Le[10], Wei-Chou Lin[11], Lin-Hung Wei[12], Yi-Chia Chiu[13], Atsushi Kaneda[14], Toyoyuki Hanazawa[15], Takuya Nakagawa[14,15], Annabel Faulkner[16], Jordan Conway[17,18,19], Georgina Long[17,18,19,20], Ines Pires da Silva[17,21], Helen McGuire[16], Hidewaki Nakagawa[22], Ruby Yun-Ju Huang[13], Yukinori Ozaki[23], Takayuki Ueno[23], Junki Koike[24], Rie Horii[24], Satoi Nagasawa[2], Masaki Shiota[25], Sotaro Chikamatsu[2,25], Shota Sasagawa[22], Yutaka Hatanaka[26], Masakazu Fujii[27], Masaya Onishi[2,28], Masahito Shimizu[28], Liang Qiao[29], Ayako Suzuki[2,30], Daisuke Matsubara[3], Jacob George[29], ASTRA Consortium[30], Yutaka Suzuki[2,*] & Ankur Sharma[1,4,31,*]

[1]Translational Genomics Program, Garvan Institute of Medical Research, Sydney, New South Wales, Australia
[2]Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Japan
[3]Department of Diagnostic Pathology, Institute of Medicine, University of Tsukuba, Ibaraki, Japan
[4]School of Clinical Medicine, Faculty of Medicine and Health, University of New South Wales, Sydney, New South Wales, Australia
[5]Division of Comprehensive Pathology, Department of Pathology, Jichi Medical University, Shimotsuke, Japan
[6]Division of Tumor Pathology, Department of Pathology, Jichi Medical University, Shimotsuke, Japan
[7] Division of Pathology, The Cancer Institute, Japanese Foundation for Cancer Research, Tokyo, Japan
[8]Cancer Ecosystem Program, Garvan Institute of Medical Research, Sydney, New South Wales, Australia
[9]Department of Gastroenterology, Graduate School of Biomedical and Health Sciences, Hiroshima University, Hiroshima 734-8553, Japan
[10]Genome and Systems Biology Degree Program, National Taiwan University and Academia Sinica, Taipei, Taiwan
[11]Department of Pathology, National Taiwan University Hospital, Taipei, Taiwan
[12]Department of Obstetrics & Gynecology, National Taiwan University Hospital, Taipei, Taiwan
[13]School of Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan.
[14]Health and Disease Omics Center, Chiba University
[15]Department of Otorhinolaryngology-Head and Neck Surgery, Graduate School of Medicine, Chiba University
[16]School of Medical Sciences, Faculty of Medicine and Health, University of Sydney, Sydney, Australia
[17]Melanoma Institute Australia, The University of Sydney, Sydney, New South Wales, Australia

[18]Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, Australia
[19]Charles Perkins Centre, The University of Sydney, Sydney, New South Wales, Australia
[20]Royal North Shore and Mater Hospitals, Sydney, New South Wales, Australia.
[21]Champalimaud Foundation, Lisbon, Portugal
[22]Laboratory for Cancer Genomics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.
[23]Breast Oncology Center Department of Advanced Medical Development, Cancer Institute Hospital, Japanese Foundation for Cancer Research
[24]Department of Pathology, St. Marianna University School of Medicine
[25]Department of Urology, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka, 812-8582, Japan
[26]Center for Development of Advanced Diagnostics, Hokkaido University Hospital, Sapporo, Hokkaido, Japan
[27]Department of Gastroenterological Surgery II, Faculty of Medicine, Hokkaido University, Sapporo, Hokkaido, Japan
[28]Department of Gastroenterology/Internal Medicine, Graduate School of Medicine, Gifu University, Gifu, Japan
[29]Storr Liver Centre, The Westmead Institute for Medical Research and Westmead
[30]A full list of ASTRA Consortium members and their affiliations appears at the end of the manuscript
[31]KK Research Centre, KK Women's and Children's Hospital, Singapore, Singapore
#These authors contributed equally: Lucy Chhuo (L.C.), Yoshiki Otsuka (Y.O.)
*To whom correspondence should be addressed;
Yutaka Suzuki (Y.S.): ysuzuki@edu.k.u-tokyo.ac.jp
Ankur Sharma (A.S.): ankur.sharma@garvan.org.au

Summary

Here we generated a pan-cancer spatial atlas of 400 tumours across 13 tissue types and 26 subtypes from the Asia-Pacific, a region under-represented in existing pan-cancer datasets. In a multimodal cohort we integrated single-cell spatial transcriptomics (1,999 genes) with more than 70 proteins measured in the same cells as built-in validation, alongside clinical data including ancestry, age, sex and survival. ASTRA1.0 revealed that each cancer subtype has distinct expression profiles and cellular compositions. Despite this heterogeneity, we identified eight recurrent spatial cellular neighbourhoods shared across tissue types, with macrophages as key organisers. We further developed a framework to quantify the co-localisation of FDA-approved drug targets with surrounding cell types. This framework classified tumours into four Spatial Theratope subtypes, each defined by a distinct signature of co-localised targets. The VEGF-A–PD-L1 pair, targeted by bevacizumab and atezolizumab, recurred across tumour types and neighbourhoods. Spatial Theratope subtypes and neighbourhood composition were each associated with progression-free survival (PFS) in the discovery cohort, and their combination was also associated with PFS in an independent cohort of 25 patients with hepatocellular carcinoma treated with atezolizumab plus bevacizumab. The ASTRA atlas provides a community resource for nominating rational, spatially informed combination therapies across cancers.

Main

The complex interaction between various cellular neighbourhoods in tumour tissue determines cancer initiation, progression and response to therapies[1–3]. Spatial profiling studies have shown that the organisation of these cells is associated with clinical outcome, independent of composition[4,5]. Beyond their prognostic value, these multicellular neighbourhoods rather than malignant cells in isolation are increasingly recognised as therapeutic targets[6–9]. Recent pan-cancer analyses have identified cellular neighbourhoods[10–13], but have largely relied on spot-based data aggregated from heterogeneous public datasets or focused on individual tumour types. As a result, whether cellular neighbourhoods defined at single-cell resolution in harmonised data are shared across cancers or specific to subtypes remains unclear.

Here, we present the ASTRA1.0 atlas, a harmonised pan-cancer spatial transcriptomic map of approximately 400 tumours spanning 13 tissue types and 26 tumour subtypes, collected at 13 institutions in three Asia-Pacific countries. Central to the resource is a curated 1,999-gene panel designed to resolve the major cancer, immune and stromal compartments. We show that data generated with this workflow are reproducible across sites and robust to batch effects, enabling all tumours to be analysed as a single cohort. Alongside the spatial transcriptomic data, we provide matched spatial proteomics, cell-type annotations, cellular neighbourhood assignments and clinical metadata. Together, the shared panel and harmonised workflow provide a foundation for future studies, machine learning models, and a framework for regional alliances to support equitable access for precision medicine.

We used the ASTRA1.0 atlas to understand association between therapeutic targets and their neighbourhoods in a spatial context. First, we quantified the expression of cancer hallmarks in epithelial cells in a pan-cancer context. Next, we identified recurrent cellular neighbourhoods (CNs) and characterised their cell type composition. Finally, we investigated whether spatial architectures of therapeutic targets, which we term Spatial Theratopes (STs), could inform tumour-agnostic therapeutic strategies. We observed that tumours could be classified into four ST subtypes across tissue types, and that specific target pairs co-localised recurrently within distinct CNs. Importantly, we identified STs present in tumours, representing CN-dependent vulnerabilities that could guide the development of novel therapeutics for precision medicine. These findings suggest that the spatial neighbourhood, rather than the tissue type, may serve as a unit of therapeutic targeting, paving the way for tissue-agnostic therapies.

ASTRA1.0: A multimodal pan-cancer atlas from Asia-Pacific

We performed Xenium *In Situ* spatial transcriptomics on approximately 400 samples (381 in the discovery and 25 in the validation cohort). The cohort consists of 26 subtypes of cancers from major tissue types including bile duct, bladder, breast, colorectum, gastro-oesophageal junction, kidney, lung, liver, lymph node, oesophagus, oropharynx, ovary and pancreas collected from institutions across Australia, Japan and Taiwan (Fig. 1a,b; Methods). For consistency of assay and coverage of tissue types, we built a custom 1,999-gene panel (ASTRA2K panel) capable of resolving various cell types and their biological functions (Fig. 1c; Supplementary Table 1). To validate findings at the protein level, we also profiled more than 70% of the cohort (299 of 406 samples) using Akoya PhenoCycler

and Hyperion XTi (Supplementary Table 2; Methods). Importantly, all haematoxylin and eosin (H&E) sections were annotated in detail by pathologists, and each tumour was classified according to its clinical subtype. Each tumour was further annotated for tissue type, subtype, stage, primary or metastatic status, treatment history, and clinical outcomes, including relapse and survival (Fig. 1d). Patient characteristics and technical statistics for the atlas are provided in Supplementary Table 3 and Source Data.

To assess technical reproducibility, we profiled 12 normal lung reference regions processed across 12 independent runs. Pairwise Pearson correlation of gene expression between every pair of regions was high, ranging from 0.97 to 1.00 (Extended Data Fig. 1a–c). We also evaluated the number of genes detected, high-quality transcripts, cell density and median transcripts[14] (Fig. 1e; Extended Data Fig. 1c–f). The relative ranking of tissue types was largely consistent across metrics, indicating that variation in detection was driven by tissue-level factors common to all metrics rather than by any single QC measure. We next benchmarked the performance of the ASTRA2K panel against a smaller tissue-specific panel (v1) and a larger off-the-shelf pan-tissue panel (Xenium 5K). Several key cancer-associated genes, including *SPP1*, *KRT18* and *ACTA2*[15–17] were absent from the Xenium 5K panel. Across the 183 overlapping genes (Fig. 1f; Supplementary Table 4), we observed that Xenium 5K had significantly lower median transcripts per cell than both v1 and ASTRA2K, which did not differ significantly from each other (Fig. 1g; Extended Data Fig. 2a–c). Additionally, in matched samples, expression measured with ASTRA2K correlated well with both v1 and Xenium 5K ($r = 0.79–0.93$; Extended Data Fig. 2d), indicating concordance across panels. Restricted to the 183 shared genes, the number of genes detected per cell differed among all three panels, highest being v1, followed by ASTRA2K and then Xenium 5K (Extended Data Fig. 2b,g). PECAM1, targeted by more than six probes in ASTRA2K, was detected at substantially higher levels than in either v1 or Xenium 5K, indicating that increasing the number of probes further improves sensitivity (Extended Data Fig. 2h–k). Having established data quality and cross-panel comparability, we employed the atlas to study cancer hallmarks, define cellular neighbourhoods and identify tumour-agnostic therapeutic targets, as outlined in Fig. 1i.

Querying the ASTRA1.0 atlas for cell types

Given the harmonised experimental platform, the use of an identical panel and the same batch of reagents, the ASTRA1.0 atlas provided an opportunity to query shared and unique cellular landscapes. We integrated 56,693,873 (~57M) cells (Methods) and observed that cells in the tumour microenvironment (TME), including fibroblasts, endothelial cells and immune cells share similar gene expression features irrespective of tumour type (Fig. 2a,b). By contrast, epithelial cells organised into discrete clusters specific to individual tissue types, indicating that, compared to TME cells, epithelial cells exhibit tissue-type-specific transcriptional profiles across solid tumours. Consistent with the UMAP embedding (Fig. 2b), the distances between cancer-type centroids were significantly greater for epithelial cells than for other cell types ($P = 3.8e-92$; Fig. 2c; Methods).

The compositions of major cell types, including epithelial cells, fibroblasts, endothelial cells, T cells, macrophages and plasma cells, were characterised for each sample across

the respective tissue types (Extended Data Fig. 3a). Epithelial fractions were highest in NET, SCLC and LCNEC, all of which are neuroendocrine neoplasms (NENs), consistent with their solid, stroma-poor histology[18] (Extended Data Fig. 3a). The NENs were further distinguished by the *MKI67*$^+$ fraction of epithelial cells (P = 4.3e-16, Kruskal–Wallis test; Extended Data Fig. 3b): SCLC and LCNEC showed the highest proliferation of all subtypes whereas NETs were among the lowest, mirroring the distinction between high-grade neuroendocrine carcinomas and well-differentiated neuroendocrine tumours observed in diagnostic practice[19]. Together, these patterns indicate that composition and proliferation estimates from ASTRA1.0 are consistent with established histopathology

The *POSTN*$^+$ fibroblast and *PLVAP*$^+$ endothelial cells have previously been reported as oncofetal in HCC[6,20]. The ASTRA1.0 atlas provides an opportunity to investigate whether these cells are specific to tissue types or represent a pan-cancer phenomenon. We observed the highest frequency of *POSTN*$^+$ fibroblasts in aggressive lung cancers (Extended Data Fig. 3c), while *PLVAP*$^+$ endothelial cells were enriched in pancreatic cancer, fetal adenocarcinoma and cholangiocarcinoma (Extended Data Fig. 3d). Immune cell states also varied across subtypes, including the fractions of *CD8A*+ T cells (P = 2.1e-07; Extended Data Fig. 3e) and *CD163*$^+$ macrophages (P = 5.6e-12; Extended Data Fig. 3f). Pleomorphic carcinoma ranked highly for both, as well as for proliferation and *POSTN*$^+$ fibroblasts. Together, these analyses highlight the utility of ASTRA1.0 for querying tissue-specific biology in a pan-cancer context.

Given that the TME is broadly shared across tissue types, we used this feature to build a machine learning model to annotate cell types in the ASTRA1.0 cohort. ASTRA-CYTE was developed as a hierarchical classifier for the ASTRA2K panel (Methods). Cells are first assigned to one of five broad classes (epithelial, endothelial, fibroblast, myeloid or lymphoid) or left unassigned. Lymphoid cells are then resolved into T, NK, B and plasma cells, and myeloid cells into monocytes, macrophages and neutrophils (Extended Data Fig. 4a; Supplementary Table 5). Resolving lineage before subtype reduces the effect of class imbalance and keeps subtype errors within the correct lineage. ASTRA-CYTE was trained on one reference sample per tissue type and achieved high accuracy in tissue-type-held-out cross-validation (Extended Data Fig. 4d). On an external test set of 115 samples not used for training, ASTRA-CYTE predictions agreed with clinician-reviewed annotations, with macro-averaged F1 scores of 0.87–0.91 across nine tissue types and 0.74 in liver (Fig. 2d,e; Supplementary Tables 6 and 7). The lower score in liver largely reflected hepatocyte transcript contamination in the manual annotations (Extended Data Fig. 4b,e). ASTRA-CYTE annotations also preserved spatial organisation, closely matching manual annotation in a representative tissue section (macro F1 = 0.92; Fig. 2f,g).

Molecular hallmark analysis

Tumours are known to exhibit certain hallmarks such as proliferation, immune evasion and phenotypic plasticity[1]. However, this framework to our knowledge, has not been applied to a large harmonised spatial cohort. We scored 13 cancer hallmarks in epithelial cells of all tissues and performed hierarchical clustering to compare cancer subtypes (Fig. 2h; Extended Data Fig. 5a; Supplementary Table 8; Methods). Because ASTRA1.0 is a transcriptomic atlas, hallmark scores reflect the expression of gene programmes

associated with each hallmark rather than the underlying genomic or functional states. Notably, cancers arising from the same tissue type are categorised into different groups, indicating that hallmark profiles are shaped by tumour subtype rather than tissue of origin.

The first group contained the NENs (NET, LCNEC and SCLC), which showed high scores for 'Genome Instability and Mutation' and low scores for 'Avoiding Immune Destruction'. For genome instability, this reflects expression of DNA repair and replication stress genes rather than mutational burden. SCLC epithelial cells showed abundant *PARP1* expression together with *RAD51* transcripts, whereas these transcripts were sparse in ovarian clear cell carcinoma (OCCC) (Fig. 2i). This is consistent with the replication stress caused by near-universal RB1 and TP53 loss in SCLC[21], and with the elevated *PARP1* expression previously reported in SCLC[22]. In LCNEC, epithelial cells formed compact nests that contained few T cells, and *CD274* transcripts were detected mainly in macrophages in the surrounding stroma rather than in tumour cells (Fig. 2j) suggesting that these tumours may evade immunity through exclusion rather than suppression[23].

The second group included LUAD, pancreatic cancer and bile duct cancers. These tumours demonstrated elevated 'Activating Invasion and Metastasis' (Extended Data Fig. 5b), consistent with the partial epithelial-mesenchymal transition reported in these adenocarcinomas[24] and the desmoplastic stroma of pancreatic and bile duct cancers[25,26]. Unlike the neuroendocrine group, these tumours scored highly for 'Avoiding Immune Destruction'. In LUAD, *CD274* transcripts were detected in epithelial cells, and T cells, including *PDCD1*-expressing cells, were interspersed among tumour cells (Fig. 2j). Thus, although LUAD and LCNEC arise in the same tissue type, they show distinct spatial patterns of immune evasion, which may underlie their different responses to immune checkpoint blockade.

The third group was defined by high scores for 'Unlocking Phenotypic Plasticity' exhibited by fetal adenocarcinoma epithelial cells expressing *POU5F1* (OCT4) and *SOX2* transcripts, whereas both were nearly absent from pulmonary pleomorphic carcinoma (PLEO) (Fig. 2k). OCT4 and *SOX2* are core pluripotency transcription factors that maintain an undifferentiated state[27]. Their expression suggests that the fetal-like histology of fetal adenocarcinoma is accompanied by an embryonic transcriptional programme.

Recurrent cellular neighbourhoods across tissue types

Spatial transcriptomics can resolve the cellular neighbourhoods of tumours. However, it remains unclear whether they are specific to individual tissue types or shared across tumours. A harmonised spatial atlas allows us to investigate the similarity of CNs across tissue types. We identified CNs by grouping cells based on the composition of their local neighbourhoods (100 nearest neighbours)[4] (Fig. 3a; Methods). This revealed eight CNs (CN1–CN8) across the atlas, each with a distinct cellular composition and gene expression signature (Fig. 3b,c; Supplementary Tables 9 and 10). The CNs were dominated by epithelial cells (CN1), macrophages (CN2), T cells (CN3), plasma cells (CN4), vascular cells (CN5), neutrophils (CN6), B cells (CN7) and stromal cells (CN8) (Fig. 3d,e). Two of these CNs (CN3 and CN7) represented different maturation stages of tertiary lymphoid structures (TLS). CN3 expressed chemokines *CCL19*, *CXCL9* and

showed co-localisation of B cells within the CN (Fig. 3b), consistent with early lymphoid aggregates[28,29]. In contrast, CN7 was more strongly enriched for B cells and expressed *CR2*, a marker of follicular dendritic cells and *CXCL13*, characteristic of mature, follicle-containing TLS[30,31].

To determine whether the eight-CN framework reliably represents the major ecosystems, we performed stability and silhouette analyses (Methods). Eight CNs gave the most stable architecture with near-identical assignments (mean pairwise adjusted Rand index ARI = 0.994; Extended Data Fig. 6a,b) and this framework was also robust to changes in neighbourhood size from 25 to 200 nearest neighbours (ARI = 0.79–0.80; Extended Data Fig. 6c). To evaluate reproducibility of CNs, we compared Australian and Japanese HCC cohorts (Methods) and observed near-identical cell-type composition within each CN (r = 0.984; Extended Data Fig. 6d–f). Finally, analysis of 365 samples from 13 tissue types indicated recurring patterns of CNs across tissue types (Fig. 3f; Supplementary Table 11). Taken together, these analyses provide a framework for identifying shared neighbourhoods in tumours across cohorts.

We next investigated whether individual cell types change their transcriptional state depending on neighbourhood occupancy. We calculated expression shifts for each cell type (Extended Data Fig. 7a; Methods) and found that shifts of cell types within the same CN were consistently correlated. Without any prior grouping, CNs naturally organised into four blocks corresponding to epithelial, immune, stromal/vascular and neutrophil neighbourhoods (Extended Data Fig. 7b). To determine which cell types, underlie these CN-specific programmes across tumours, we ranked them by a conservation score (Methods). Macrophages were the most conserved across tumours, followed by fibroblasts and T cells, whereas epithelial cells were the least conserved, (Fig. 3g,h; Extended Data Fig. 7c; Supplementary Table 12) consistent with their tissue-type-specific transcriptional identity (Fig. 2b,c).

To understand what these conserved shifts represent, we examined the genes associated with macrophages in each CN (Methods) and observed that macrophages in the T cell CN (CN3) upregulated MHC class II (*CD74*, *HLA-DRA*, *HLA-DPA1*, *HLA-DPB1*), together with *STAT1* and *CXCL9*, consistent with an interferon-primed, antigen-presenting state[15,32]. By contrast, macrophages in the stromal CN (CN8) were enriched for scavenger receptors *CD163*, *STAB1* and *MRC1*, along with *RNASE1*, *FOLR2* and complement *C1Q*, a programme associated with immunoregulatory and resident-like macrophage states in human tumours[33,34] (Extended Data Fig. 7d; Supplementary Table 13). Collectively, these results resemble the concept of foundation species in ecology[35], indicating that macrophages are key organisers of CNs and suggest that therapeutic targeting of these cells could disrupt tumour-associated ecosystems.

Stratification of tumours based on Spatial Theratopes

Antibody-drug conjugates (ADCs), bispecific antibodies and rational combinations of immune checkpoint inhibitors are reshaping cancer therapy. However, patient selection largely lacks robust stratification, which contributes to poor outcomes[36,37]. Although these therapies work by engaging their targets, it remains unclear how the local neighbourhood of these targets affects clinical outcomes across tissue types. We therefore assessed co-

localisation between tumour-associated targets (e.g. HER2, TROP-2, EGFR, EpCAM, VEGF-A) and immune checkpoint targets (e.g. PD-L1, PD-1, CTLA-4) across 340 tumours from 12 tissue types and 23 histological subtypes, excluding lymph nodes and normal samples (full target list in Supplementary Table 14). A cell was scored as co-localised for a target pair if it expressed the tumour target and at least one of its 100 nearest neighbours expressed the checkpoint target (Fig. 4a; Methods). We next applied k-means clustering to determine the patterns of target co-localisation across tumours (Methods) and identified four Spatial Theratopes (STs; Extended Data Fig. 8a–d).

Because targeted therapies are developed and approved for specific tissue types, we next asked whether STs reflected the tissue of origin. ST1 comprised lung, ovarian and breast tumours, and ST2 and ST4 included tumours from several tissue types, whereas ST3 consisted predominantly of primary and metastatic liver tumours (Fig. 4b; Supplementary Table 15). Moreover, tumours from the same tissue type, such as lung, were distributed across multiple STs, indicating intertumoural heterogeneity in target organisation. ST1, ST2 and ST4 were each characterised by a distinct co-localised target pair (TROP-2–PD-L1, EpCAM–PD-L1 and tissue factor–PD-L1, respectively), whereas VEGF-A–PD-L1 was present across all four STs (Fig. 4c; Extended Data Fig. 8e; Supplementary Table 16). Notably, combined blockade of the VEGF and PD-1/PD-L1 pathways has shown clinical benefit in several solid tumours[38–40]. Imaging mass cytometry on the same sections (Methods) confirmed VEGF-A and PD-L1 protein in adjacent cells in representative samples, consistent with the RNA-based ST annotation (Fig. 4d). To identify the determinants of ST membership, we trained classifiers to predict each ST from tissue type, histological subtype and CN composition (Methods). ST2 and ST3 membership was largely determined by histological subtype, whereas adding CN composition modestly improved prediction of ST1 and ST4 (mean area under the curve (AUC) gains of 0.05 and 0.08, respectively; Extended Data Fig. 8f; Supplementary Table 17). This suggests that, for some tumours, the spatial context of therapeutic targets carries information beyond tumour type.

As tumours within an ST share similar target co-localisation profile, we asked whether local neighbourhoods further distinguished clinical outcomes within an ST. First, in a subset of 48 resected lung cancers of various subtypes assigned to ST1 (Fig. 4e,f; Supplementary Table 18; Methods), the presence of a TLS-like neighbourhood (CN7) was associated with longer recurrence-free survival (Fig. 4g; Methods), and CN7 abundance was lower in patients who experienced recurrence (Fig. 4h). By contrast, the hazard ratio for the epithelial-enriched CN1 neighbourhood, which showed enrichment of *TACSTD2* (TROP-2) expression (Supplementary Table 10), pointed in the opposite direction, although this was not statistically significant (Supplementary Table 19). This raises the hypothesis that patients with CN7-low, CN1-rich tumours could be candidates for evaluation of TROP-2-directed therapy. Next, we examined a cohort of patients with HCC treated with atezolizumab plus bevacizumab, which target PD-L1 and VEGF-A, respectively. Of 25 tumours, 23 were assigned to ST3 (one each to ST2 and ST4); the 22 ST3 tumours with complete follow-up were analysed (Fig. 4i,j; Extended Data Fig. 9d; Supplementary Tables 20 and 21). In this cohort, co-localisation of the drug targets (VEGF-A and PD-L1) alone was not associated with progression-free survival (PFS; Fig. 4k) or best response (Extended Data Fig. 9g,h). However, a higher proportion of the

stromal neighbourhood (CN8) was associated with shorter PFS (Fig. 4k,l; Extended Data Fig. 9f). These results suggest that, within a shared Spatial Theratope, neighbourhood context may further stratify patients, a hypothesis that warrants prospective evaluation.

Discussion

Large-scale cancer genomics atlases make it possible to characterise tumours and their clinical features at population scale[41]. When such resources represent diverse ancestries and health systems, they can also help ensure that the benefits of precision oncology are distributed equitably[42]. Here we present ASTRA1.0, a multimodal pan-cancer atlas comprising approximately 400 tumours across 13 tissue types and 26 subtypes from three Asia-Pacific countries, profiled on a shared 1,999-gene panel with matched spatial proteomics and clinical annotation. We demonstrated that the ASTRA2K panel achieved higher sensitivity than the off-the-shelf pan-tissue panel and that the resulting data were reproducible across cohorts and robust to batch effects. This harmonisation allowed us to identify microenvironment features shared across tissue types and to build a cell-annotation classifier, ASTRA-CYTE, that enables the community to identify cell types in current and future ASTRA cohorts.

We illustrate the utility of this resource by examining oncofetal stromal cells in a pan-cancer setting. Previously, we identified *POSTN*$^+$ fibroblasts and *PLVAP*$^+$ endothelial cells as oncofetal cells in HCC[6,20]. Analysis of the ASTRA1.0 cohort revealed that these cells are present across multiple tissue types in the atlas, with the highest proportion of *POSTN*$^+$ fibroblasts observed in aggressive lung cancers. Our hallmark analysis extends a framework largely derived from bulk data into a spatial context. We observed co-localisation of PD-L1$^+$ tumour cells and PD-1$^+$ T cells in LUAD, whereas lung neuroendocrine tumours exhibited the physical exclusion of T cells and PD-L1 expression on tumour-associated macrophages. These findings suggest distinct modes of immune evasion within the same tissue type, with potential implications for immunotherapy. This analysis also highlighted phenotypic plasticity as a hallmark of fetal adenocarcinoma, suggesting a molecular basis for the fetal-like morphology associated with these tumours.

Notably, the recurrence of eight cellular neighbourhoods across tissue types supports the notion that niches are a shared organisational unit of solid tumours[10–13]. Among these neighbourhoods, macrophages and fibroblasts were the most conserved building blocks, reminiscent of foundation species in ecosystems. Macrophages also exhibited CN-specific states, such as an antigen-presenting state in T cell CN (CN3) and tissue-resident signature in the stromal CN (CN8). By contrast, epithelial cells displayed cancer-type-specific responses. Notably, macrophages are among the earliest cells to emerge from yolk sac during development[43], raising the possibility that they act as organisers of cellular neighbourhoods. Furthermore, Spatial Theratopes grouped tumours by target co-localisation across tissue types, and within STs, neighbourhood composition provided additional information about clinical outcome. In resected ST1 lung cancers, the presence of a TLS-like neighbourhood (CN7) was associated with longer recurrence-free survival after adjustment for stage and histology, in line with the established prognostic value of mature TLS[28]. In patients with HCC treated with atezolizumab plus bevacizumab, a higher proportion of the stromal neighbourhood (CN8) was associated with shorter progression-free survival, whereas co-localisation of the drug targets alone was not,

consistent with prior reports that fibroblast-rich stroma is associated with resistance to PD-L1 blockade[6,44].

The following limitations should be considered. The targeted 1,999-gene panel enables detection of cancer-specific genes and harmonisation across sites but limits discovery of programmes involving genes outside the panel. Spatial Theratopes were defined at the transcript level, and although imaging mass cytometry confirmed key pairs, therapeutic response depends on protein abundance and accessibility, which transcripts alone does not capture completely. The outcome analyses were retrospective, and these findings should be validated in independent prospective cohorts.

The harmonised data and clinical annotations further provide a resource for machine learning models that integrate cell composition, neighbourhoods and clinical data to predict treatment response and clinical outcomes. Spatial Theratopes could, if validated prospectively, inform tumour-agnostic stratification in clinical trials of ADCs, bispecific antibodies and immunotherapy combinations by selecting patients based on spatial target architecture rather than tumour type alone. Finally, as new countries and centres join the ASTRA consortium, larger and more diverse cohorts will enable ancestry to be disentangled from environmental exposures, such as pollution and infection, and from disparities in access to treatment, clarifying how each shapes clinical response.

Figure Legends

Fig.1: Asia-Pacific spatial multimodal atlas of human cancers.

a, ASTRA cohort overview demonstrating the approximate number of cells and samples (*n*) per tissue type, with histological subtypes indicated. **b**, Geographical location of contributing institutions and the number of samples. **c**, Composition of the Xenium ASTRA2K gene panel (1,999 genes) by gene ontology. **d**, Heatmap of clinical and technical annotations for all samples including tissue type, histological subtype, tumour stage, primary or metastatic status, treatment before sampling, relapse, survival status and availability of matched spatial proteomics. Samples are ordered by tissue type, then subtype. Grey, not available. **e**, Median genes detected per cell in each sample (n = 365), grouped by tissue type. Box plots show median and interquartile range (IQR), whiskers extend to 1.5× IQR, points represent individual samples. **f**, Venn diagram of gene overlap among the Xenium v1 (302 genes), ASTRA2K and Xenium 5K (5,101 genes) panels. **g**, Median transcripts detected per cell for six samples, each profiled with all three panels. *P* values, two-sided paired Wilcoxon signed-rank tests with Benjamini–Hochberg correction. **h,** Representative Xenium Explorer images of *YAP1* transcripts (green; one dot per transcript) in sample jichihgnec27 (LCNEC) profiled with each panel. Images show different fields; background shows DAPI (v1) or multimodal cell segmentation staining (ASTRA2K, 5K). p95, 95th percentile of *YAP1* transcripts per cell across the field shown. Scale bar, 50 μm. **i**, Overview of the analytical workflow: cancer hallmark scoring, cellular neighbourhood (CN) analysis and identification of tumour-agnostic therapeutic targets.

Fig. 2: A pan-cancer spatial single-cell atlas and automated cell-type annotation with ASTRA-CYTE.

a, UMAP embedding of the ASTRA1.0 cohort coloured by cancer subtype. **b**, UMAP embedding coloured by annotated cell type. **c**, Distances between cancer-type centroids in the reference principal component space (50 PCs), shown per cell type. Each dot represents one pair of tissue types (n = 78–276 pairs per cell type; n = 231 pairs for epithelial cells). *P* = 3.8e-92, one-sided Mann–Whitney *U* test (epithelial cells versus all other cell types pooled). **d,e,** ASTRA-CYTE annotation performance on an external test set of 115 samples across ten tissues, none of which were used for training. **d,** Per-cell-type F1 scores for broad classes (epithelial cells, fibroblasts and endothelial cells), lymphoid cells (NK/T cells, B cells and plasma cells) and myeloid cells (macrophages, monocytes and neutrophils). Dot size indicates the number of cells and dot colour indicates F1 score. **e,** Macro F1 score per tissue. **f,g,** A representative colorectal cancer section (P24_4891_CRC) coloured by **f,** clinician-reviewed annotations and **g,** ASTRA-CYTE predictions (macro F1 = 0.92; scale bar, 1 mm). **h,** Heatmap showing the profile of 13 cancer hallmark signatures across 25 tissue types. Hallmark scores were calculated for each epithelial cell, averaged within each sample and then within each tissue type, and z-scored across tissue types for each hallmark. Rows (hallmarks) and columns (tissue types) were ordered by hierarchical clustering using Euclidean distance and Ward's linkage. **i**, Representative Xenium Explorer images of the 'Genome Instability and Mutation' hallmark shown by *BRCA1* (purple), *RAD51* (green) and *PARP1* (orange) transcripts in SCLC (tkbhgnec34) and ovarian clear cell carcinoma (4_040-L2), overlaid on epithelial cells (Scale bar, 100 μm). **j**, The 'avoiding immune destruction' hallmark illustrated by *CD8A* (yellow), *CD274* (orange) and *PDCD1* (blue) transcripts in

LUAD (tkbadeno33) and LCNEC (tkbhgnec32a), overlaid on the spatial distribution of epithelial cells, T cells and macrophages (Scale bar, 20 μm). **k**, Representative images of the 'Unlocking Phenotypic Plasticity' hallmark by *POU5F1* (green) and *SOX2* (blue) transcripts in fetal adenocarcinoma (tkbfetal3_1) and pleomorphic carcinoma (tkbpleo18), overlaid on epithelial cells (Scale bar, 50 μm).

Fig.3: Recurrent cellular neighbourhoods and conserved CN-induced cell states across tissue types.

a, Schematic of cellular neighbourhood (CN) analysis in tumour tissue. The neighbourhood of each cell is defined by its 100 spatial nearest neighbours, and cells are clustered into CNs based on cell-type composition. **b**, Cell-type composition of the eight CNs, each annotated by its dominant cell type: epithelial (CN1), macrophage (CN2), T cell (CN3), plasma cell (CN4), vascular (CN5), neutrophil (CN6), B-cell (CN7) and stroma (CN8). **c**, Prominent marker genes for each CN reflecting the dominant cell type of each CN. The heatmap shows the median log2 fold change in expression for each CN compared with all other neighbourhoods. Selected marker genes are annotated in heatmap. **d**, Spatial map of CNs in a representative lung adenocarcinoma section tkbadeno28. Boxes mark the representative cellular neighbourhoods highlighted in panel e (Scale bar, 500 μm). **e**, Representative regions for each CN. Top, cells coloured by cell type annotations. Bottom, transcripts of selected marker genes for each CN (Scale bar, 100 μm). **f**, Recurrence analysis for each CN across tissue types (n = 365 samples). Dot size shows prevalence, defined as the percentage of samples in which the CN makes up more than 1% of cells. Dot colour intensity represents the median CN abundance in positive samples. **g,h,** Correlation of CN-induced expression shifts across CNs and tissue types for **g,** macrophages and **h,** epithelial cells. Each block corresponds to a pair of CNs; rows and columns within blocks are tissue types. Macrophage programmes are strongly correlated across tissue types, whereas epithelial programmes show weaker correlation.

Fig.4: Spatial Theratopes identify co-targetable tumour–immune interactions and cellular neighbourhoods associated with clinical outcome.

a, Schematic of Spatial Theratope (ST) identification in 340 tumours (12 tissue types, 23 subtypes). For each target pair, a cell was scored positive if it expressed a tumour target and at least one of its 100 nearest neighbours expressed an immune checkpoint target. The resulting tumour-level co-localisation profiles were clustered by k-means into four STs. **b,c**, n = 340 tumours from 12 tissue types and 23 subtypes (ST1, n = 61; ST2, n = 105; ST3, n = 111; ST4, n = 63). **b**, Proportion of tumours from each subtype within ST1–ST4. **c**, Percentage of tumours in each ST with a min–max-scaled score > 0.5 for VEGF-A–PD-L1 and for the top enriched target pairs for each ST. **d**, Imaging mass cytometry on the same sections of two hepatocellular carcinoma (HCC) samples (M365T and M356T), showing VEGF-A and PD-L1 protein co-localisation consistent with ST3. Segmented cells are coloured by cell type; PD-1 (white), PD-L1 (green) and VEGF-A (brown) protein signals are overlaid on *PDCD1* (purple), *CD274* (green) and *VEGFA* (magenta) transcripts. Scale bar, 20 μm. **e**, University of Tsukuba resected lung cancer cohort assigned to ST1 (n = 48 patients; seven histological subtypes; n = 47 with recurrence data). **f**, Cellular neighbourhood (CN) composition and clinical features (relapse, stage and subtype) for each patient. **g**, Kaplan–Meier curves of recurrence-free survival stratified by CN7 presence (proportion > 0; n = 39) and absence (n = 8). Shaded

regions, 95% confidence intervals (CIs); *P* values, log-rank tests; hazard ratios (HRs), Cox models. **h**, CN7 proportion in patients with (n = 19) and without (n = 28) recurrence. Box plots show the median and interquartile range, with whiskers extending to 1.5× the interquartile range; each point represents one patient. P value, two-sided Mann–Whitney U test. **i–l**, Independent cohort of patients with HCC treated with atezolizumab (anti-PD-L1) plus bevacizumab (anti-VEGF-A) at Gifu University (n = 22 patients). **i**, Assignment to STs. 23 of 25 tumours were assigned to ST3 and 22 were analysed. **j**, CN composition (log10 proportion), progression and best response by mRECIST (CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease). **k**, Univariable Cox models for PFS, with VEGF-A–PD-L1 score and CN8 proportion each modelled as median-split and continuous variables. Points, hazard ratios; error bars, 95% CIs; P, two-sided Wald test. Dashed line, HR = 1. **l**, Kaplan–Meier curves of progression-free survival (PFS) by CN8 proportion, split at the median (high, n = 10; low, n = 12); shading and P value as in **g**.

Methods

Clinical samples

The ASTRA1.0 discovery cohort comprises 381 samples spanning 13 tissue types and 26 subtypes, including tumours from the lung, liver, bile duct, pancreas, oropharynx, oesophagus, gastro-oesophageal junction, colorectum, breast, ovary, bladder and kidney, and lymph node tissues. An independent cohort of 25 HCC samples was used for validation of Spatial Theratopes, totalling 406 samples. Samples were contributed by thirteen institutions across Australia, Japan and Taiwan as depicted in Fig. 1b. Pathological diagnosis was made by certified clinicians at each institution according to the World Health Organization classification of tumours for the corresponding tissue type[45]. De-identified clinical and pathological information was obtained from medical records. Ethical sample collection and informed consent followed the procedures of each institution. The study was approved by the institutional review board of each participating institution; approval numbers are listed in Supplementary Table 23.

In situ RNA expression profiling by Xenium

Xenium in situ expression analysis was performed using the Xenium Slides & Sample Prep Reagents (PN-1000460, 10x Genomics)[46]. FFPE tissue blocks were sectioned at 5 μm thickness and mounted on Xenium slides (10x Genomics) following the manufacturer's instructions (CG000578, 10x Genomics). The slides were then incubated at 42°C for 3 h. Deparaffinisation and decrosslinking were performed following the manufacturer's instructions (CG000580, 10x Genomics). The subsequent steps followed the user guide for 'Xenium In Situ Gene Expression' (CG000582, 10x Genomics). The custom-designed ASTRA2K panel, which targets 1,999 genes (design ID: CHZBUW), was hybridised to the tissue at 50°C overnight, followed by washing, ligation, and amplification. Autofluorescence quenching and nuclei staining were performed in the dark. The slides were run on the Xenium Analyzer (10x Genomics). The full list of target genes in the ASTRA2K panel is provided in Supplementary Table 1.

Spatial protein profiling

Spatial protein expression was profiled by two complementary methods. Multiplexed immunofluorescence was performed with the PhenoCycler system (Akoya Biosciences)[47] (n = 283), using the same tissue sections analysed by Xenium. After the Xenium run, sections were washed, subjected to antigen retrieval in a pressure cooker for 20 min and stained with a panel of 79 antibodies for 3 h (Supplementary Table 2). PhenoCycler-inventoried antibodies validated by Akoya Biosciences across multiple FFPE tissue types were prioritised, and the staining quality of custom-conjugated antibodies was verified on test tissues. Stained sections were washed and the antibodies were fixed. Imaging was performed on the PhenoCycler instrument with a BZ-X810 fluorescence microscope (Keyence).

Imaging mass cytometry was performed on a Hyperion XTi Imaging System (Standard BioTools; RRID:SCR_023195)[48] on the same sections analysed by Xenium[49]. Sixteen sections underwent antigen retrieval in AR9 buffer (pH 9, Akoya Biosciences) at 96°C for 30 min, were blocked with 3% BSA for 45 min and were stained overnight at 4 °C with a panel of 43 metal-conjugated antibodies validated by immunohistochemistry (Supplementary Table 2). Nuclei were stained with an iridium intercalator.

For visual comparison with RNA data, IMC images were exported as OME-TIFF files and registered to the corresponding Xenium sections in Xenium Explorer (v5, 10x Genomics) using landmark-based alignment, and VEGF-A and PD-L1 protein signals were overlaid on Xenium transcript and cell segmentation data (Fig. 4d).

Evaluation of the ASTRA2K panel

To assess batch effects, serial sections of a normal lung sample were run on 12 Xenium slides (Extended Data Fig. 1a,b). Per-gene mean transcript counts (log1p-transformed) were compared between slides by pairwise Pearson correlation, alongside standard quality metrics (transcripts and cells per μm², and median genes and transcripts per cell). For cross-panel comparison, sections from six large-cell neuroendocrine carcinomas were profiled with the Xenium v1, ASTRA2K and Xenium 5K panels. Using the 183 genes shared by all three panels, per-sample median transcripts and genes per cell were compared by paired Wilcoxon signed-rank tests with Benjamini–Hochberg correction[50] (Extended Data Fig. 2a–d; Supplementary Table 4). Gene-wise consistency was assessed by Pearson correlation of $\log_2$ mean transcript counts and by fold changes relative to the matched 5K sample.

Xenium data preprocessing and quality control

Cells with fewer than 10 transcripts, fewer than five detected genes or a segmented area below 15 μm² were excluded. Thresholds were chosen from the pooled distributions of these metrics across samples, with each sample weighted equally, and applied uniformly to all samples. This quality control was applied for cell-type classification, cellular neighbourhood analysis and Spatial Theratope classification; additional analysis-specific filters are described in the relevant sections.

Cell type annotation

Cell types were first annotated with SingleR (v2.14.1)[51] using the Human Primary Cell Atlas reference (celldex v2024-02-26)[52]. The reference was restricted to twelve main labels relevant to the tumour microenvironment, and its fine-grained labels were used to separate plasma cells from B cells and lymphatic endothelial cells from endothelial cells. Labels were assigned to individual cells and to Seurat clusters (Seurat v5.1.0)[53]. Clusters were flagged as low confidence when the cluster-level label differed from the most frequent cell-level label, or when that label accounted for less than 50% of the cluster's cells. Annotations and flagged clusters were reviewed by the clinician responsible for each sample against H&E staining and pathological findings, and labels inconsistent with histology were revised.

These reviewed annotations were used to train ASTRA-CYTE, a hierarchical classifier of gradient-boosted decision trees (XGBoost[54]). ASTRA-CYTE first assigns cells to five broad classes (epithelial cells, fibroblasts, endothelial cells, myeloid cells and lymphoid cells) and then resolves lymphoid and myeloid cells into seven subtypes (Extended Data Fig. 4a). ASTRA-CYTE was trained on 2,090,457 cells from ten tissue types and cells with prediction confidence below 0.7 at the subtype level retained their broad-class label, and cells with a confidence below 0.7 at the broad-class level were labelled unassigned. Applied to 115 samples not used for training, ASTRA-CYTE predictions agreed with the

clinician-reviewed annotations, with macro-averaged F1 scores of 0.87–0.91 across nine tissue types and 0.74 in liver (Fig. 2d,e; Supplementary Table 7). Details of the reference data, training and evaluation are provided in the Supplementary Methods.

Reference embedding and projection

Data from all samples were combined into a single dataset of 56,693,873 cells. A reference embedding was built from 10,000 randomly selected cells per subtype (250,000 cells in total) by normalisation, principal component analysis (50 components), UMAP[55] and Leiden clustering (resolution 0.5)[56]. No batch correction was applied. All remaining cells were then projected onto this reference with the ingest function of scanpy (v1.11.4)[57]. For centroid analysis, cells were grouped by sample and cell type, downsampled to 500 cells per group and averaged in this principal component space. The centroid of a cell type in a tissue type was the unweighted mean of its sample centroids. Distances between cancer-type centroids were compared between epithelial cells and pooled tumour microenvironment cell types by a one-sided Mann–Whitney U test.

Cell-type composition and marker positivity

Cell-type frequencies were calculated per sample as the proportion of all cells. Normal lung samples (n = 17) and lymph node samples (n = 5) were excluded, leaving 336 tumours across 25 subtypes. For individual genes, a cell was considered positive if it had at least one transcript, and positive fractions were computed per sample and cell type, excluding groups of fewer than 20 cells. Tissue types were compared by Kruskal–Wallis tests with Benjamini–Hochberg correction.

Cancer hallmark analysis

Hallmark activity was scored per cell with UCell (v2.10.1, maxRank = 500)[58] in the cells annotated as epithelial. The gene sets for the 13 cancer hallmarks were taken from a previous study[59], and each set was restricted to genes present in the 1,999-gene panel (Supplementary Table 8). Per-cell scores were averaged within each sample, and per-sample means were averaged within each tissue type. The resulting matrix was z-scored per hallmark and then centred per tissue type, so that the values shown reflect the relative activity of each hallmark within a tissue type rather than its absolute level. Tissue types were clustered on both axes by Ward linkage on Euclidean distances[60]. For the radar chart (Extended Data Fig. 5), these scores were averaged across tissue types within each tissue of origin, weighted by the number of samples, and z-scored per hallmark.

Cellular neighbourhood identification

For each cell, the local neighbourhood was defined as its 100 nearest cells, excluding itself, by Euclidean distance between cell centroids within each sample[4]. A composition vector was computed giving the fraction of these neighbours assigned to each of ten ASTRA-CYTE cell types. Unassigned and unresolved lymphoid and myeloid neighbours were excluded from the calculation, and cells with no assigned neighbours were removed. Composition vectors from all cells across 365 samples, including normal lung, were z-scored per cell type and clustered jointly by k-means (scikit-learn v1.7.2[61] k = 8; 30 initialisations; random seed 42) to define eight cellular neighbourhoods (CNs). The model was fitted on a random subsample of 2,000,000 cells, and all cells were then assigned to the nearest centroid. Each CN was named after the dominant cell types in the mean

neighbourhood composition of its cells (Figure 3b; Supplementary Table 9). Selection of k and robustness analyses are described in the Supplementary Methods.

Identification of neighbourhood-associated genes

To identify genes associated with each CN independently of tissue type, raw counts were summed across all cells of each CN in the four samples with the most cells per tissue type, to limit the influence of tissue types with many samples. Pseudobulk profiles from fewer than 100 cells were discarded, and the remainder were converted to $\log_2(\text{CPM} + 1)$. Within each tissue type, the $\log_2$ fold change of each CN was computed relative to the mean of the remaining CNs and summarised across tissue types by the median. Genes were considered CN-associated if the CN had a pseudobulk profile in at least four tissue types and the gene's fold change was positive in at least 75% of them. This analysis was descriptive, and no significance testing was performed. Top genes per CN are shown in Fig. 3c, and complete ranked lists are provided in Supplementary Table 10.

Cellular neighbourhood prevalence across tissue types

A CN was considered present in a sample when it comprised more than 1% of that sample's cells. For each tissue type, prevalence was the fraction of samples meeting this threshold, and abundance was the median fraction of cells assigned to the CN among samples in which it was present (Fig. 3f; Supplementary Table 11).

CN-induced expression shifts

To test whether CNs affect their resident cells consistently across tissue types, pseudobulk expression profiles were computed for each combination of sample, CN and cell type with at least 30 cells, using the ten final ASTRA-CYTE cell types. To limit the influence of transcripts misassigned from neighbouring cells, each cell type was analysed only for genes expressed predominantly in that cell type. For each cell type in each sample, the CN-induced shift was defined as the difference between its pseudobulk profile in a given CN and its mean profile across CNs in that sample, removing patient-, tissue- and cell type-specific baselines. Shifts were standardised per gene before downstream analyses.

Conservation of CN-induced shifts across tissue types

For each cell type, shifts were averaged across samples within each tissue type, and Pearson correlations were calculated between shift profiles from different tissue types (Extended Data Fig. 7a; Supplementary Table 12). The conservation score was defined as the median correlation between profiles of the same CN minus that between profiles of different CNs. Significance was assessed by permuting CN labels within samples (200 permutations), with Benjamini–Hochberg correction across cell types. One-sided P values were computed as the proportion of permuted conservation scores greater than or equal to the observed score. CN families were defined by inspection of the block structure and confirmed by mean between-block correlations. Further details, including a sample-level analysis, are provided in the Supplementary Methods.

Transfer of CNs to the HCC cohort

Neighbourhood compositions in the HCC cohort were computed as in the discovery cohort (100 nearest neighbours, ASTRA-CYTE cell types). Compositions were z-scored using the mean and standard deviation of the discovery cohort, and each cell was assigned

to the discovery CN with the nearest centroid (Supplementary Table 20).

Target-pair co-expression scoring

Twelve tumour-associated targets and four immune checkpoint targets were selected based on approved or clinical-stage antibody-based therapies (e.g. bispecific antibodies and antibody–drug conjugates). The tumour-associated targets were HER2 (*ERBB2*), HER3 (*ERBB3*), EGFR, TROP-2 (*TACSTD2*), FRα (*FOLR1*), EpCAM, VEGF-A, VEGFR2 (*KDR*), gp100 (*PMEL*), tissue factor (*F3*), MET and DLL3; the immune targets were PD-1 (*PDCD1*), PD-L1 (*CD274*), CTLA-4 and LAG-3. A cell was considered positive for a target if at least one transcript of the corresponding gene was detected. For each tumour–immune pair, a cell was scored as positive if it expressed the tumour target and at least one of its 100 nearest neighbours (Euclidean distance between cell centroids, excluding the cell itself) expressed the immune target. For immune × immune pairs, a cell was scored as positive if it expressed either target and at least one neighbour expressed the other. PD-1 × PD-L1 was excluded as a receptor–ligand interaction rather than a combination of therapeutic targets. The score for each pair in a sample was the fraction of all cells scored positive. Samples with fewer than 50 cells were excluded. Scores for all samples can be found in Supplementary Table 15.

Spatial Theratope classification

Lymph node samples (n = 5), normal lung samples (n = 17) and normal liver samples (n = 3) were excluded, leaving 340 tumours. Scores for 53 target pairs were min–max scaled within each tumour to a range of 0 to 1, so that clustering reflected the relative ranking of target pairs within each tumour rather than absolute levels. Spatial Theratopes (STs) were defined by k-means clustering[61] of the scaled score matrix (k = 4; 50 random initialisations; random seed 0). Silhouette widths were low across k = 2–10, indicating gradual rather than discrete structure; k = 4 was selected because it had the highest silhouette width among $k \leq 5$ and high subsampling stability, and because it separated a group of tumours (ST4) that was merged with ST1 at k = 3 but recovered almost intact at k = 6 (Extended Data Fig. 8a–d). Selection of k and stability analyses are described in the Supplementary Methods.

To assess the contribution of tissue of origin and neighbourhood composition to ST membership, random forest classifiers were trained to predict each ST from tissue type, histological subtype and CN proportions, alone and in combination, and evaluated by repeated cross-validation (Extended Data Fig. 8f; Supplementary Table 17; Supplementary Methods).

Target-pair enrichment across Spatial Theratopes

To identify target pairs characteristic of each ST, the min–max-scaled pair scores of tumours in each ST were compared with those of all remaining tumours by two-sided Mann–Whitney U tests, separately for each target pair, with Benjamini–Hochberg correction across pairs within each ST. As a measure of effect size, the fraction of tumours with a pair score >0.5 was calculated in each ST and in the remaining tumours, and the difference was taken. A target pair was considered enriched in an ST if it had an FDR < 0.05 and a fraction of high-scoring tumours at least 10 percentage points greater than in the remaining tumours (Fig. 4c; Extended Data Fig. 8e; Supplementary Table 16).

Enriched pairs were ranked by this difference, with FDR used to break ties. Because STs were defined from the same pair scores, these comparisons were used to rank and describe pairs rather than as independent tests of differences between STs.

Assignment of HCC tumours to Spatial Theratopes

Target-pair scores for HCC tumours were computed and min–max scaled within each tumour as in the discovery cohort, and each tumour was assigned to the ST with the nearest discovery centroid by Euclidean distance (Extended Data Fig. 9d; Supplementary Table 21). Twenty-three of 25 tumours were assigned to ST3, one to ST2 and one to ST4. Survival analyses were restricted to ST3 tumours with complete progression-free survival data (n = 22); one ST3 patient was excluded because of missing follow-up information. Assignment confidence was assessed as described in the Supplementary Methods.

Survival analysis

Per-sample proportions of eight cellular neighbourhoods (CN1–CN8) were analysed in two cohorts. The first comprised resected ST1 lung cancers (n = 48), of which 47 with available recurrence-free survival data were included in survival analyses. The second consisted of 22 ICI-treated hepatocellular carcinomas assigned to ST3 with complete follow-up (of 25 in the validation cohort) also with PFS and mRECIST response (responders: CR/PR) as endpoints. Each CN was dichotomised as present/absent when enough patients had a proportion of zero (≥5 for lung, ≥10 for HCC), and otherwise at the median. CNs were tested only if each group had at least five patients and five events. Survival associations were estimated with Cox models, adjusted for stage and histology in the lung cohort and univariable in the HCC cohort. The proportional hazards assumption was assessed using a test based on scaled Schoenfeld residuals with rank-transformed time. P values were corrected using the Benjamini–Hochberg procedure, and Kaplan–Meier curves were compared by log-rank test (Supplementary Table 19). In the HCC cohort, CN8 proportion and the VEGF-A–PD-L1 pair score were compared between responders and non-responders using two-sided Mann–Whitney *U* tests, correlated with mRECIST category (PD < SD < PR < CR) using Spearman's rank correlation (two-sided), and tested for association with PFS in univariable Cox models, with each variable modelled as median-split and continuous (pair score per s.d.; CN8 per $\log_{10}$ unit). Cox *P* values were from two-sided Wald tests. Analyses used Python (v3.11.14) with lifelines (v0.30.3), SciPy (v1.16.2) and statsmodels (v0.14.5).

Data availability

Raw and processed Xenium data, PhenoCycler and IMC images, cell-type annotations, CN assignments and de-identified clinical metadata will deposited in DNA Data Bank of Japan [accession number-TBC] and will be made publicly available upon publication. Source data are provided with this paper.

Code availability

Code for all analyses, including ASTRA-CYTE, is available at https://github.com/OncofetalEcosystemLab/ASTRA1.0_analysis and archived at Zenodo ([DOI to be provided upon publication]).

References

1. Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discov.* **12**, 31–46 (2022).
2. Binnewies, M. *et al.* Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat. Med.* **24**, 541–550 (2018).
3. Anderson, N. M. & Simon, M. C. The tumor microenvironment. *Curr. Biol.* **30**, R921–R925 (2020).
4. Schürch, C. M. *et al.* Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. *Cell* **182**, 1341-1359.e19 (2020).
5. Nirmal, A. J. *et al.* The Spatial Landscape of Progression and Immunoediting in Primary Melanoma at Single-Cell Resolution. *Cancer Discov.* **12**, 1518–1541 (2022).
6. Li, Z. *et al.* Presence of onco-fetal neighborhoods in hepatocellular carcinoma is associated with relapse and response to immunotherapy. *Nat. Cancer* **5**, 167–186 (2024).
7. Wang, X. Q. *et al.* Spatial predictors of immunotherapy response in triple-negative breast cancer. *Nature* **621**, 868–876 (2023).
8. Bagaev, A. *et al.* Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* **39**, 845-865.e7 (2021).
9. Jia, G. *et al.* Spatial immune scoring system predicts hepatocellular carcinoma recurrence. *Nature* **640**, 1031–1041 (2025).
10. Li, J. *et al.* Pan-cancer analysis of spatial transcriptomics reveals heterogeneous tumor spatial microenvironment. *Cell Rep. Med.* **7**, 102751 (2026).
11. Liu, Y. *et al.* Conserved spatial subtypes and cellular neighborhoods of cancer-associated fibroblasts revealed by single-cell spatial multi-omics. *Cancer Cell* **43**, 905-924.e6 (2025).
12. Cho, K. S. *et al.* Pan-cancer spatial atlas of tertiary lymphoid structures. *Science* **392**, eadz2742 (2026).
13. Behera, V. *et al.* Pan-tumor spatial transcriptomics reveals conserved properties of tumor organization. *Cell Rep. Med.* **7**, (2026).
14. Tateishi, M. *et al.* Molecular characterization of large cell neuroendocrine carcinoma of the lung by multilayered omics analysis. *Nat. Commun.* https://doi.org/10.1038/s41467-026-78485-x (2026) doi:10.1038/s41467-026-78485-x.
15. Bill, R. *et al.* CXCL9:SPP1 macrophage polarity identifies a network of cellular programs that control human cancers. *Science* **381**, 515–524 (2023).

16. Öhlund, D. *et al.* Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J. Exp. Med.* **214**, 579–596 (2017).
17. Moll, R., Divo, M. & Langbein, L. The human keratins: biology and pathology. *Histochem. Cell Biol.* **129**, 705–733 (2008).
18. Vocino Trucco, G., Righi, L., Volante, M. & Papotti, M. Updates on lung neuroendocrine neoplasm classification. *Histopathology* **84**, 67–85 (2024).
19. Rindi, G. *et al.* A common classification framework for neuroendocrine neoplasms: an International Agency for Research on Cancer (IARC) and World Health Organization (WHO) expert consensus proposal. *Mod. Pathol.* **31**, 1770–1786 (2018).
20. Sharma, A. *et al.* Onco-fetal Reprogramming of Endothelial Cells Drives Immunosuppressive Macrophages in Hepatocellular Carcinoma. *Cell* **183**, 377-394.e21 (2020).
21. George, J. *et al.* Integrative genomic profiling of large-cell neuroendocrine carcinomas reveals distinct subtypes of high-grade neuroendocrine lung tumors. *Nat. Commun.* **9**, 1048 (2018).
22. Byers, L. A. *et al.* Proteomic Profiling Identifies Dysregulated Pathways in Small Cell Lung Cancer and Novel Therapeutic Targets Including PARP1. *Cancer Discov.* **2**, 798–811 (2012).
23. Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer–immune set point. *Nature* **541**, 321–330 (2017).
24. Nieto, M. A., Huang, R. Y.-J., Jackson, R. A. & Thiery, J. P. EMT: 2016. *Cell* **166**, 21–45 (2016).
25. Sirica, A. E. & Gores, G. J. Desmoplastic Stroma and Cholangiocarcinoma: Clinical Implications and Therapeutic Targeting. *Hepatol. Baltim. Md* **59**, 2397–2402 (2014).
26. Korc, M. Pancreatic cancer–associated stroma production. *Am. J. Surg.* **194**, S84–S86 (2007).
27. Ben-Porath, I. *et al.* An embryonic stem cell–like gene expression signature in poorly differentiated aggressive human tumors. *Nat. Genet.* **40**, 499–507 (2008).
28. Schumacher, T. N. & Thommen, D. S. Tertiary lymphoid structures in cancer. *Science* **375**, eabf9419 (2022).
29. Siliņa, K. *et al.* Germinal Centers Determine the Prognostic Relevance of Tertiary Lymphoid Structures and Are Impaired by Corticosteroids in Lung Squamous Cell Carcinoma. *Cancer Res.* **78**, 1308–1320 (2018).
30. Sautès-Fridman, C., Petitprez, F., Calderaro, J. & Fridman, W. H. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat. Rev. Cancer* **19**, 307–325 (2019).
31. Vanhersecke, L. *et al.* Mature tertiary lymphoid structures predict immune checkpoint inhibitor efficacy in solid tumors independently of PD-L1 expression. *Nat. Cancer* **2**, 794–802 (2021).
32. Cheng, S. *et al.* A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell* **184**, 792-809.e23 (2021).
33. Miller, T. E. *et al.* Programs, origins and immunomodulatory functions of myeloid cells in glioma. *Nature* **640**, 1072–1082 (2025).
34. Ramos, R. N. *et al.* Tissue-resident FOLR2+ macrophages associate with CD8+ T cell infiltration in human breast cancer. *Cell* **185**, 1189-1207.e25 (2022).
35. Dayton, P. K. & Hessler, R. R. Role of biological disturbance in maintaining

diversity in the deep sea. *Deep Sea Res. Oceanogr. Abstr.* **19**, 199–208 (1972).
36. Liu, J. *et al.* A modular framework for antibody–drug conjugate biomarkers. *Nat. Rev. Clin. Oncol.* 1–20 (2026) doi:10.1038/s41571-026-01213-3.
37. Pistilli, B., Colombo, R., Mosele, M. F. & Chandarlapaty, S. Antibody-drug conjugates in precision oncology. *Cancer Cell* **44**, 1126–1146 (2026).
38. Finn, R. S. *et al.* Atezolizumab plus Bevacizumab in Unresectable Hepatocellular Carcinoma. *N. Engl. J. Med.* **382**, 1894–1905 (2020).
39. Makker, V. *et al.* Lenvatinib plus Pembrolizumab for Advanced Endometrial Cancer. *N. Engl. J. Med.* **386**, 437–448 (2022).
40. Rini, B. I. *et al.* Pembrolizumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma. *N. Engl. J. Med.* **380**, 1116–1127 (2019).
41. Aaltonen, L. A. *et al.* Pan-cancer analysis of whole genomes. *Nature* **578**, 82–93 (2020).
42. Popejoy, A. B. & Fullerton, S. M. Genomics is failing on diversity. *Nature* **538**, 161–164 (2016).
43. Ginhoux, F. *et al.* Fate Mapping Analysis Reveals That Adult Microglia Derive from Primitive Macrophages. *Science* **330**, 841–845 (2010).
44. Mariathasan, S. *et al.* TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* **554**, 544–548 (2018).
45. IARC Publications Website - WHO Classification of Tumours. https://publications.iarc.who.int/Book-And-Report-Series/Who-Classification-Of-Tumours.
46. Janesick, A. *et al.* High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. *Nat. Commun.* **14**, 8353 (2023).
47. Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968-981.e15 (2018).
48. Giesen, C. *et al.* Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11**, 417–422 (2014).
49. Eapen, M. M. *et al.* An Integrated Spatial Multi-Omics Workflow for Sequential RNA and Protein Profiling in FFPE Tumor Tissue. *Cancer Res. Commun.* **6**, 1963–1975 (2026).
50. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* **57**, 289–300 (1995).
51. Aran, D. *et al.* Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
52. Mabbott, N. A., Baillie, J. K., Brown, H., Freeman, T. C. & Hume, D. A. An expression atlas of human primary cells: inference of gene function from coexpression networks. *BMC Genomics* **14**, 632 (2013).
53. Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* **42**, 293–304 (2024).
54. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (2016). doi:10.1145/2939672.2939785.
55. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. Preprint at https://doi.org/10.48550/arXiv.1802.03426 (2020).

56. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
57. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
58. Andreatta, M. & Carmona, S. J. UCell: Robust and scalable single-cell gene signature scoring. *Comput. Struct. Biotechnol. J.* **19**, 3796–3798 (2021).
59. Sibai, M. *et al.* The spatial landscape of cancer hallmarks reveals patterns of tumor ecological dynamics and drug sensitivity. *Cell Rep.* **44**, 115229 (2025).
60. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
61. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. Preprint at https://doi.org/10.48550/arXiv.1201.0490 (2018).

Acknowledgements

ASTRA is supported by the Adopting Sustainable Partnerships for Innovative Research Ecosystem (ASPIRE) program grant to A. Sharma, Y.S. and H.N., jointly funded by the National Health and Medical Research Council (NHMRC; GNT2043672) and the Japan Agency for Medical Research and Development (AMED; JP25jf0126017). A.M., J.G. and L.C. are also supported by this grant. A. Sharma is supported by a CSL Centenary Fellowship, an NHMRC Ideas Grant (GNT2010795), the Medical Research Future Fund (MRFF; MRF2016215), an NHMRC Clinical Trials and Cohort Studies Grant (GNT2032407) and catalytic funding from the Kinghorn Foundation. Y.S. is supported by AMED ASPIRE (JP23jf0126003) and the Japan Society for the Promotion of Science (JSPS) KAKENHI Platform for Advanced Genome Science (PAGS; JP22H04925). A. Suzuki is supported by the AMED Project for Promotion of Cancer Research and Therapeutic Evolution (P-PROMOTE; JP25ama221548), H.N. by the Uehara Memorial Foundation and D.M. by JSPS KAKENHI (JP26K23362). We thank the patients and their families. We thank the 10x Genomics and Standard BioTools teams for technical input on panel design. We acknowledge the National Computational Infrastructure (NCI Australia) for computational resources. The supercomputing resource was provided by the Human Genome Center, the Institute of Medical Science, the University of Tokyo. We thank V. Deswaerte for critical reading of the manuscript.

ASTRA Consortium

Jacob George[29], Florent Ginhoux[32,33,34], Ruby Yun-Ju Huang[13], Mitsuho Imai[35], Stephanie Ma[36,37], Archita Mishra[16,38], Hidewaki Nakagawa[22], Woong-Yang Park[39,40], Ankur Sharma[1,4,31] & Yutaka Suzuki[2]

[32]Université Paris-Saclay, Gustave Roussy, U1356 Next Generation Immuno-Oncology Research, Villejuif, France

[33]Translational Immunology Institute, SingHealth Duke-NUS Academic Medical Centre, Singapore, Singapore

[34]Shanghai Institute of Immunology, Department of Immunology and Microbiology, and Center for Human Translational Immunology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[35]Translational Research Support Office, National Cancer Center Hospital East, Chiba, Japan

[36]School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

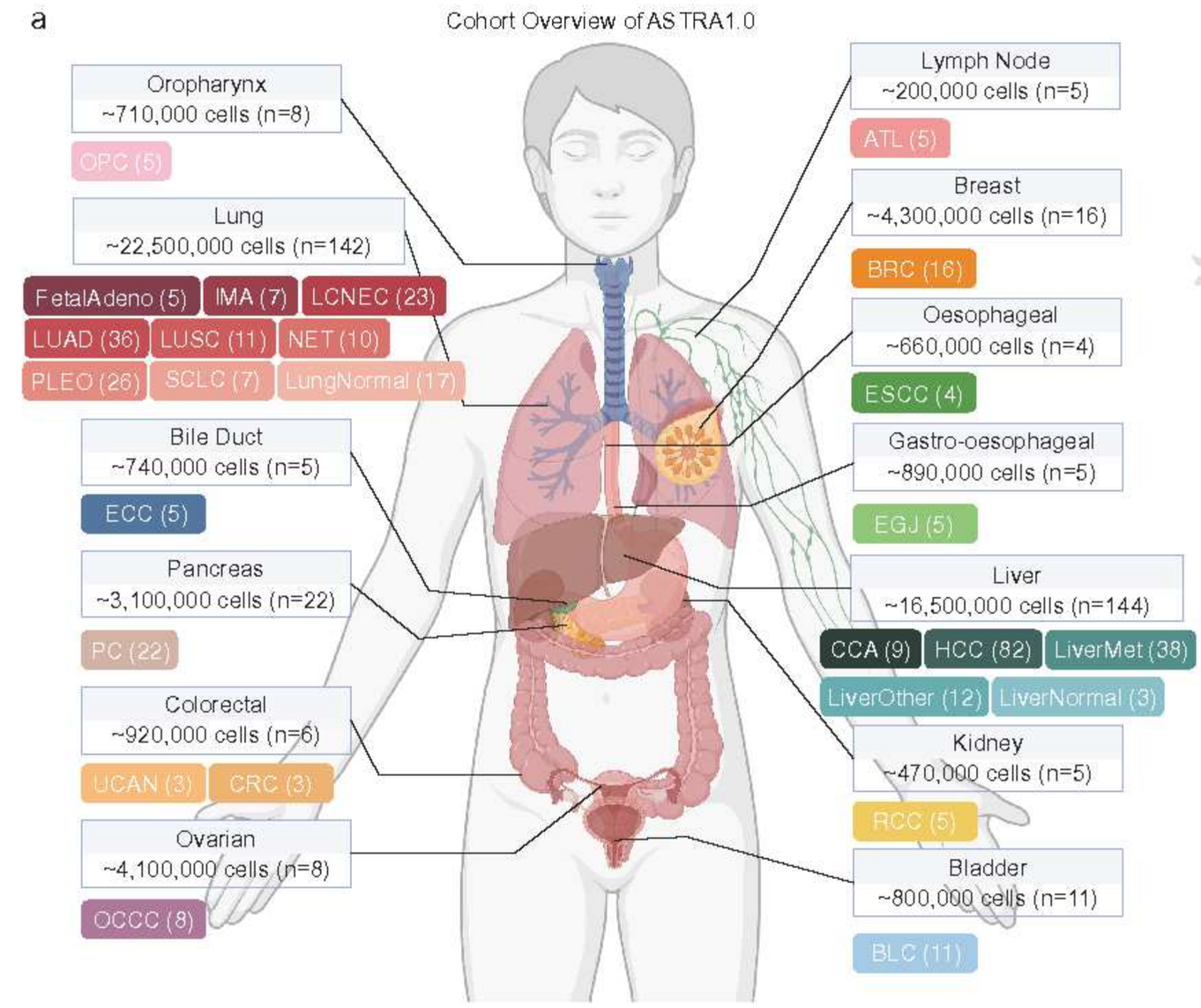
[37]State Key Laboratory of Liver Research, The University of Hong Kong, Hong Kong SAR, China

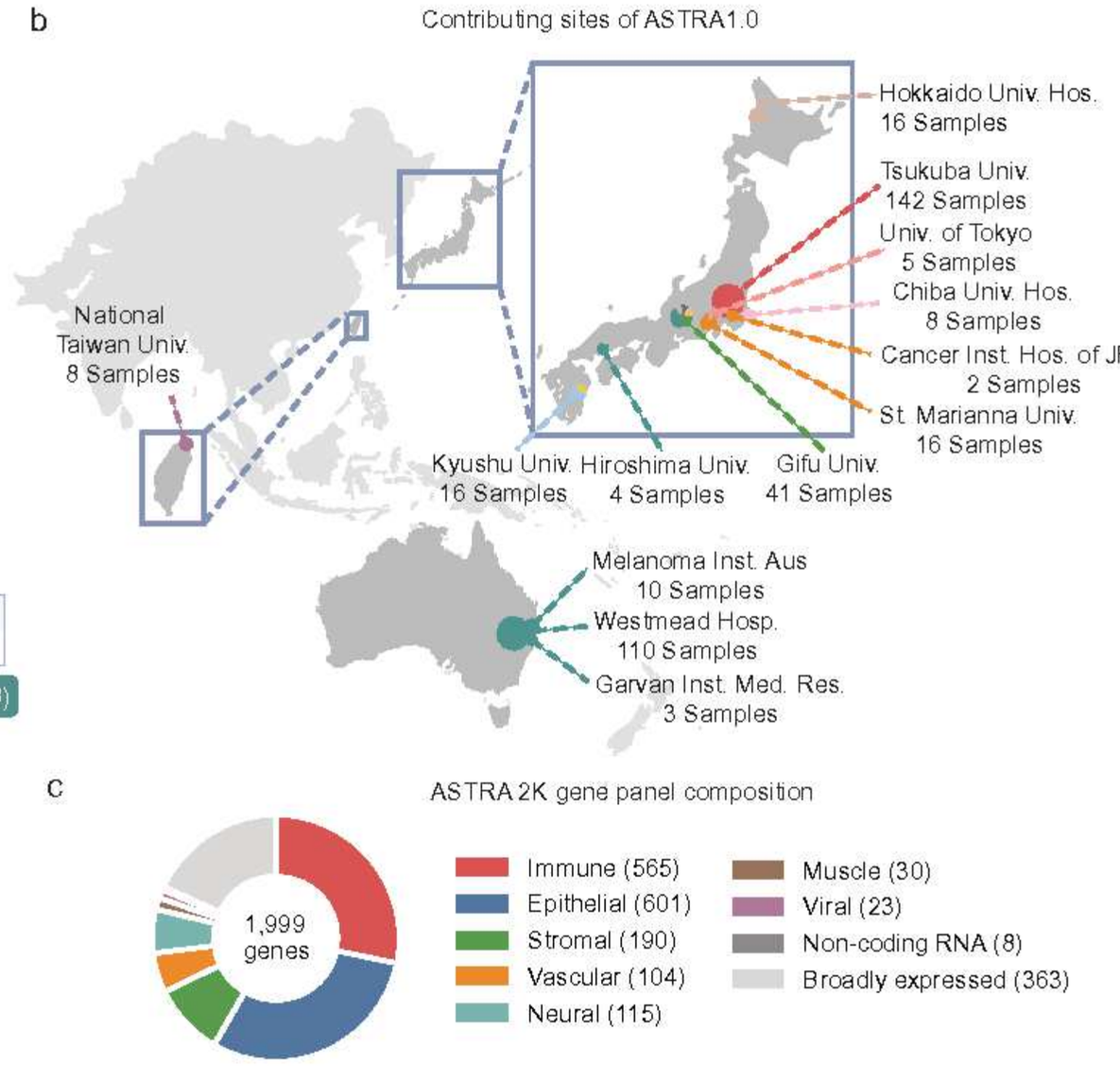
[38]Sydney Infectious Diseases Institute (Sydney ID), Faculty of Medicine and Health, The University of Sydney, Sydney, NSW, Australia

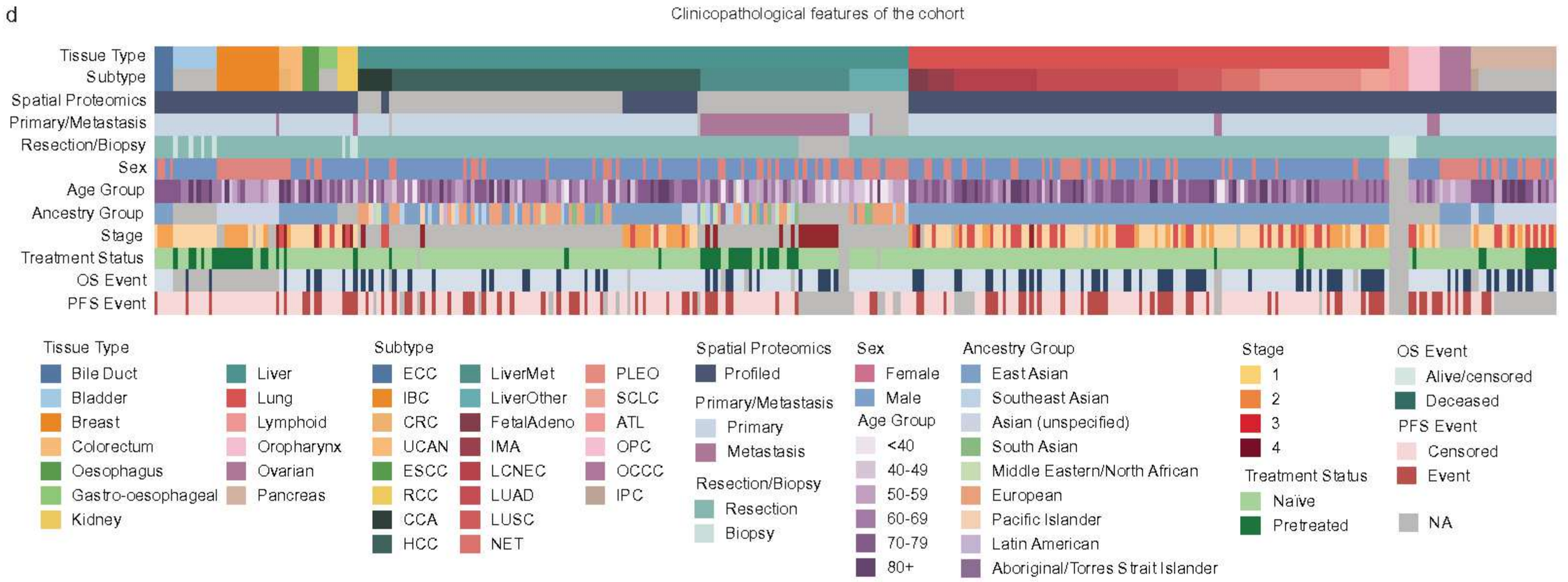
[39]Department of Health Science and Technology, Samsung Advanced Institute of Health Sciences and Technology (SAIHST), Sungkyunkwan University, Seoul, Republic of Korea

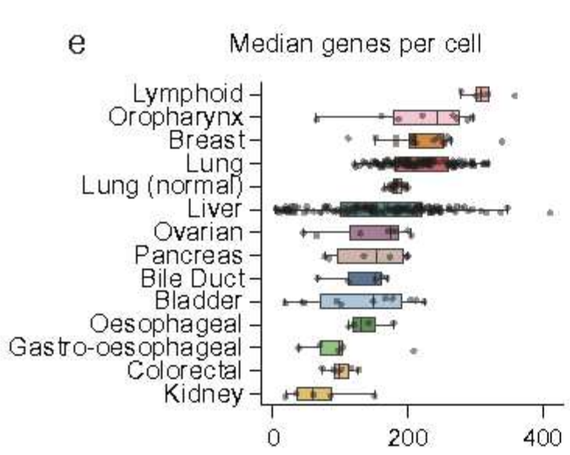
[40]Department of Precision Medicine, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

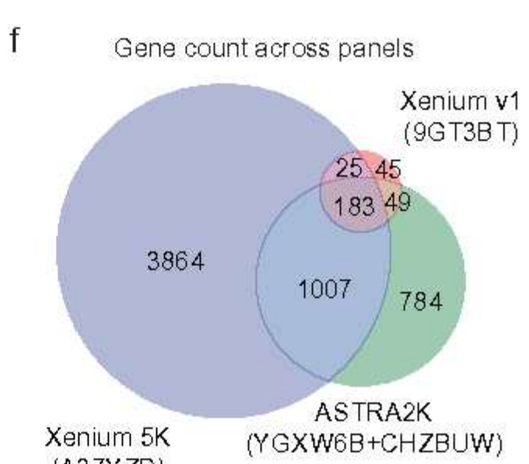
Fig. 1 | Asia-Pacific spatial multimodal atlas of human cancers.

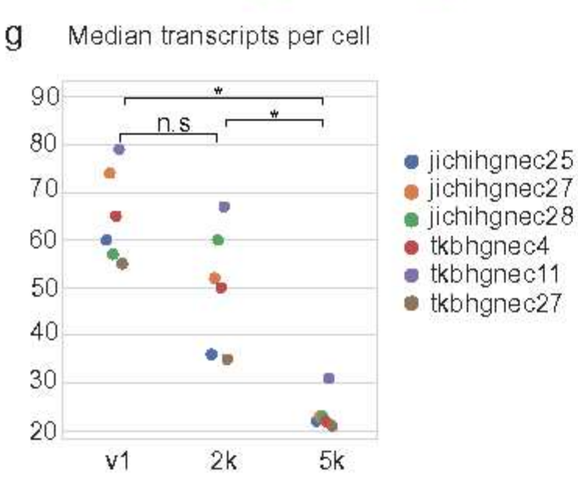


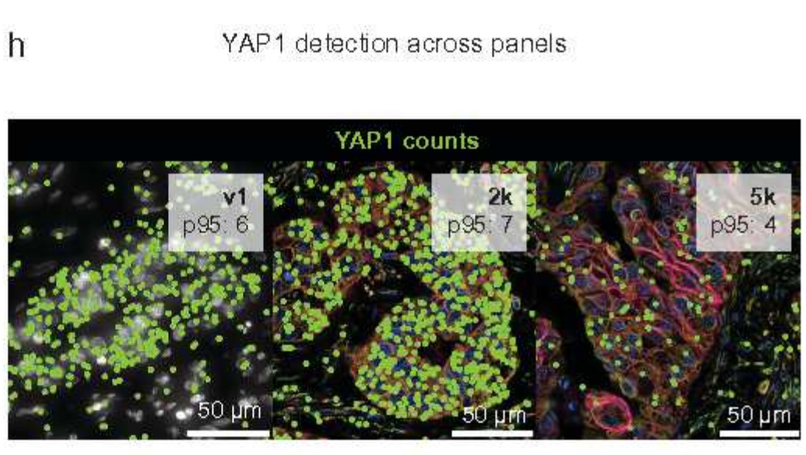












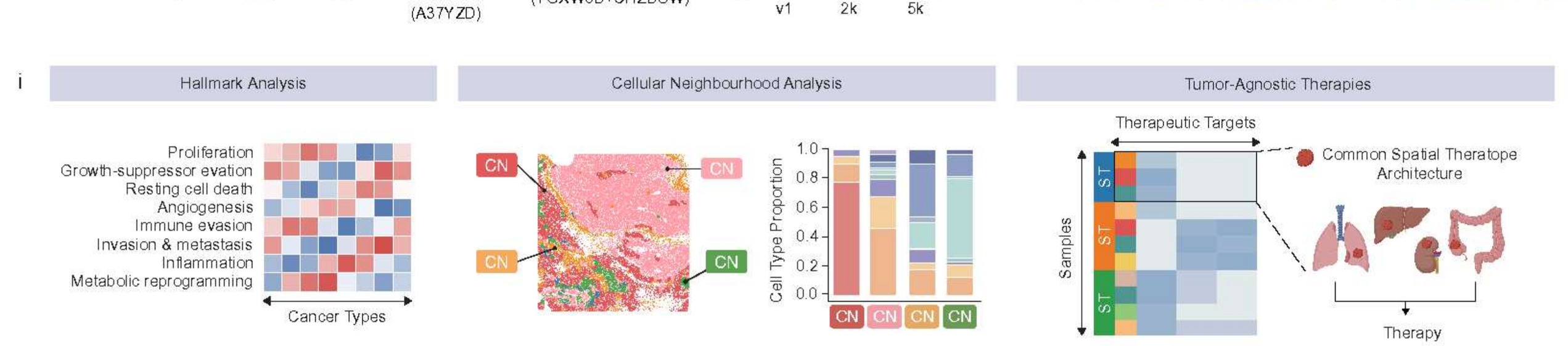
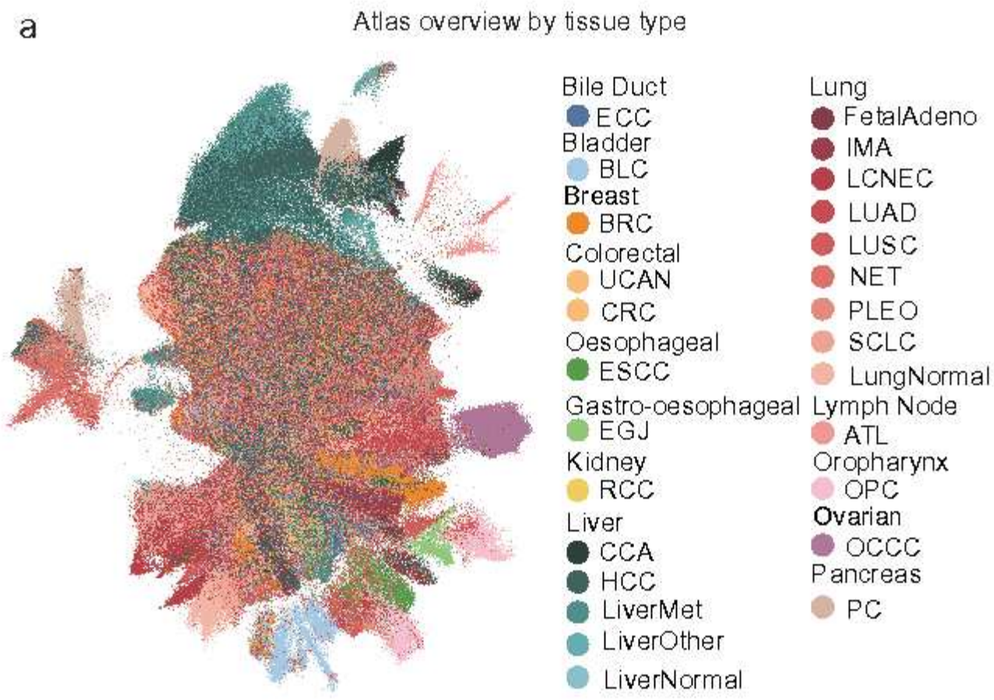
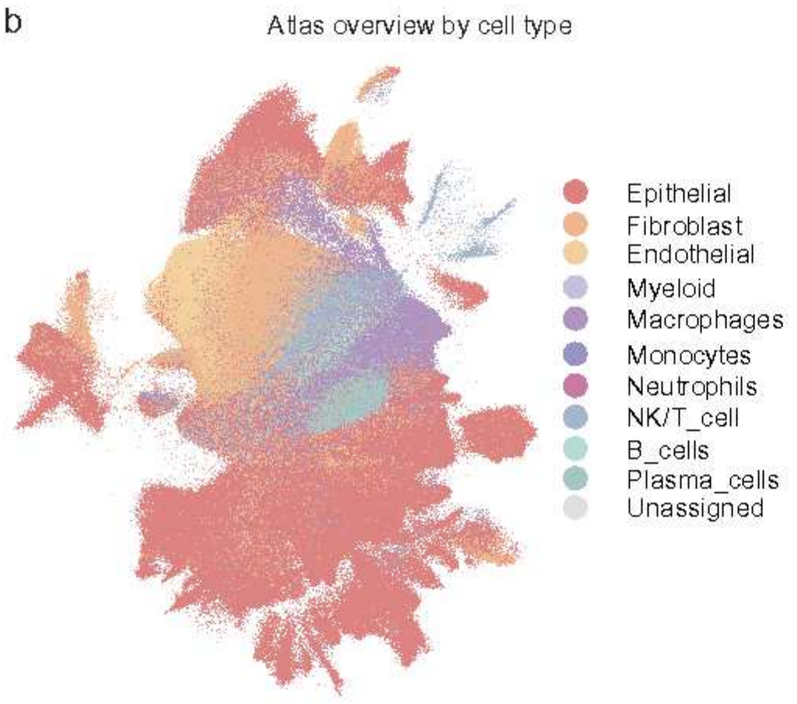
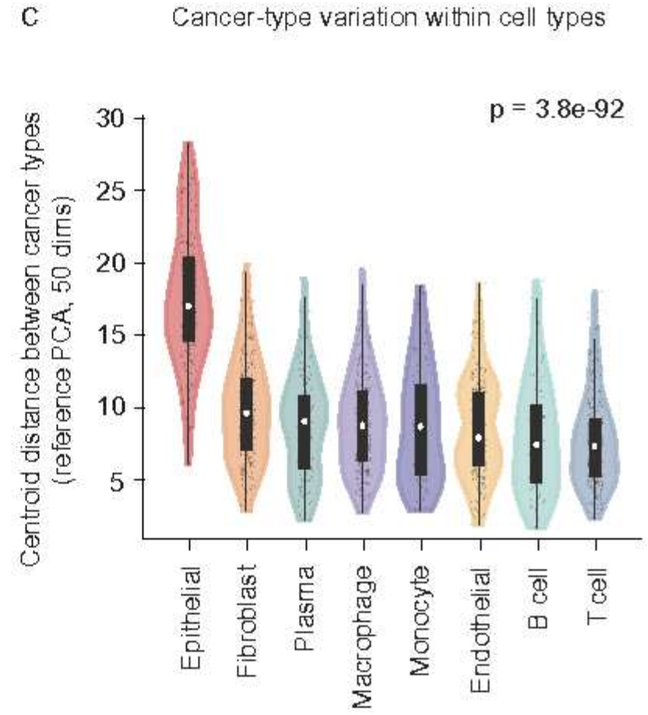
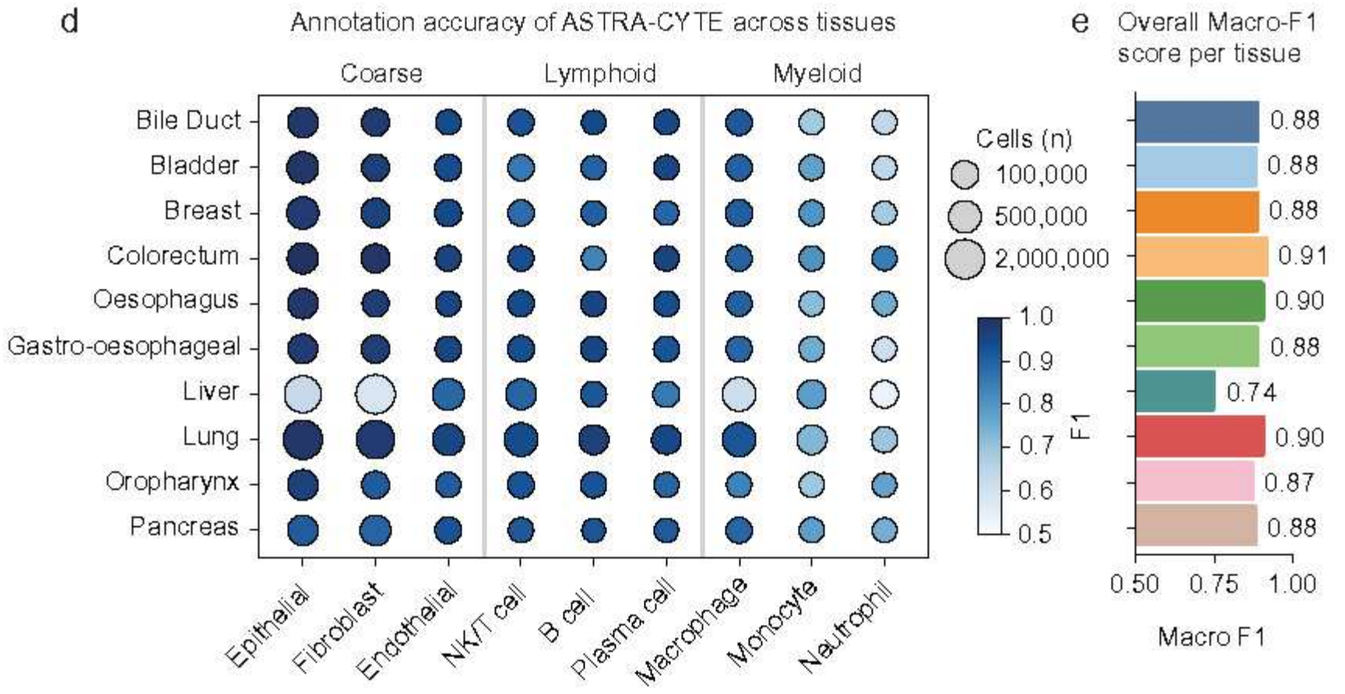


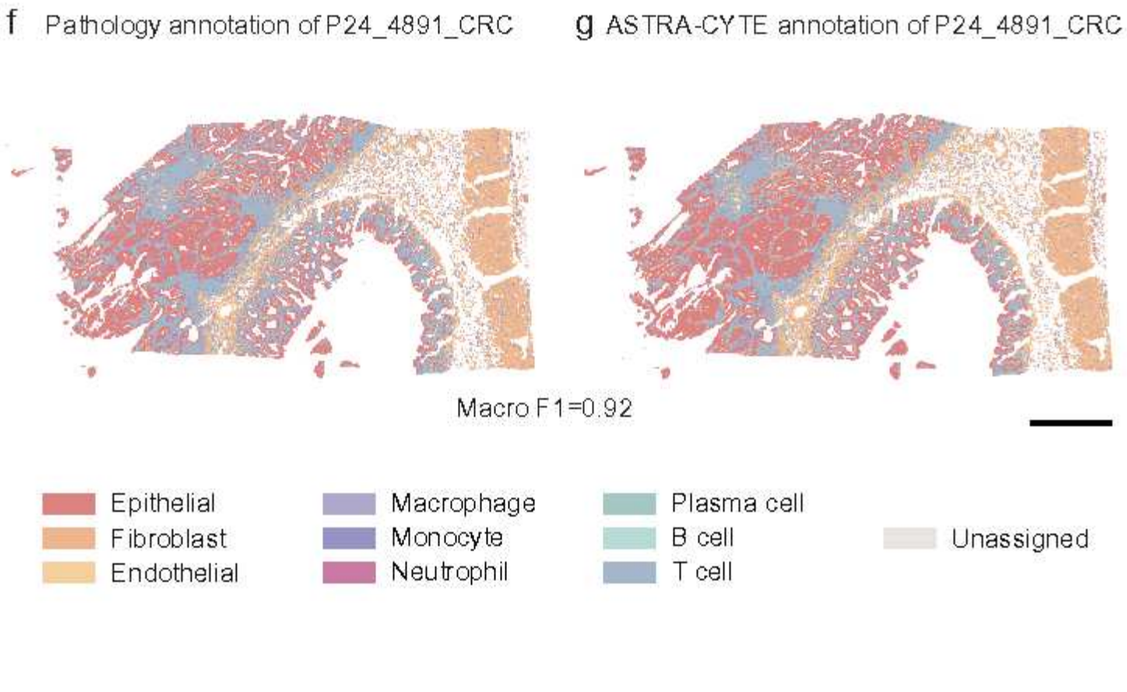
Fig. 2 | A pan-cancer spatial single-cell atlas and automated cell-type annotation with ASTRA-CYTE.

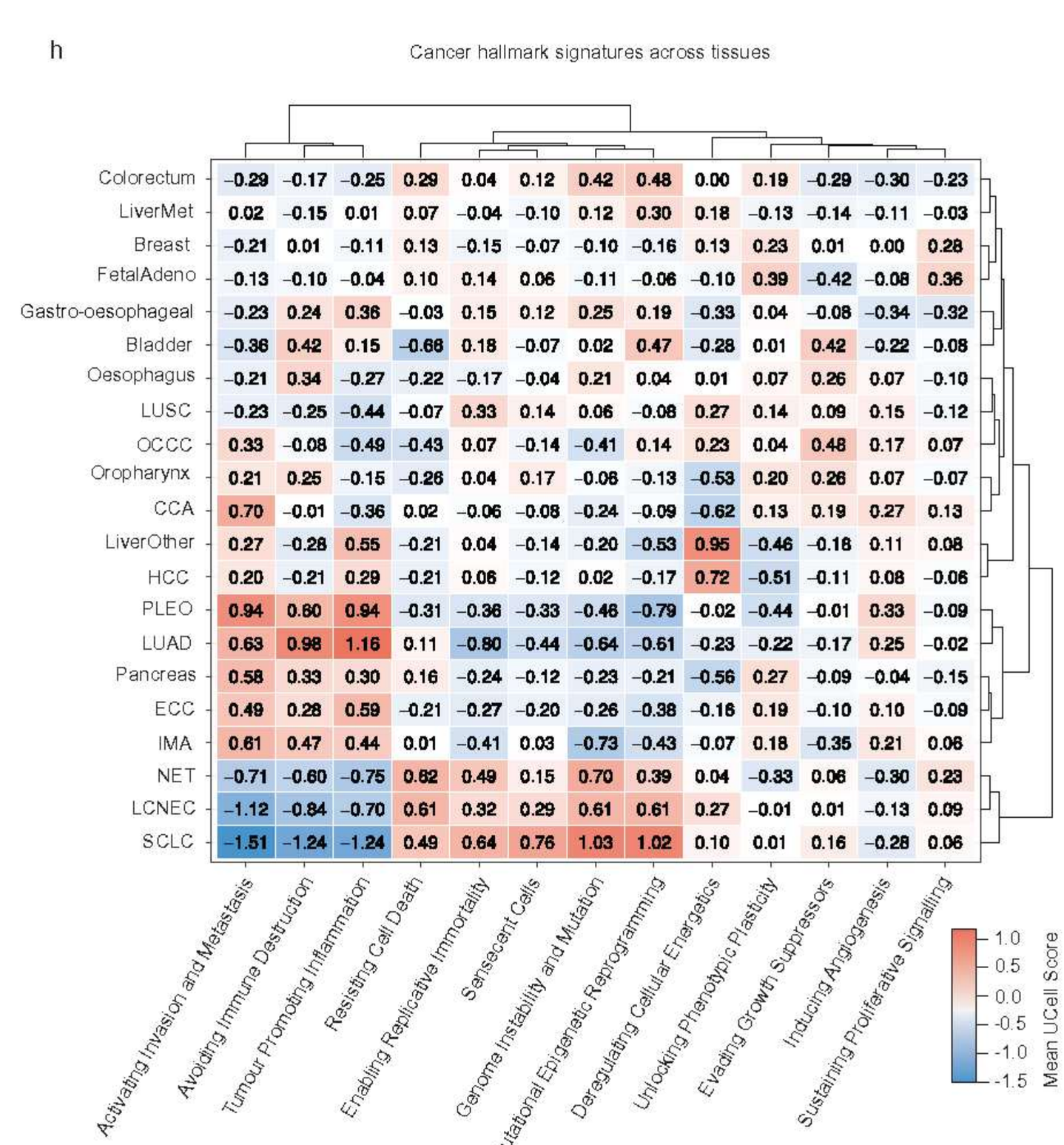












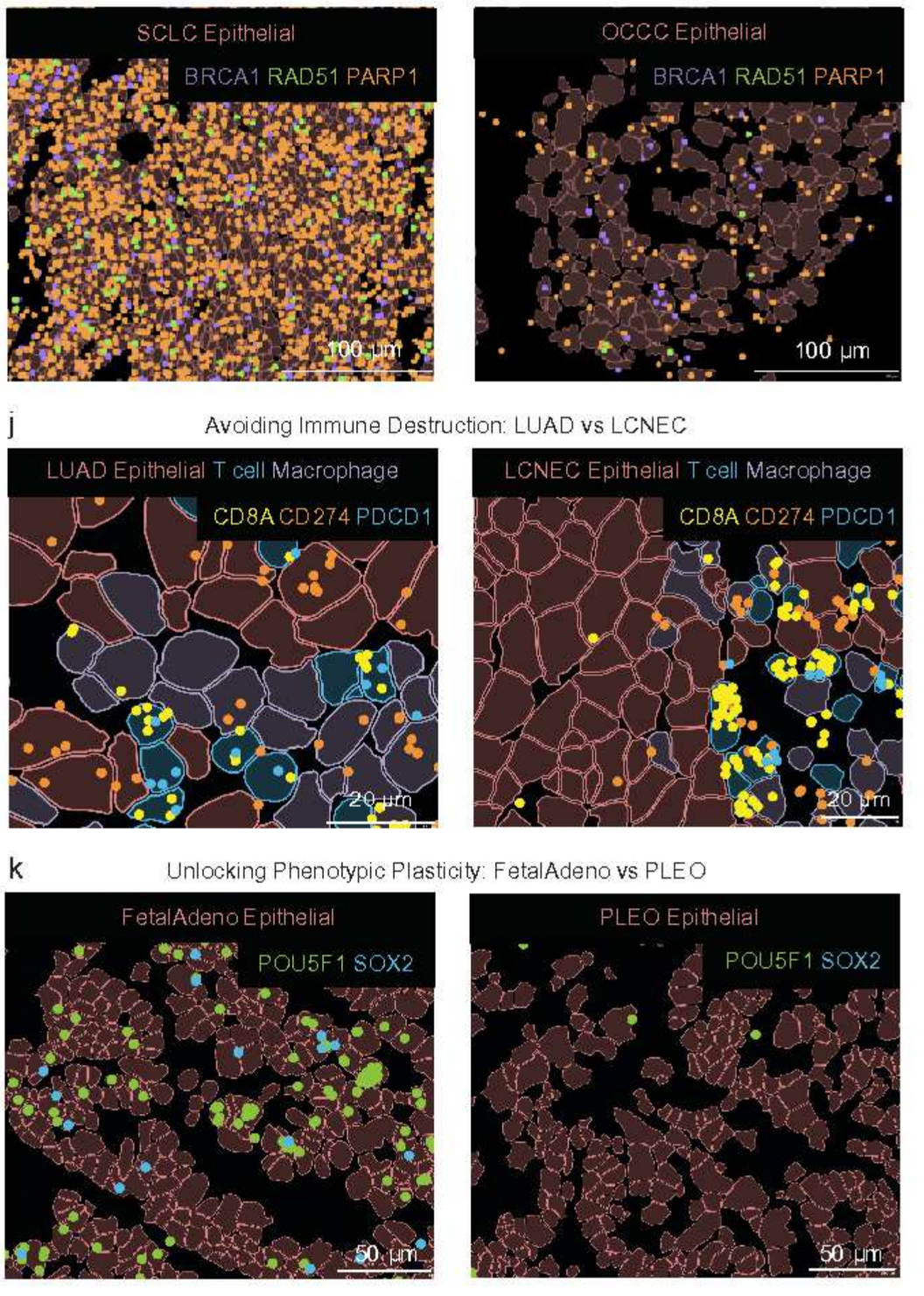


Fig. 3 | Recurrent cellular neighbourhoods and conserved CN-induced cell states across tissue types.

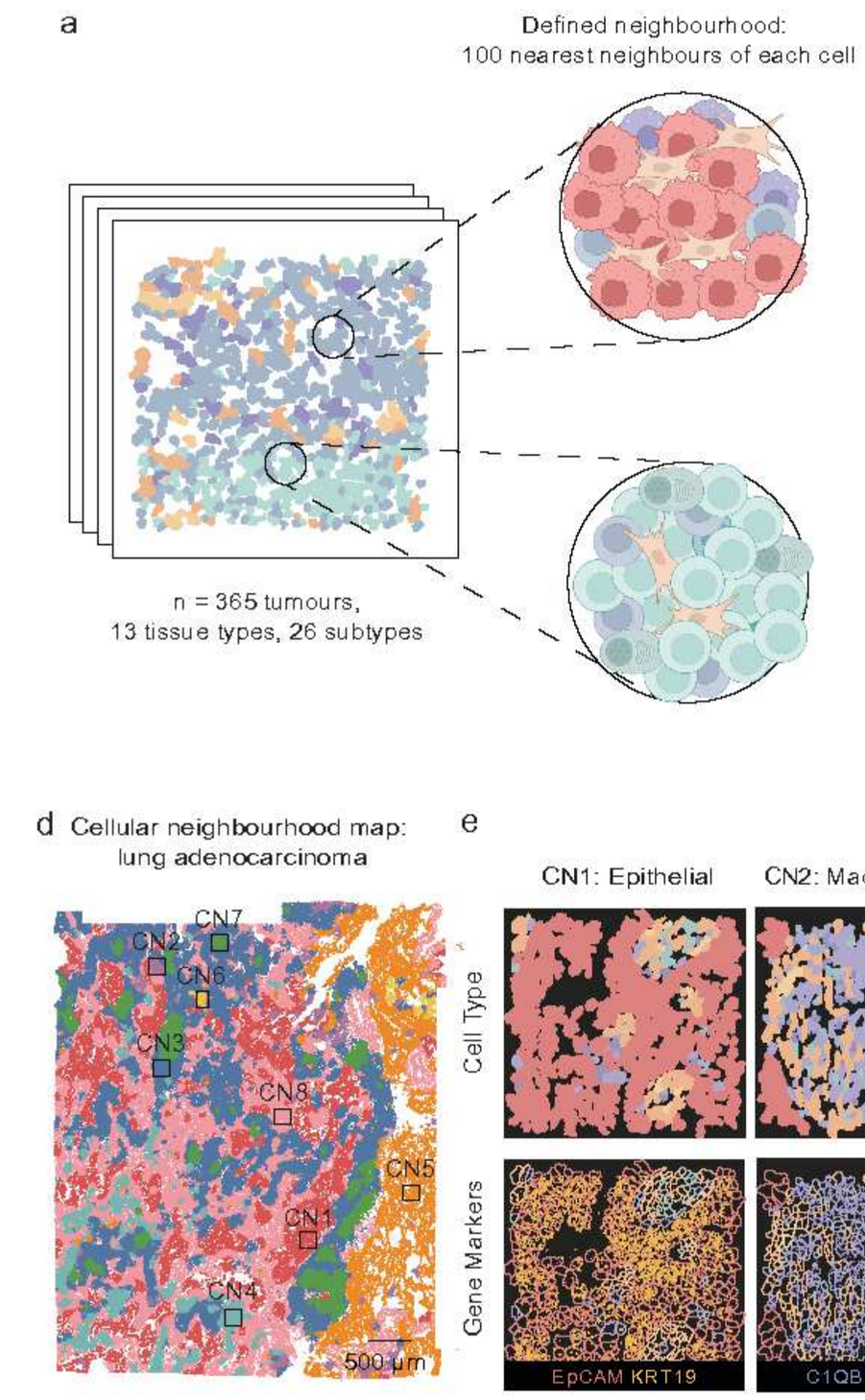
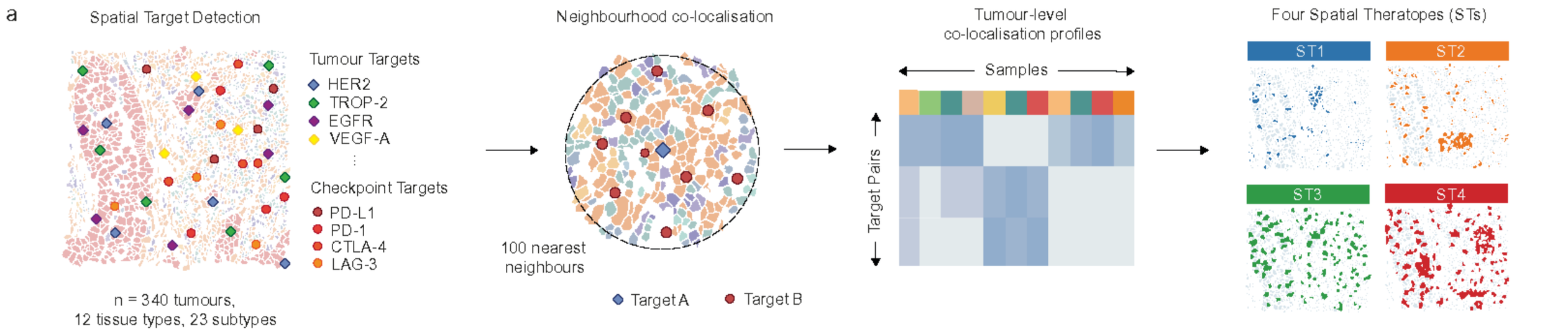


Fig. 4 | Spatial theratopes identify co-targetable tumour–immune interactions and niches associated with clinical outcome.



Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ExtendedDataFig6.pdf
- ExtendedDataFigureLegends.pdf
- ExtendedDataFig7.pdf
- ExtendedDataFig5.pdf
- ExtendedDataFig1.pdf
- ExtendedDataFig9.pdf
- ExtendedDataFig4.pdf
- SupplementaryTables.xlsx
- ExtendedDataFig3.pdf
- ExtendedDataFig8.pdf
- SupplementaryMethods.docx
- ExtendedDataFig2.pdf