Article

The single-cell pathology landscape of breast cancer

Hartland W. Jackson[1,2,10], Jana R. Fischer[1,2,3,10], Vito R. T. Zanotelli[1,2,3], H. Raza Ali[1,2,4], Robert Mechera[5], Savas D. Soysal[6,7], Holger Moch[8], Simone Muenst[9], Zsuzsanna Varga[8], Walter P. Weber[5] & Bernd Bodenmiller[1,2*]

Single-cell analyses have revealed extensive heterogeneity between and within human tumours[1–4], but complex single-cell phenotypes and their spatial context are not at present reflected in the histological stratification that is the foundation of many clinical decisions. Here we use imaging mass cytometry[5] to simultaneously quantify 35 biomarkers, resulting in 720 high-dimensional pathology images of tumour tissue from 352 patients with breast cancer, with long-term survival data available for 281 patients. Spatially resolved, single-cell analysis identified the phenotypes of tumour and stromal single cells, their organization and their heterogeneity, and enabled the cellular architecture of breast cancer tissue to be characterized on the basis of cellular composition and tissue organization. Our analysis reveals multicellular features of the tumour microenvironment and novel subgroups of breast cancer that are associated with distinct clinical outcomes. Thus, spatially resolved, single-cell analysis can characterize intratumour phenotypic heterogeneity in a disease-relevant manner, with the potential to inform patient-specific diagnosis.

Histological and phenotypical differences between tumours guide cancer diagnosis, prognosis and the selection of treatment. At present, breast cancers are graded on the basis of tumour structure and cellular morphology, and subcategorized when more than 1% of tumour cells contain hormone receptors or more than 10% express high levels of HER2 protein or exhibit amplification of the *HER2* gene[6–8]. This leaves a large portion of cells uncharacterized, despite the fact that additional molecular subclasses and morphological features have previously been identified as prognostic[9–12]. It is clear that clonal evolution and spatially distinct tumour microenvironments drive inter- and intrapatient cellular heterogeneity and hinder effective treatment[2–4,13,14]. Using highly multiplexed imaging[5,15,16], multiple complex cellular phenotypes have been identified within the context of the tumour microenvironment, and this has enabled refined histopathological classification of clinical tissue samples[17–20]. Here, using a topological single-cell network analysis of high-dimensional mass cytometry images, we quantified spatial inter- and intratumour heterogeneity on a single-cell level and identified spatially resolved features and novel subtypes of breast cancer that are associated with clinical outcome.

Spatially resolved single-cell phenotypes

To comprehensively quantify the cellular heterogeneity and spatial organization of breast cancer tissue, we designed an imaging mass cytometry (IMC) panel specific to breast histology (Extended Data Fig. 1) and used this to image samples from 281 tumours that represent all clinical subtypes and grades of pathology (Supplementary Table 1). IMC combines immunohistochemistry staining using metal isotope-labelled antibodies with laser ablation and mass-spectrometry-based detection to produce high-dimensional images[5] (Fig. 1a). Our panel of 35 antibodies simultaneously quantified the clinically established breast cancer targets oestrogen receptor (ER), progesterone receptor (PR) and HER2; the proliferation marker Ki-67; markers of epithelial, mesenchymal, immune and endothelial lineages; and targets that provide insight into signalling pathways, oncogenes and epigenetics (Extended Data Fig. 1, Supplementary Table 2). IMC generates images that are comparable to those produced by immunofluorescence or immunohistochemistry, but that have the capacity for highly multiplexed staining[5,21] (Extended Data Figs. 2, 3). Images were segmented into single cells and tumour and stromal regions using a random forest pixel classifier (Ilastik) and CellProfiler[17,20,22]. We identified 855,668 cells in 381 images (289 tumour, 87 healthy breast and 5 liver controls), and quantified the expression of both marker genes and the spatial features of each cell (Fig. 1a). Clustering with PhenoGraph[23] identified cell phenotype clusters (hereafter referred to as phenotypes) of endothelial, T and B cell, macrophage and stromal cell populations, as well as 59 diverse tumour cell phenotypes. As was previously observed[1], some tumour phenotypes were unique to individual patients (Fig. 1b, c, Extended Data Fig. 4a). To identify common cellular subtypes within this spectrum, we defined 14 tumour cell metaclusters by hierarchical clustering of the tumour single-cell phenotypes defined by PhenoGraph (Fig. 1c, Extended Data Fig. 5a).

Tumours from every clinical subtype contained populations of fibroblasts, endothelial and immune cells at similar densities, but were enriched in populations of tumour cells that showed variable expression

[1]Department of Quantitative Biomedicine, University of Zurich, Zurich, Switzerland. [2]Institute of Molecular Life Sciences, University of Zurich, Zurich, Switzerland. [3]Life Science Zurich Graduate School, ETH Zurich and University of Zurich, Zurich, Switzerland. [4]CRUK Cambridge Institute, University of Cambridge, Cambridge, UK. [5]Department of Surgery, University of Basel and University Hospital Basel, Basel, Switzerland. [6]Department of Biomedicine, University of Basel, Basel, Switzerland. [7]Visceral Surgery Research Laboratory, Department of Surgery, Clarunis University Center for Gastrointestinal and Liver Diseases Basel, Basel, Switzerland. [8]Institute of Pathology and Molecular Pathology, University Hospital Zurich, Zurich, Switzerland. [9]Institute of Pathology and Genetics, University Hospital Basel, Basel, Switzerland. [10]These authors contributed equally: Hartland W. Jackson, Jana R. Fischer. *e-mail: bernd.bodenmiller@uzh.ch
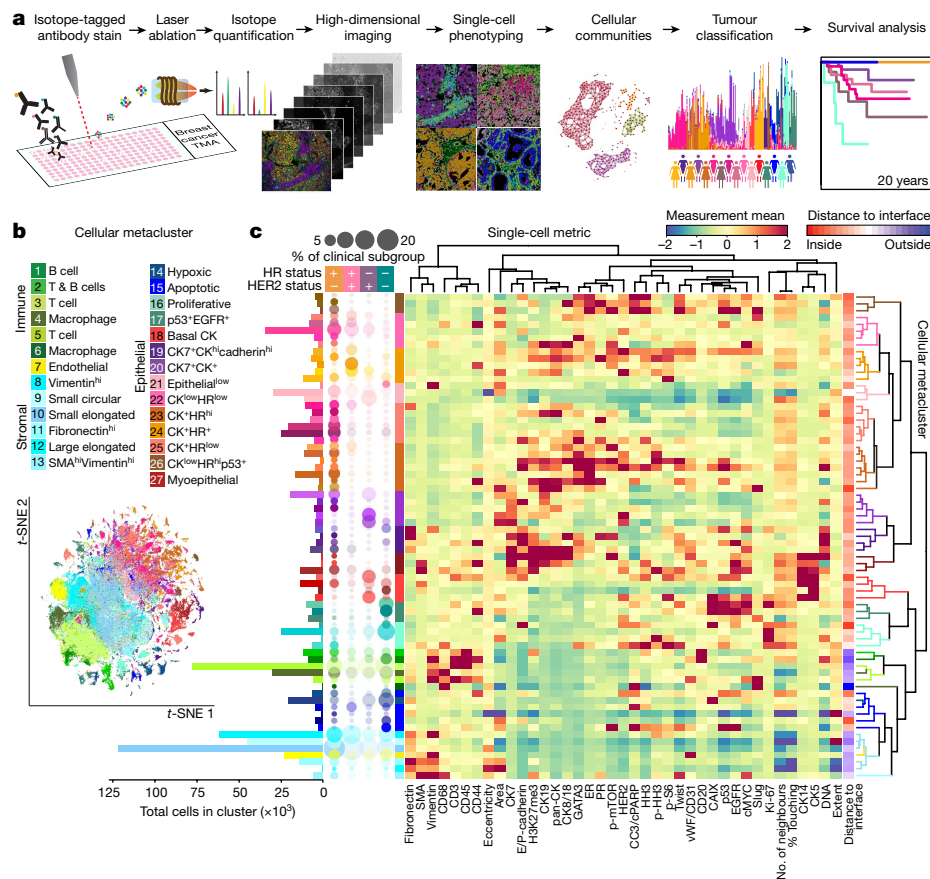
Article

Fig. 1 | Single-cell phenotypes in high-dimensional histopathology of breast cancer. a, Schematic of IMC acquisition of multiplexed images from 281 patients with breast cancer and the analyses of single-cell phenotypes, metaclusters, stromal-cell organization and architecture, tumour and patient subclassification and patient overall survival. **b**, Map using *t*-distributed stochastic neighbour embedding (*t*-SNE) of 171,288 subsampled single cells from high-dimensional images of breast tumours coloured by cell-type

metacluster identifier. **c**, Heat map showing the *z*-scored mean marker expression or distance to tumour–stroma interface for each PhenoGraph cluster, coloured by metacluster identifier. The absolute cell counts of each PhenoGraph cluster are displayed as a bar plot (left). In the bubble plot, circle size shows the relative proportion of all cells in a clinical subtype that come from each cluster, and circle opacity shows the proportion of each cluster that is present in the different clinical subtypes.

levels of cytokeratins, hormone receptors and HER2, reflective of clinical subtype (Fig. 1c, Extended Data Fig. 6a). Across all patients, immune cells were excluded from the cohesive tumour mass—although immune cells and fibroblasts did occasionally infiltrate the tumour mass and rare HR$^{low/-}$ cells (that is, cells with low or no expression of hormone receptors) lacking cytokeratins invaded past the tumour–stroma front in some samples (Fig. 1c, Extended Data Fig. 4). Tumour regions contained various luminal HR$^+$ epithelial cell phenotypes that were identified by combinations of expression of ER, PR, GATA3, E-cadherin and multiple cytokeratins, but hormone receptors were also expressed without cytokeratins in a few cases (metacluster 26) (Fig. 1c, Extended Data Fig. 6b). Of the luminal cytokeratins (CK7, CK8/18 and CK19), only CK7 was associated with specific luminal subsets of tumour cells (metaclusters 19 and 20) (Fig. 1c). Expression of HER2 was not a defining feature of metaclusters but was observed at different levels in multiple phenotypes. Phenotypes without hormone receptor and HER2 receptor expression (characteristics of triple-negative breast cancer (TNBC)) included metaclusters with high levels of Ki-67, p53, EGFR and the hypoxia marker CAIX (metaclusters 15–17); basal cytokeratins (metacluster 18); and even luminal cytokeratins (PhenoGraph clusters within metaclusters 19 and 22) (Fig. 1c).

Multicellular breast cancer architecture

On the basis of these single-cell phenotypes, we sought to define patterns of multicellular architecture in breast tumour tissue (Fig. 2a).

We evaluated regional correlations between cellular metaclusters to determine whether cells co-occurred across all images, and used neighbourhood analysis based on permutation tests[17] to quantify cell colocalization and identify statistically significant interaction or avoidance between pairs of cell phenotypes (Fig. 2b). Tumour cell phenotypes were almost never correlated: each individual tumour contained many homotypic interactions between similar cells and few heterotypic interactions between tumour cells (Fig. 2b, highlighted area 1). Generally, heterotypic interactions were associated with regionally specific structures such as blood vessels (Fig. 2b, highlighted area 2) or with distinct epithelial areas (highlighted area 1) or stromal areas with a high density of cells, in which immune cells interacted (highlighted area 3). Directional interactions were also observed where supporting fibroblasts enclosed endothelial cells in large blood vessels and where stromal cells surrounded tumour cells at the tumour–stroma interface (Fig. 2b, highlighted areas 2 and 4). T cells and proliferating epithelial cells were observed in the vicinity of endothelial cells, and their proportions were correlated across images (highlighted areas 5 and 6, respectively, in Fig. 2b); proliferating cells surrounded endothelial cells, not the reverse, representing their location outside vessels. The neighbourhood analysis revealed interaction signatures that distinguished well-separated or stromal-interactive tumour architectures, which were related to the tumour grade as scored by a pathologist[17] (Extended Data Fig. 7).

Tissue function is implemented by multicellular units that we refer to as communities, which consist of higher-order (rather than paired)
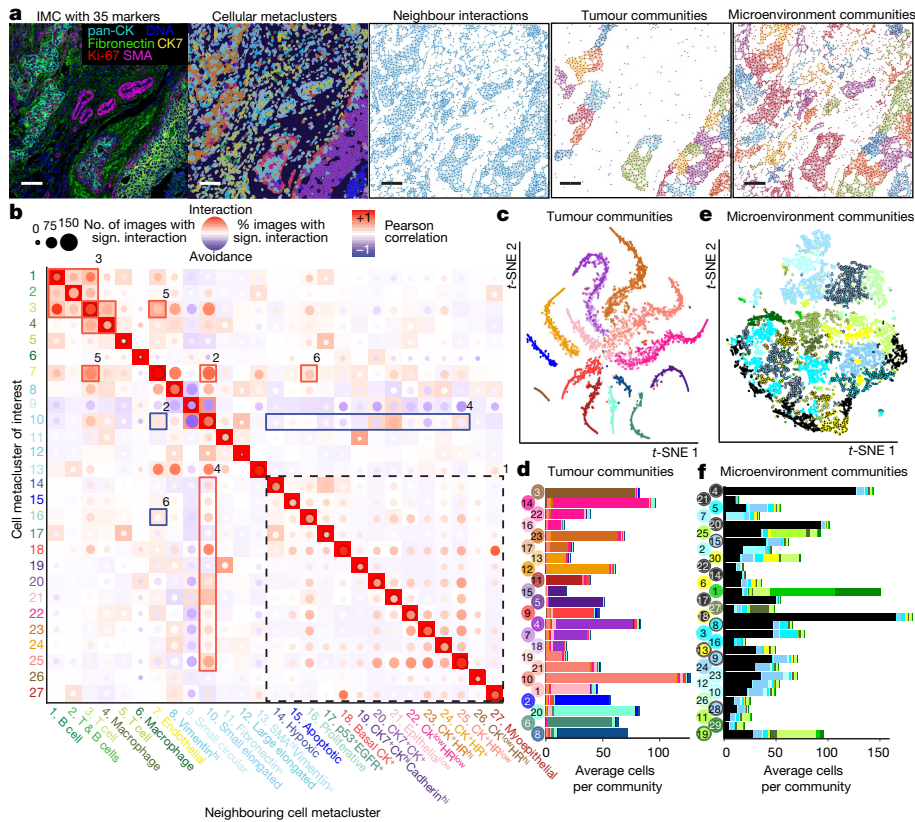
Fig. 2 | A global map of the cellular neighbourhoods and interaction networks of breast cancer. a, Representative images depicting the different steps in the spatial analysis. Images show (left to right): pseudocoloured IMC; single-cell mask of the same field of view labelled by cellular metacluster identifier; the topological cell interaction network; modular regions of the tumour network identified as epithelial communities; and modular regions of the tumour–stroma network identified as microenvironment communities. Scale bars, 100 μm. **b**, Heat map in which squares indicate the Pearson correlation of cell phenotype proportions across all measured tissue regions ($n = 367$ images) and circles indicate significant pairwise cell-type interaction or avoidance summarized across the two-sided permutation tests on the individual images ($n = 367$ images, 1,000 permutations each). Circle colour

indicates the percentage of images and size represents the number of images with a significant cell–cell interaction or avoidance ($P < 0.01$). Highlighted interactions (numbered boxes) include (1) tumour epithelium; (2) endothelium; (3) immune cells; (4) surrounding stroma; (5) endothelium and T cells; and (6) proliferating epithelium surrounding endothelial cells. Highlight colours that are not symmetrical indicate examples of directional interaction. **c–f**, Individually coloured epithelial communities ($n = 8,495$) (**c**, **d**) and microenvironment communities ($n = 12,854$) (**e**, **f**) clustered by PhenoGraph on the basis of the minimum to maximum normalized absolute number of cells from each cellular metacluster, and visualized on a t-SNE map (**c**, **e**) and in stacked bar plots (**d**, **f**) that indicate the average number of cells from each cellular metacluster; black represents tumour cell phenotypes.

interactions between one or more cell phenotypes. We identified communities by first constructing a topological network of the interactions between neighbouring cells and then applying a graph-based approach to detect communities using the Louvain algorithm[24] (Fig. 2a). Applied only to tumour cells, community detection identified dense epithelial patches of different sizes, which we term tumour communities (TCs). When applied to all cells, microenvironment communities (MCs) that contain tumour and stromal cell components were identified (Fig. 2a). Using PhenoGraph, we grouped multicell communities according to only their tumour cell phenotypes (TCs), or all cells agnostic to tumour cell phenotype (MCs). Tumour communities were mostly dominated by a single cellular metacluster and were separated on the basis of the absolute number of cells (Fig. 2c, d (for example, TCs 4, 7 and 18)) (Supplementary Figs. 1, 2). Some microenvironment communities consisted of fibroblasts that interacted with a variety of tumour cells (MCs 2, 5 and 8); others showed sparse stroma content (MCs 14, 17, 18 and 20–22) or were enriched for T cells (MCs 19, 25 and 30), macrophages (MC 27), large networks of T and B cells (MC 1) or endothelial cells (MCs 6, 7, 13, 25 and 30) (Fig. 2e, f). Communities that were enriched in fibroblasts had few interacting immune cells, in accordance with the known roles of fibroblasts as agents of tumour desmoplasia and immune exclusion[25].

Single-cell pathology and risk

We next investigated how the organization of single cells into communities contributes to the tissue architecture of breast cancer and its subtypes[6]. Cells from multiple cellular metaclusters were found in every clinically defined subtype of breast cancer (Extended Data Fig. 6a), which indicates that the general classification based on pathology does not fully elucidate inter- and intrapatient cellular heterogeneity[1,9,11]. We reasoned that examining the landscape of pathology at a single-cell level would provide a higher-resolution classification of patients than that achieved using current clinical subtypes that are based on classic histology of single stains. Using unsupervised clustering, we grouped patient tumours on the basis of their tumour cell metacluster composition and identified 18 single-cell pathology (SCP) subgroups that split the classic clinical subtypes (Fig. 3a, Extended Data Fig. 8a, Supplementary Table 3). SCP subgroups had various proportions of the epithelial communities (Fig. 3b), and individual SCP subgroups had distinct clinical outcomes when compared to all other patients, to SCP subgroups of the same clinical classification and to other SCP subgroups that contain similar cellular metaclusters but different architectures (Fig. 4a–h, Supplementary Tables 4, 5).

Tumours that are clinically defined by their expression of hormone receptors (HR+ tumours) were divided into those that were strongly
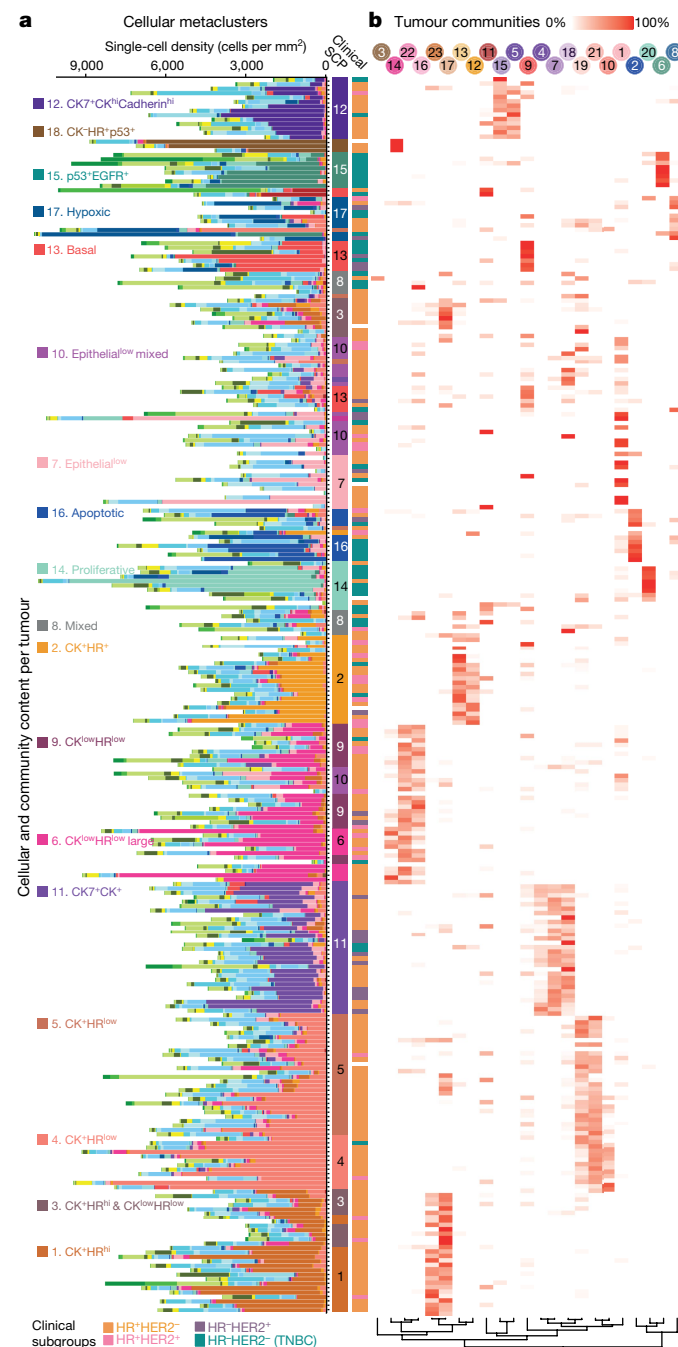
Fig. 3 | Single-cell pathology identifies subgroups of patients with breast cancer. a, Hierarchically clustered stacked bar plot of cell-type metacluster densities in each tumour. Coloured columns (right) indicate clinical subtype and SCP subgroup. **b**, Heat map indicating the proportions of different epithelial communities that are present within each image.

CK[+]HR[+] cells), was significantly enriched in the HR[+]HER2[+] clinical subtype, which was otherwise dominated by the CK[low]HR[low] metacluster 22 (Fig. 3a, Extended Data Figs. 6, 8a). SCPs 11 and 12 were characterized by CK7[+] cells primarily from metaclusters 20 and 19, respectively. SCP 11 overlapped with the clinically assigned HR[−]HER2[+] tumour type, and—although this clinical subtype usually has poor outcomes[6,8]—patients with SCP 11 tumours had significantly better outcomes than other patients in this cohort (Fig. 4e). By contrast, the small number of patients with CK7[+] tumours of the SCP 12 subgroup, which were predominantly clinically assigned as HR[+]HER2[−], did not survive in the long term (Figs. 3a, 4a, c, e, Extended Data Figs. 6, 8a). Tumours from patients with high-risk TNBC contained distinct cell types including cells with a pattern of cytokeratin expression suggestive of a luminal rather than a myoepithelial cell of origin (Figs. 1c, 3a, Extended Data Figs. 6, 8a). TNBC phenotypes without luminal epithelial markers and with high levels of hypoxic, p53[+]EGFR[+], basal or proliferative markers distinguished SCPs 13, 14, 15 and 17, which were all associated with poor outcome (Figs. 1c, 3a, 4f, Extended Data Figs. 6, 8a). SCP 16 tumours were p53[+] and expressed apoptotic markers, and, notably, patients with tumours of this subgroup did not succumb to disease even though they were clinically classified as having TNBC (Fig. 4f, Extended Data 8a).

By mapping the cellular spatial organization of these tumours, we observed variable structures and cellular densities, and relationships between cellular phenotype and tissue organization (Fig. 3a, b). Heterogeneous tumours consist of multiple phenotypically pure communities, as indicated by many bands on the heat map, whereas homogeneous tumours that are organized in one epithelial sheet or with similar communities of different sizes have only a few clustered bands (Fig. 3b). Most tumours were dominated by a single tumour cell metacluster and few community types, but tumours in SCP 8 and some in SCP 10 were unusually heterogeneous—consisting of multiple epithelial cellular metaclusters at similar proportions that were localized to spatially distinct communities (Fig. 3b, Extended Data Fig. 6). Patients with these heterogeneous tumours of subgroup SCP 8 had very poor outcomes (Fig. 4d). Overall, intratumour phenotypic heterogeneity was spatially segregated into separate tumour communities as opposed to heterogeneous tumour masses, and patients with tumours with spatiophenotypic heterogeneity had poorer outcomes.

Unlike tumour cell phenotypes, the stromal cell phenotypes that we identified were present in every clinical subtype at similar densities (Figs. 1c, 3a, Extended Data Fig. 6). We therefore investigated whether the tumour–stromal microenvironment communities were more informative than stromal phenotype content alone. When we hierarchically clustered images according to the presence of multiple microenvironment communities, 11 groups (which we term stromal environments (SEs)), were revealed; some were enriched in one microenvironment community (single column in Extended Data Fig. 7a), whereas others contained mixtures of communities (multiple columns in Extended Data Fig. 7a). Neighbour analysis detected distinct cell–cell interactions within each stromal environment (Extended Data Fig. 7b). Some stromal environments included large epithelial networks with sparse stroma (SEs 7 and 10, which are made of MCs 17 or 18); others involved vascularized regions (SE9, which involves MCs 6 and/or 13), showed different fibroblast phenotypes such as vimentin[hi] (SEs 4, 6 and 9, which involve metacluster 8 in many communities) or fibronectin[hi] (SEs 1 and 2, which include metacluster 11); and others were poorly cohesive and made up of many small communities (SEs 2, 5 and 6).

We found that stromal environments were associated with SCP subgroups and specific tumour cell phenotypes. For example, hypoxic SCP 17 TNBCs were commonly classified as large, stroma-deficient epithelial sheets (SE 7), and SCP 13–16 TNBCs were associated with T-cell-enriched (SEs 5 and 8) or macrophage-enriched (SE 3) stromal environments (Extended Data Figs. 7, 8b). HR[+] tumours are more likely to be immune-cold, but some contained rare and localized immune-enriched
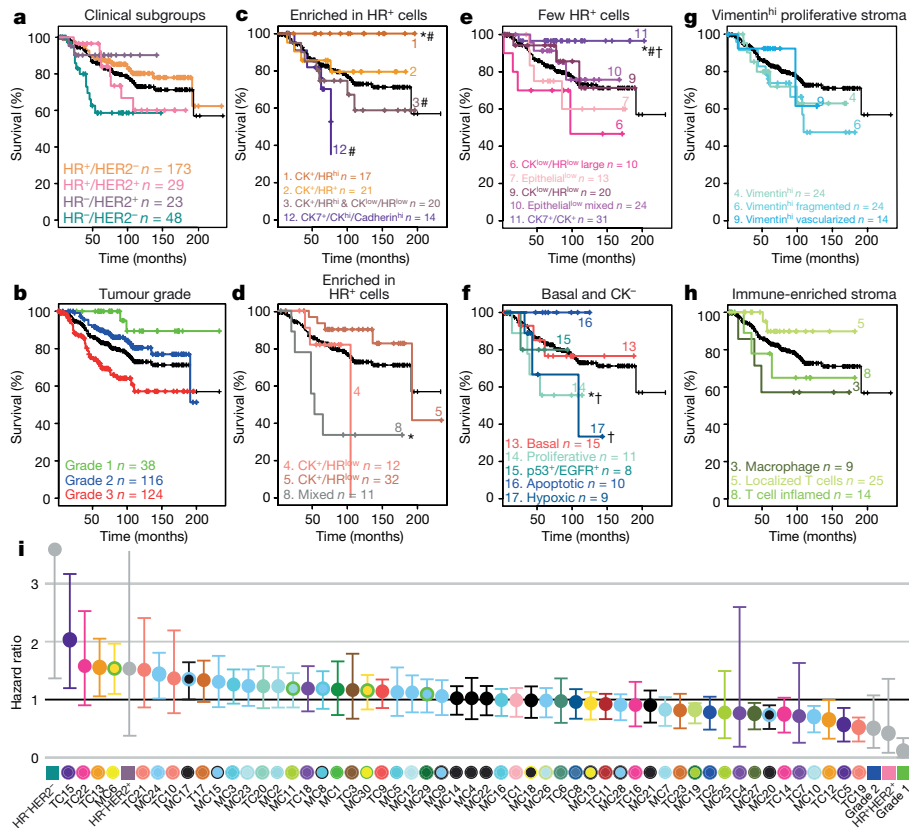
enriched in HR[hi/+] cells (SCPs 1–5 and 12), and those with few HR[hi/+] cells surrounded by many cells that expressed only low levels or lacked HRs (SCPs 6–10 and 11), which are currently not clinically classified (Fig. 3a, Extended Data Figs. 6, 8a, Supplementary Fig. 3). SCP 1, which predominantly contains metacluster 23 (CK[+]HR[hi] tumour cells), was only associated with patients who did not succumb to disease. Conversely, SCP 3—which contains the same cellular metaclusters but differs in structure, with smaller communities and relatively higher proportions of CK[low]HR[low] cells of metaclusters 22 and 25—was associated with poor prognosis, as were SCPs 6 and 9, which involve predominantly CK[low]HR[low] cells (Figs. 3a, 4c, e, Extended Data Fig. 9). SCP 2 (containing

Fig. 4 | Single-cell pathology subgroups have distinct clinical outcomes.
a–**h**, Kaplan–Meier curves of overall survival for each patient group (n = 278 patients in total) on the basis of clinical subtype (**a**), clinical grade (**b**), SCP subgroup (**c**–**f**) or stromal environment (**g**, **h**). *P < 0.05 compared to all other samples, #P < 0.05 compared to similar subgroups, †P < 0.05 compared to other HR+HER2− patients (log-rank test). For exact P values, see Supplementary Tables 4 and 6. **i**, Relative hazard ratios and 95% confidence intervals of disease-specific overall survival for densities of tumour communities and microenvironment communities and clinical categories (molecular subtype and grade) estimated using a Cox proportional hazards model (n = 266 patients, reference group HR+HER2− grade 3 for molecular subtypes and grades, n = 15 were excluded as only communities with fewer than 10 cells were identified in these patients)).

communities (SE 5)[1]. The HR+ tumours never contained an exclusively immune-hot stromal environment (SE 8) like those observed in TNBCs and the highly heterogeneous SCP 8 (Extended Data Figs. 7, 8b). HR+ tumours were associated with a range of fibroblast-enriched stromal environments, with small elongated fibroblasts interacting with CK+HR^low epithelial cells in SEs 1, 2 and 5 and vimentin^hi stromal cells associating with HR^hi tumour cells in SEs 4 and 6 (Fig. 2b, c, Extended Data Fig. 8b). Stromal environments were related to tumour phenotype, but only fragmented SE 6 (containing proliferative vimentin^hi fibroblasts) was independently associated with poorer disease-free and overall survival (Fig. 4h, Extended Data Figs. 8b, 9, Supplementary Tables 6, 7).

Compared to clinically defined subtypes, single-cell pathology grouping improved the ability to predict the overall survival of a patient using Cox proportional hazards modelling (Supplementary Table 8). To identify features associated with patient risk that are not captured by clinical grading and classification, we investigated the epithelial and stromal single-cell and community contributions to the model. Almost no single-cell phenotypes or cellular metaclusters were independently associated with outcome (not shown); however, spatially defined cell communities were (Fig. 4i). For some cell types, large communities of tumour cells were related to a better outcome, whereas similar networks of a small size were related to a poor outcome (Fig. 4i (TCs 12 versus 13; 17 versus 23; and 5 versus 15). In addition, the microenvironment community MC 6—which is characterized by vascularization with T cell involvement—was significantly associated with an increased risk of death, even though it was more commonly found in the low-risk HR+ clinical subgroups than other subgroups (Fig. 4i,

Extended Data Fig. 7). By contrast, highly T-cell-infiltrated MC 19 and macrophage-enriched MC 27 were significantly associated with better patient outcomes, even though inflammation is more common in high-risk TNBC tumours than other clinical subgroups (Fig. 4i, Extended Data Fig. 7). Thus, tumour types and tumour and stromal architectures that are defined by single-cell pathology could inform prognosis beyond the current clinical classifications.

Quantification of intratumour heterogeneity

We investigated the reproducibility and spatial variability of SCP classifications in two central and two peripheral tumour regions from 72 patients in an independent cohort, which resulted in 344 additional images containing a total of 411,410 cells (Supplementary Table 9). We used the same analytical approach described above to independently define single-cell phenotypes, match them to cellular metaclusters and classify each imaged region into SCP subgroups and stromal architectures (Extended Data Figs. 5c, 10a). All cellular metaclusters and SCP subgroups that were identified in the first cohort were present in the second cohort. Probably owing to the patient-selection strategy (which enriched the second cohort in metastatic low-grade tumours), we observed a higher proportion of subgroups that involve CK^low HR^low single-cell phenotypes—such as SCP 3, 6 and 9—in this cohort (Extended Data Fig. 10a). For each region of each tumour, we quantified the spatial variability of cell phenotypes using Shannon entropy (Extended Data Fig. 10a colour bar), and determined the difference in cellular content relative to the patient's overall distribution of single-cell phenotypes

(all regions) using Kullback–Leibler divergence (Extended Data Fig. 10b, c). Most patients had only moderate inter-region heterogeneity (Extended Data Fig. 10b). Distinctly divergent regions were more commonly identified as SCP 2, which consists of CK$^+$HR$^+$ cells, or SCP 7, containing epithelial^{low} cells (Extended Data Fig. 10b, c). Approximately 40% of tumours had identical classification in all regions, and 60% had one or more regions that did not agree with the whole tumour classification (Extended Data Fig. 10d). However, in the majority of cases, SCP classification of individual regions matched the tumour-wide classification.

SCP subgroups had varying levels of regional heterogeneity. Most SCP subgroups that were TNBCs were phenotypically homogeneous (the exception being basal CK$^+$ SCP 13). SCP subgroups that were associated with HR$^+$ cancer with relatively poor prognosis—such as CK^{low}HR^{low} SCPs 6, 7, 9 and 10—had substantial spatial heterogeneity (Extended Data Fig. 10e). Different SCP subgroups containing similar cell types occasionally co-occurred in the same tumour, but SCP 9 was always accompanied by SCP 6 regions, and the two were associated with similar outcomes (Fig. 4e, Extended Data Fig. 10e). Spatially heterogeneous HR^{low} and CK^{low} SCP subgroups 6–10 have outcomes that are not statistically different from the population average, and on the basis of their regional heterogeneity it is probable that unsampled phenotypes influenced the accuracy of classification (Extended Data Fig. 10e). We observed that heterogeneous regions classified as SCP 8 or 10 were always accompanied by multiple additional tumour subtypes (Extended Data Fig. 10e). These may be multiclonal or highly plastic tumours. Subsampling was sufficient to identify and stratify homogeneous low-risk HR^{hi} and high-risk TNBCs, but increased sampling may be needed to accurately identify HR^{low} tumour phenotypes and tumours with considerable intratumour heterogeneity.

Discussion

This systematic, multidimensional interrogation of breast cancer histology has generated a detailed spatial map of single-cell phenotypes and cellular communities and their relationships with disease. We demonstrated that single-cell pathology can better segregate patients with distinct clinical outcomes than can the current strategy of clinical subtyping. Analysis of multicellular structures revealed that phenotypic heterogeneity in tumours was spatially localized to distinct regions or lesions. Moreover, in relation to patient outcome, the information yielded by the multicellular structures was superior to that yielded by single-cell data alone. We observed that phenotypic and spatial heterogeneity varied between clinically established subtypes and identified breast cancer phenotypes that co-occur. This work suggests that multicellular spatial information is medically relevant and provides a basis for future study of how spatial and phenotypic tissue features influence disease outcome.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1876-x.

1. Wagner, J. et al. A single-cell atlas of the tumor and immune ecosystem of human breast cancer. *Cell* **177**, 1330–1345 (2019).
2. Bedard, P. L., Hansen, A. R., Ratain, M. J. & Siu, L. L. Tumour heterogeneity in the clinic. *Nature* **501**, 355–364 (2013).
3. Dagogo-Jack, I. & Shaw, A. T. Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **15**, 81–94 (2018).
4. McGranahan, N. & Swanton, C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer Cell* **27**, 15–26 (2015).
5. Giesen, C. et al. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11**, 417–422 (2014).
6. Coates, A. S. et al. Tailoring therapies—improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann. Oncol.* **26**, 1533–1546 (2015).
7. Hammond, M. E. H. et al. American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J. Clin. Oncol.* **28**, 2784–2795 (2010).
8. Wolff, A. C. et al. Human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline focused update. *J. Clin. Oncol.* **36**, 2105–2122 (2018).
9. Curtis, C. et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).
10. Sørlie, T. et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl Acad. Sci. USA* **98**, 10869–10874 (2001).
11. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
12. Beck, A. H. et al. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Sci. Transl. Med.* **3**, 108ra113 (2011).
13. Focke, C. M., Decker, T. & van Diest, P. J. Intratumoral heterogeneity of Ki67 expression in early breast cancers exceeds variability between individual tumours. *Histopathology* **69**, 849–861 (2016).
14. Rye, I. H. et al. Intratumor heterogeneity defines treatment-resistant HER2+ breast tumors. *Mol. Oncol.* **12**, 1838–1855 (2018).
15. Angelo, M. et al. Multiplexed ion beam imaging of human breast tumors. *Nat. Med.* **20**, 436–442 (2014).
16. Gerdes, M. J. et al. Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. *Proc. Natl Acad. Sci. USA* **110**, 11982–11987 (2013).
17. Schapiro, D. et al. histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nat. Methods* **14**, 873–876 (2017).
18. Keren, L. et al. A structured tumor-immune microenvironment in triple negative breast cancer revealed by multiplexed ion beam imaging. *Cell* **174**, 1373–1387 (2018).
19. Carvajal-Hausdorf, D. E. et al. Multiplexed (18-plex) measurement of signaling targets and cytotoxic T cells in trastuzumab-treated patients using imaging mass cytometry. *Clin. Cancer Res.* **25**, 3054–3062 (2019).
20. Damond, N. et al. A map of human type 1 diabetes progression by imaging mass cytometry. *Cell Metab.* **29**, 755–768 (2019).
21. Bodenmiller, B. Multiplexed epitope-based tissue imaging for discovery and healthcare applications. *Cell Syst.* **2**, 225–238 (2016).
22. Lin, J.-R. et al. Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. *eLife* **7**, e31657 (2018).
23. Levine, J. H. et al. Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell* **162**, 184–197 (2015).
24. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, P10008 (2008).
25. Kalluri, R. The biology and function of fibroblasts in cancer. *Nat. Rev. Cancer* **16**, 582–598 (2016).

Methods

Data reporting

No statistical methods were used to predetermine sample size. All samples were stained simultaneously. Image-acquisition order was distributed spatially and independently of patient or tumour replicates. Image acquisition and single-cell quantification and clustering were blinded to patient identifiers and clinical metadata.

Clinical data

The samples of tumours from the described patient cohorts were obtained from University Hospital Basel and University Hospital Zurich. The cohort from University Hospital Basel includes 281 patients who were not selected for any clinical or histological features. Pathologists recorded the available patient metadata (Supplementary Table 1) and evaluated the suitability of tissue sections for tissue microarray (TMA) construction[26]. The TMA contains one 0.8-mm tumour core per patient, in some cases an additional matched sample of healthy breast tissue, and a few control samples (liver tissue). The cohort from University Hospital Zurich comprised 72 patients; the samples include four 0.6-mm cores from four different regions of each tumour as previously described[27]. Tumour cores were punched from two central and two peripheral areas that averaged 1 cm in distance between regions. Samples were selected to contain equal proportions of the different tumour grades and to include patients with and without lymph node metastases (Supplementary Table 9). In total 720 images were acquired that varied in size and localization in the tumour. This project was approved by the local Commission of Ethics (ref. numbers 2014-397 and 2012-0553).

Panel

An antibody panel was designed to target epitopes specific for breast cancer as well as markers for cell cycle and phosphorylation-based signalling, and to distinguish epithelial, endothelial, mesenchymal and immune cell types (Extended Data Fig. 1, Supplementary Table 2). Clone information is available in Supplementary Table 2.

Preparation and staining

Tissue samples were formalin-fixed and paraffin-embedded at the University Hospitals of Basel and Zurich. The antibody panel described in Extended Data Fig. 1 was used to stain the tissue sections[5]. Tissue sections were dewaxed in xylene overnight and rehydrated in a graded series of alcohol (ethanol:deionized water 100:0, 90:10, 80:20, 70:30, 50:50, 0:100; 5 min each). In a 95-°C water bath, heat-induced epitope retrieval was conducted in Tris-EDTA buffer at pH 9 for 20 min. The TMAs were immediately cooled and then blocked with 3% BSA and 5% goat serum in TBS for 1 h. Samples were incubated overnight at 4 °C in primary antibody at 7.5 g l^{-1} diluted in TBS, 0.1% Triton X-100 and 1% BSA. Tissue samples were washed twice with TBS and 0.1% Triton X-100, and twice with TBS, and dried before IMC measurements.

For combined immunofluorescence and IMC staining, tissues were stained overnight at 4 °C with primary metal-conjugated mouse HER2 (^{151}Eu) and rabbit pan-cytokeratin (^{175}Lu) antibodies before washing and the mixed addition of fluorescent and metal-conjugated anti-mouse (AF488, ^{165}Ho) and anti-rabbit (AF555, ^{159}Tb) secondary stains for 1 h at room temperature. A cover slip was added, and tissue was imaged for fluorescence signal. Subsequently, the cover slip was removed, and samples were washed, dried and subjected to mass cytometry laser ablation and acquisition.

IMC

Images were acquired using a Hyperion Imaging System (Fluidigm). The largest square area from each core of a TMA was laser-ablated in a rastered pattern at 200 Hz, and preprocessing of the raw data was completed using commercial acquisition software (Fluidigm). IMC acquisition stability was monitored by interspersed acquisition of isotope-containing polymer (Fludigm). All successful image acquisitions were processed, and images containing pan-marker staining variation specific to TMA location were removed. In a few cases the acquisition was interrupted and later continued, resulting in two tumour images from the same TMA core. Therefore, the cohort of 281 patients resulted in 289 tumour, 87 healthy breast and 5 liver control images. Where applicable, signal spillover between channels was corrected using functions from the CATALYST R package[28] (v.1.5.6). The cohort of 72 patients resulted in 263 tumour, 68 healthy breast and 6 control images that were used for analysis.

Data processing

Data were converted to TIFF format and segmented into single cells using the flexible analysis pipeline available at https://github.com/BodenmillerGroup/ImcSegmentationPipeline. In brief, individual cells and regions of tumour and stroma were segmented using a combination of Ilastik v.1.1.9[29] and CellProfiler v.2.1.1[30]. Ilastik was used to generate a probability map by classifying pixels (single cells: nuclei, membrane and background; tumour/stroma: tumour, stroma and background) on the basis of a combination of antibody stains to identify membranes and nuclei. Probability maps were then segmented into single-cell or tumour and stroma object masks using CellProfiler.

Single-cell segmentation masks and TIFF images of the 35 channels were overlaid and the mean expression levels of markers and spatial features of single cells were extracted using the MATLAB toolbox regionprops. Even with very-good-quality segmentation, the imaging of tissue segments results in single-cell data of tissue slices and overlapping cell fragments that do not always capture the nucleus of a cell, and therefore nuclei-mismatched signal can be assigned to neighbouring cells in densely packed areas. This can lead to rare cases in which data assigned to one cell contains marker expression from the neighbourhood.

The single-cell IDs of the direct neighbours of each cell (that is, cells within 4 pixels (4 μm) of the cell of interest) were detected and recorded using histoCAT software. The number of pixels expanded to detect neighbours was chosen such that small gaps in segmentation would be bridged, but no cells after the direct neighbour would be recorded (cell minor axis lengths: 5th–95th percentile 4.84–14.59 pixels; average 9.51 pixels).

Individual cell locations inside or outside of a tumour mask were identified and the distance of each cell to the tumour boundary (from inside and outside of the tumour region) was calculated using the MATLAB toolbox regionprops. Distances were measured between the closest pixels of the objects in question.

Data transformation and normalization

The presented data were not transformed, and all analyses were based on raw IMC measurements. Single-cell marker expressions are summarized by mean pixel values for each channel. The single-cell data were censored at the 99th percentile to remove outliers, and z-scored cluster means were visualized in heat maps. For t-SNE and PhenoGraph the data were normalized to the 99th percentile, as is suggested for these algorithms[23,31]. To visualize the number of cells per image or patient and for survival modelling, the cell counts were normalized by the image area (total number of pixels) and displayed as cell density. For Cox proportional hazards survival modelling, these densities were multiplied by a factor of 10^7 to yield values larger than 1 and then log-transformed.

Analysis workflow

The single-cell analysis pipeline was implemented in R, but image-analysis steps were performed in MATLAB. All statistical tests were performed using common functions in R.

Article

Clustering and metaclustering

Single cells of the large cohort from University Hospital Basel were clustered into groups of phenotypically similar cells using a combination of PhenoGraph[23] for initial, unsupervised clustering and an aggregation of these clusters into larger groups on the basis of their mean marker correlations to identify cellular metaclusters. In a first step, the data were over-clustered to detect and separate rare cell subpopulations. PhenoGraph (v.2.0) was used with default parameters (as implemented in histoCAT and Cyt) and 20 nearest neighbours. For high-dimensional clustering, 29 markers and 4 cell-shape features were used: iridium, histone, phosphorylated histone, CK14, CK5, CK8/18, CK19, CK7, pan-CK, E/P-cadherin, ER, PR, HER2, GATA3, SMA, vimentin, fibronectin, vWF/CD31, CD44, CD45, CD68, CD3, CD20, cleaved caspase 3/cleaved PARP, carbonic anhydrase, phosphorylated S6, Ki67, p53, EGFR, area, eccentricity, extent and number of neighbours. Of the resulting 71 clusters, the 59 epithelial clusters were aggregated into larger groups following the hierarchical clustering (Euclidean distance and Ward's linkage) of their mean marker correlations. Multiscale bootstrap resampling was used to assess the uncertainty of each subtree (R package pvclust, v.2.0), and separation of the hierarchy was assigned so that significant epithelial subtrees were maintained and known biological differences were separated. This resulted in 14 tumour cell metaclusters of varying size and subtree robustness (Extended Data Fig. 5a). Clusters that showed a marker expression typical of stromal and immune cells (these clusters were limited in number owing to our tumour-marker-focused panel) were kept as in the original PhenoGraph clustering and not aggregated into larger groups. This metaclustering yielded 27 cellular subgroups, which represent various immune, stromal and epithelial cell types. The granularity, the level and the detail at which phenotypes are divided or clustered for the studied cell types depends on the selection of both the panel and the choice of parameters. Although a more-granular distinction of cell types might enable even more subtle differences in the expression of markers to be elucidated, it would limit comparability between tumours as many tumour cell types would be patient-specific.

Cluster matching across cohorts

Single cells from the second cohort from University Hospital Zurich were clustered unsupervised and independently using PhenoGraph[23] with the same settings described for the first cohort above and a nearest neighbour parameter of 30. The clusters were matched to the most-similar metacluster of the previous cohort using Pearson correlation of the z-scored mean marker expression. In two specific cases (clusters 8 and 15) in which the cluster in question was poorly correlated with all metaclusters but was most correlated with a stromal cell type, we manually re-assigned the cluster because on visual inspection of the images, those clusters represented cells that form clear tumour bulks.

Barnes–Hut t-SNE

For visualization, high-dimensional single-cell data were reduced to two dimensions using the nonlinear dimensionality reduction algorithm t-SNE[31]. We applied the Barnes–Hut implementation of t-SNE to 99th-percentile normalized data with default parameters (initial dimensions, 110; perplexity, 30; θ, 0.5). The algorithm was run on a randomly subsampled set of cells (20% from each image) so that visible patterns in crowded plots were not obscured and for better computational performance.

Neighbourhood analysis

To identify significantly enriched or depleted pairwise neighbour interactions between cell types, histoCAT functions were used to perform a permutation-test-based analysis of spatial single-cell neighbourhoods[17]. Neighbouring cells were defined as those within 4 pixels (4 μm). P values smaller than 0.01 were considered as significant.

Single-cell pathology patient grouping

Patients were grouped on the basis of the proportions of tumour cell metaclusters using the cytofkit R implementation of PhenoGraph[23] (v.1.10.0) with 8 nearest neighbours and default parameters. The number of nearest neighbours was chosen such that small groups of patients for whom a distinct cell type was predominant could be separated. A choice of a higher value for this parameter would lead to fewer groups, and hence patients with entirely unrelated predominant phenotypes would be grouped together. A lower value of the nearest neighbour parameter might capture more subtle differences in the cellular composition of tumour types, but would severely limit statistical power for group comparison and survival analysis. The composition of each patient group by their clinically assigned metadata is available in Supplementary Table 3. Patient group 18 was removed from further downstream analysis owing to lack of statistical power, as it contained only three patients with distinct tumours that were strongly dominated by a rare HR^+CK^- cell type.

Single-cell pathology group matching

Tumour cores from the second cohort from University Hospital of Zurich were assigned to the most-similar previously defined SCP group on the basis of their matched tumour cell-type components. The inverse of Pearson correlation was used as a distance metric.

Spatial heterogeneity

In the cohort from University Hospital Zurich that contains multiple cores per tumour, the intracore heterogeneity of tumour cells and that of stromal cells were separately quantified using the entropy-based Shannon index on the amounts of the different cell types within each core. Shannon entropy has been shown to serve as a measure of diversity and homogeneity in various contexts[32]. It can also be considered as a measure for the information content of a string; in our case, every cell of an image is represented by a letter according to its cell type. The most-compressible string is obtained if every cell is of the same type, and the string with most information is obtained if every cell is unique according to the following formula: $H(X) = -\sum P_i \log_2(P_i)$, in which P_i is the probability of a given symbol and $H(X)$ is the Shannon entropy. Inter-core heterogeneity within a tumour was approximated by calculating the Kullback–Leibler divergence from the cell-type distribution (proportions of each cell type) of an individual core to the average cell-type distribution across all cores from a patient. Kullback–Leibler divergence describes the loss of information that occurs when going from an original distribution to a summary distribution[33]. Hence, if all cores of a tumour are composed of identical proportions of the same cell types, the Kullback–Leibler divergences of every individual core to the patient average will be minimal. The R package entropy (v.1.2.1) was used for the calculation of both Shannon entropy and Kullback–Leibler divergence. Intratumour heterogeneity and the consistency of SCP group assignment of images of the same patient are visualized in Extended Data Fig. 10.

Spatial communities

The images were converted into topological neighbourhood graphs in which every cell is represented by a node (visualized at the centroid), and the nodes are connected by an edge if the cells directly neighbour each other (Fig. 2). Neighbouring cells were defined as those within 4 pixels (4 μm) of the outermost pixel assigned to a cell. The Louvain community detection algorithm[24] (C implementation by Lefebvre and Guillaume, v.0.2, wrapped by MATLAB as used by the implementation of PhenoGraph v.2.0 that was used by histoCAT and Cyt) was then applied to identify highly interconnected spatial subunits in the tissue graph. Although using community detection algorithms on spatially constrained networks is known to hide underlying non-spatially driven solutions, the only aim of applying the algorithm here was to extract

spatial information and identify communities on the basis of physical proximity[34]. This analysis was performed on epithelial cells only to identify tumour communities (without including stromal or immune cells in the graph) and again on all cells of a tissue to identify communities of the tumour microenvironment. A tumour-specific cohesiveness score was calculated on the basis of the average sizes of the identified tumour communities. Communities that contained fewer than ten cells were excluded from further analysis to focus on cohesive cell patches and not individual disconnected cells. Fifteen patients were excluded from the analysis based on tumour communities because the imaged regions did not contain any tumour communities that consisted of at least ten cells. To identify recurring similar spatial cell-type communities, the cytofkit PhenoGraph[23] (v.1.10.0) was run on the minimum to maximum normalized, absolute numbers of cells of each cell metacluster in each community. This analysis was conducted separately for the tumour communities (on the basis of only the epithelial cell types ($k = 80$)) and for the microenvironment communities (on the basis of all cells but only taking into account the individual stromal cell types, and aggregating all tumour cell types into one label (cell type group 100: including all tumour cells, $k = 30$)). This analysis was conducted separately for each cohort but was based on the matched metacluster cell types.

Stromal environments

On the basis of their microenvironment community compositions, images were grouped into 11 different stromal environments using hierarchical clustering (Euclidean distance and Ward's linkage). This analysis was conducted separately for each cohort but was based on the matched metacluster cell types.

Overlapping classifications and enrichments

The Fisher's exact test was used to identify SCP patient groups enriched for a specific stromal environment (Extended Data Fig. 8). The test was performed using the R function fisher.test (with parameter enrichment = "greater") for every potential stromal region of a patient group. The P values were corrected for multiple testing using the Bonferroni method. This enrichment analysis was also conducted with different combinations of SCP subgroups, stromal environments and clinical classifications (Extended Data Fig. 8).

Survival curves and Cox proportional hazard regression models

Kaplan–Meier survival curves and Cox proportional hazards survival regression models were generated using the R package survival (v.2.42-4). The overall survival as well as the disease-free survival of patients in different clinical or single-cell-defined subgroups was analysed (Fig. 4, Extended Data Fig. 9, Supplementary Tables 4–8). Both log-rank tests and Cox proportional hazards models were used to investigate whether a patient subgroup significantly deviated from the survival of the remaining patients or from the survival of other patients of similar SCP groups or the same clinical classification (Fig. 4, Supplementary Tables 4–7). Log-transformed densities of communities or single cells, together with the clinical subgrouping and grading, were provided to a Cox proportional hazards survival model to find significant associations of certain community or single-cell types with patient risk and to investigate the hazard ratios (Fig. 4). Nested Cox proportional hazards models were compared using likelihood ratio tests (anova.coxph) to assess whether additional variables improved the survival model (Supplementary Table 8).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The data supporting the findings of this study (including high-dimensional TIFF images, single-cell and tumour and stroma masks, single-cell and patient data) are available online at Zenodo (https://doi.org/10.5281/zenodo.3518284).

Code availability

All of the code used to produce the results of this study is available at https://github.com/BodenmillerGroup/SCPathology_publication.

26. Kononen, J. et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat. Med.* **4**, 844–847 (1998).
27. Kündig, P. et al. Limited utility of tissue micro-arrays in detecting intra-tumoral heterogeneity in stem cell characteristics and tumor progression markers in breast cancer. *J. Transl. Med.* **16**, 118 (2018).
28. Chevrier, S. et al. Compensation of signal spillover in suspension and imaging mass cytometry. *Cell Syst.* **6**, 612–620 (2018).
29. Sommer, C., Straehle, C., Köthe, U. & Hamprecht, F. A. Ilastic: interactive learning and segmentation toolkit. In *Proc. 2011 8th IEEE International Symposium on Biomedical Imaging: From Nano to Macro* 230–233 (IEEE, 2011).
30. Carpenter, A. E. et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006).
31. Amir, A. D. et al. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nat. Biotechnol.* **31**, 545–552 (2013).
32. Angel Martín, M., & Rey, J.-M. (2000). On the role of Shannon's entropy as a measure of heterogeneity. *Geoderma* **98**, 1–3 (2000).
33. Kullback, S. & Leibler, R. A. On information and sufficiency. *Ann. Math. Stat.* **22**, 79–86 (1951).
34. Expert, P., Evans, T. S., Blondel, V. D. & Lambiotte, R. Uncovering space-independent communities in spatial networks. *Proc. Natl Acad. Sci. USA*, **108**, 7663–7668 (2011).

Author contributions H.W.J. and B.B. conceived the study. H.W.J. performed all image quantification and IMC experiments. Z.V. performed immunohistochemical staining. J.R.F. performed data analysis. V.R.T.Z. constructed image processing and analysis tools. H.W.J. and J.R.F. performed the biological analysis and interpretation with input from the co-authors. H.R.A. provided input on clinical interpretation and survival analysis. H.M., Z.V., R.M., S.D.S., S.M. and W.P.W. provided patient samples and clinical input throughout the study. H.W.J., J.R.F. and B.B. wrote the manuscript.

Article
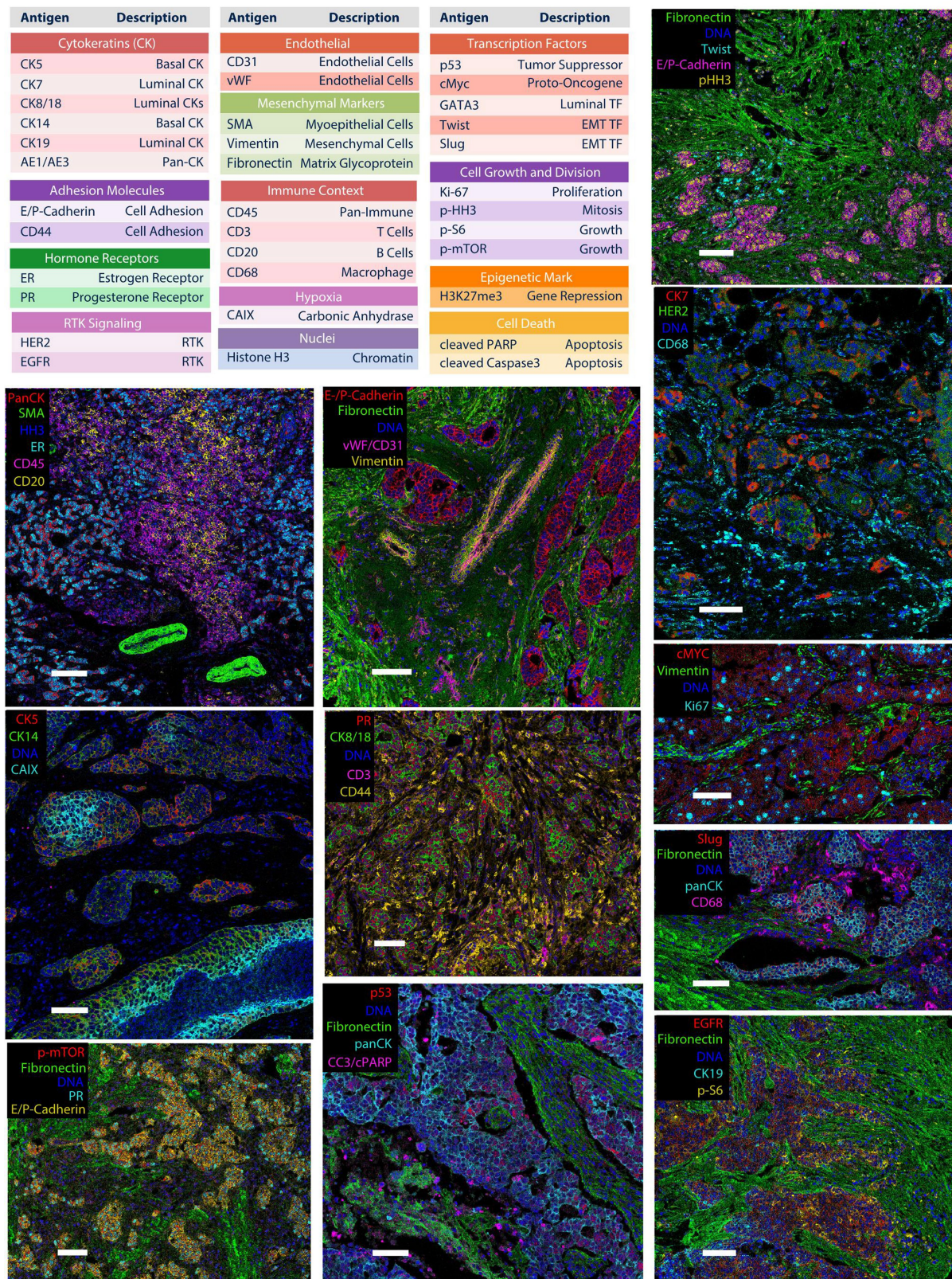
Extended Data Fig. 1 | Antibody panel and example pseudocoloured images of markers. Antigens targeted by the antibodies in the panel of 35 isotope-conjugated antibodies that was used to stain the breast cancer tissue, and representative marker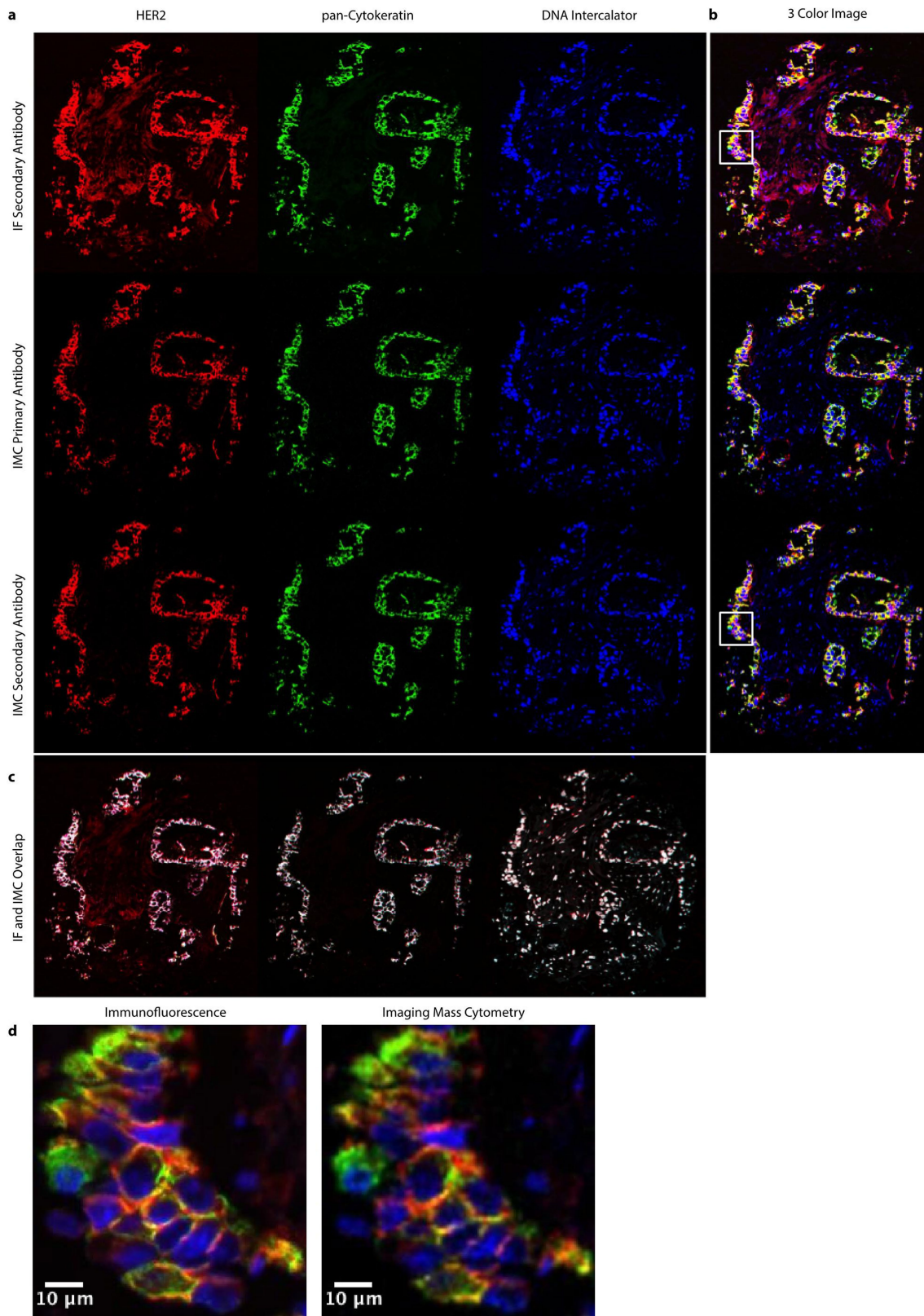 images from the analysed cohort generated by IMC. Every marker is visualized at least once. Each image represents a different tumour of the analysed cohort. Each marker was individually scaled to enable visualization. RTK, receptor tyrosine kinase; EMT, epithelial–mesenchymal transition; TF, transcription factor. Scale bars, 100 μm.
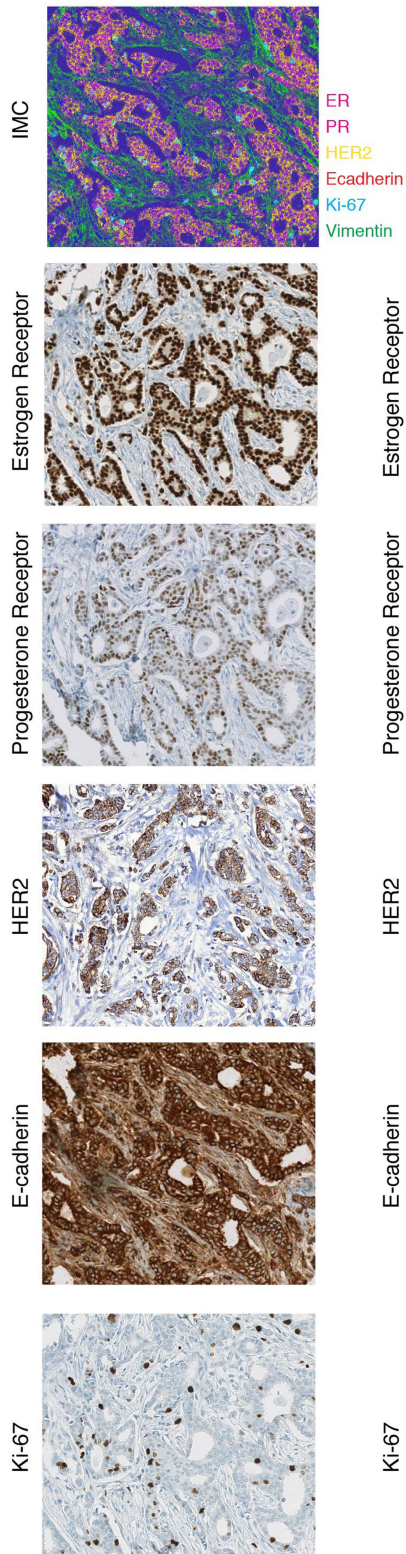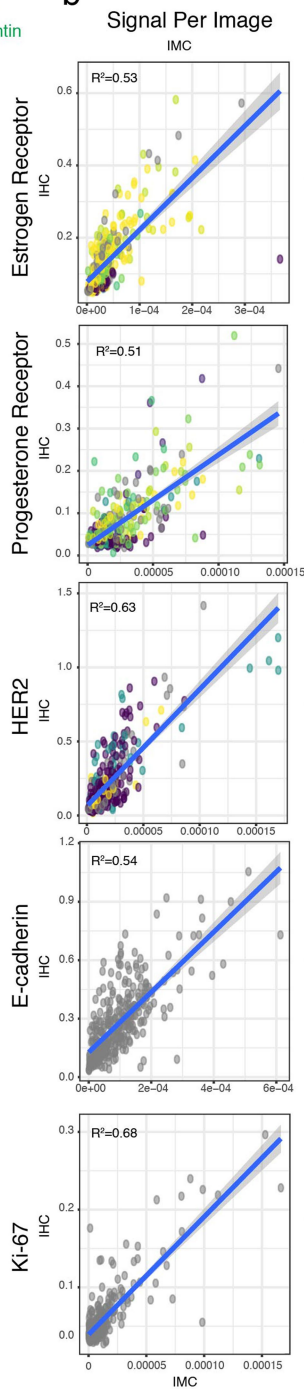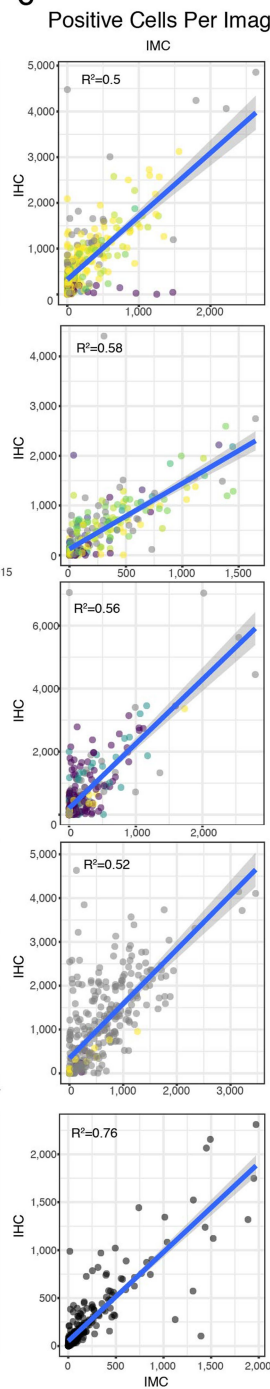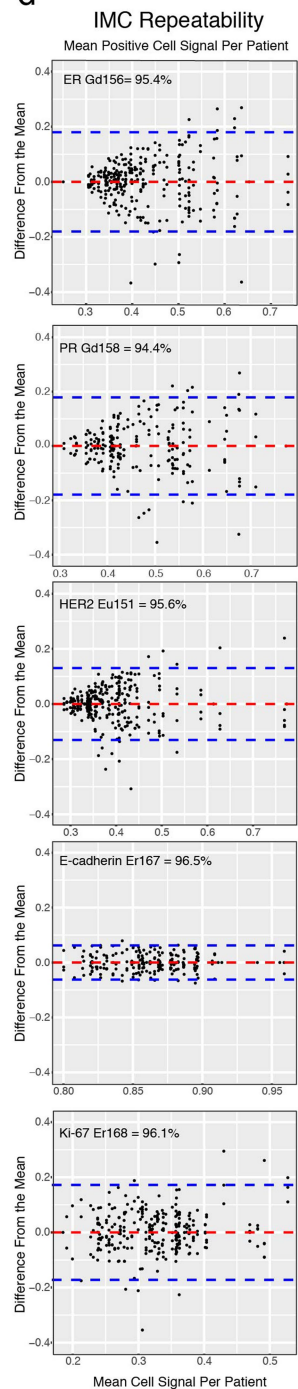
**Extended Data Fig. 2** | See next page for caption.

Article

Extended Data Fig. 2 | Comparison and reproducibility analysis of immunohistochemistry and IMC. a, Representative IMC and immunohistochemistry images of the quantified stains in sections of the same tumour core. **b**, Scatter plot and correlation of total immunohistochemistry (IHC) and IMC signal in sections of the same tumour core (IHC, optical density per μm^2, IMC, ion counts per μm^2; $n = 319$ cores). **c**, Scatter plot and correlation of the number of positively stained cells in sections from the same tumour core ($n = 319$ cores). **d**, Bland–Altman plots for reproducibility of the IMC signal in positively stained cells across images from different regions of the same tumour, adapted to visualize the average across four samples on the x axis and the difference of every individual sample to the tumour average on the y axis. Only images that contained positively stained cells and more than 200 cells in total were taken into account for this analysis (ER, $n = 280$ cores from 72 patients; PR, $n = 213$ cores from 66 patients; HER2, $n = 291$ cores from 72 patients; E/P-cadherin, $n = 200$ cores from 65 patients; Ki67, $n = 281$ cores from 72 patients). Red line represents the overall average of the differences to the tumour mean; blue lines represent the 95% confidence interval (1.96 × s.d.). The percentage of observations that fall within the confidence interval is indicated at the top of each plot.

a HER2 | pan-Cytokeratin | DNA Intercalator | **b** 3 Color Image

IF Secondary Antibody

IMC Primary Antibody

IMC Secondary Antibody

c IF and IMC Overlap

**d** Immunofluorescence | Imaging Mass Cytometry

10 µm | 10 µm

Extended Data Fig. 3 | Simultaneous immunofluorescence and mass cytometry imaging. Immunofluorescence (IF) and mass cytometry (IMC) imaging of the same tissue sample using metal-conjugated HER2 and pan-CK primary antibodies and both fluorescent and metal-conjugated secondary stains. Pseudocolour images of individual channels (**a**), three-marker images produced from each label type (**b**) (white, overlap; red, HER2; green, pan-CK; blue, DNA intercalator), as well an overlay of the same marker from all three label types (**c**) (white, overlap; red, secondary immunofluorescence; green, secondary IMC; blue, primary IMC). **d**, High-magnification images of the regions labelled with white squares in **b**, comparing the resolution, expression and similarity of immunofluorescence and IMC.
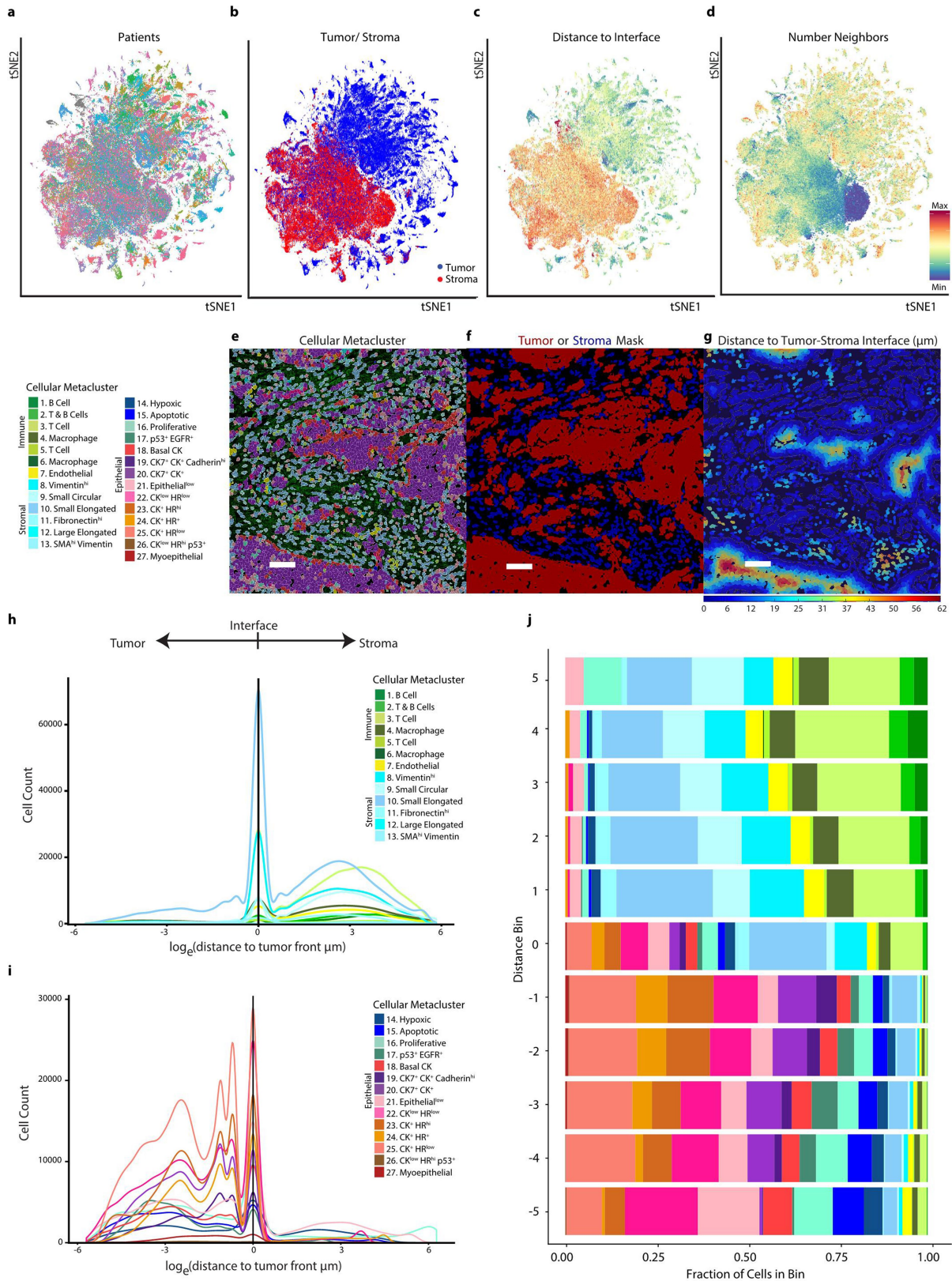
Extended Data Fig. 4 | Single-cell localization relative to the tumour–stroma interface. a–d, *t*-SNE maps of 171,288 subsampled single cells from high-dimensional images of breast tumours, coloured by patient (**a**), localization relative to the tumour–stroma interface (**b**), single-cell distance to the tumour–stroma interface (**c**) and number of neighbouring cells (**d**). **e–g**, Representative images with single-cell mask, labelled by metacluster identifier (**e**), tumour and stroma masks (**f**) and heat map that represents the distances of single cells to

the tumour–stroma interface from each side (**g**). Scale bar, 100 μm. **h**, **i**, Log-transformed distances to tumour front of stromal cell clusters (**h**) and tumour cell metaclusters (**i**). **j**, Binned distances of all metaclusters to the tumour front. Bin number 0 contains all cells that are directly touching the interface. Negative distances represent the distance to the tumour boundary from inside the tumour and positive values indicate the distance outside the tumour.
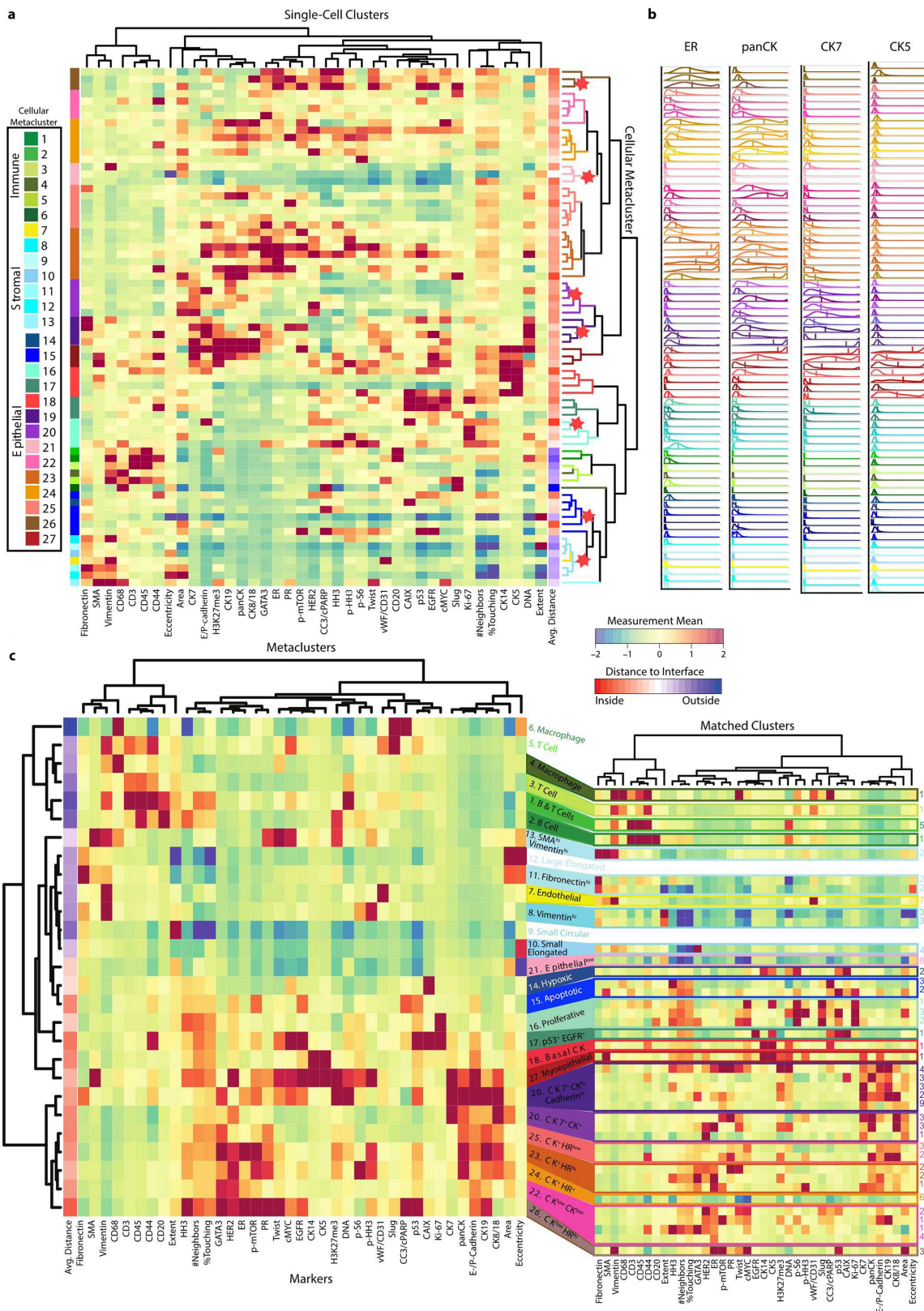
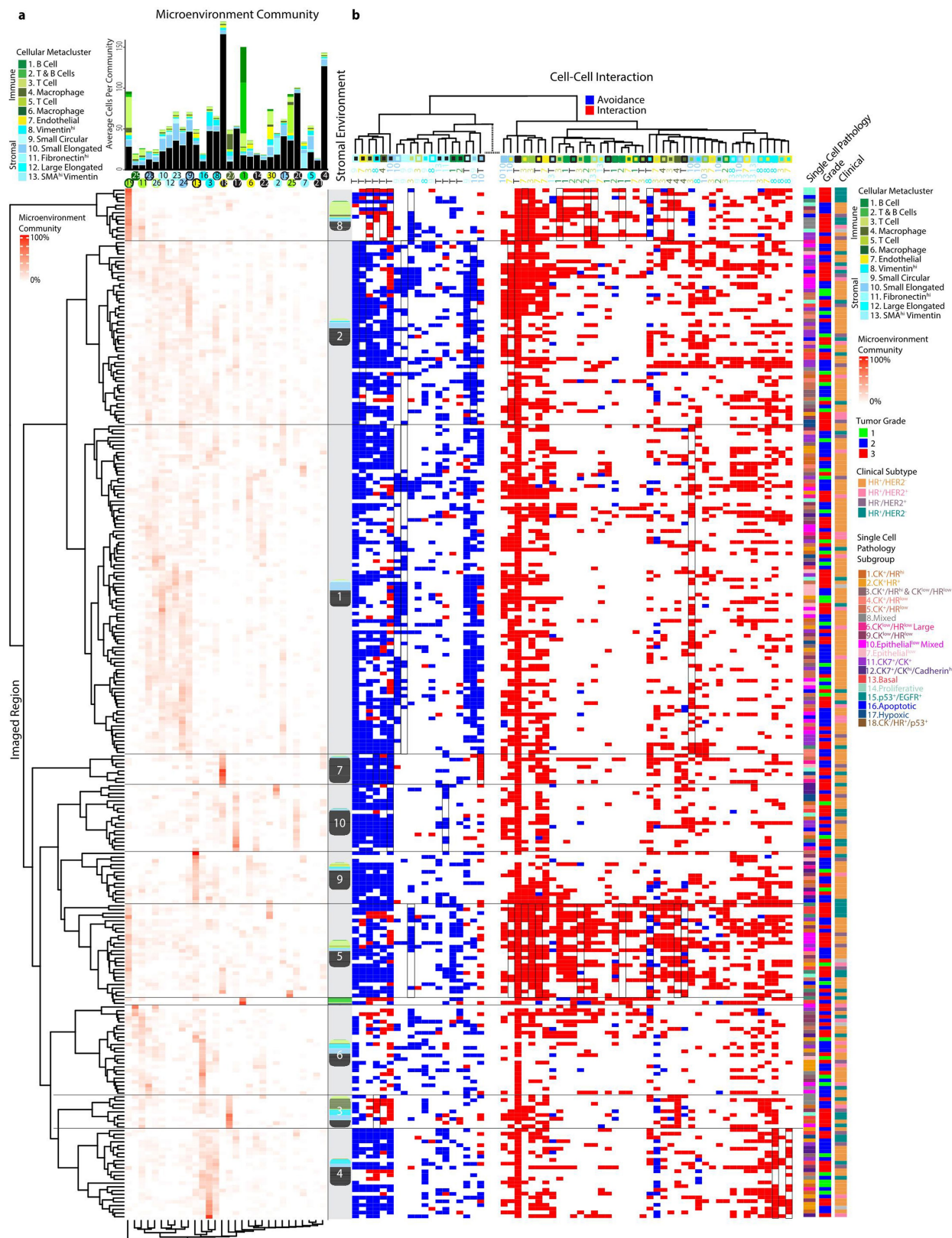**Extended Data Fig. 5 | Metaclustering and cluster matching across cohorts.**
a, Heat map showing z-scored mean marker expression of single-cell phenotypic clusters identified by PhenoGraph (Fig. 1) with colours on the colour bar and hierarchical clustering indicating the corresponding metacluster. Red stars on the hierarchical clustering tree indicate subgroups that robustly reappear as separate groups using multiscale bootstrap resampling (P < 0.05 (R function pvclust)). **b**, Examples of untransformed distributions of cluster marker expressions that differ between metaclusters.

c, Heat maps showing the z-scored mean marker expression or distance to the tumour–stroma interface for each metacluster defined in the cohort of 281 patients from University Hospital Basel and each matched PhenoGraph cluster from the multicore cohort of 71 patients from University Hospital Zurich. PhenoGraph clusters of the Zurich cohort were matched to the metaclusters of the Basel cohort on the basis of the Pearson correlation of the mean marker expression (Methods).
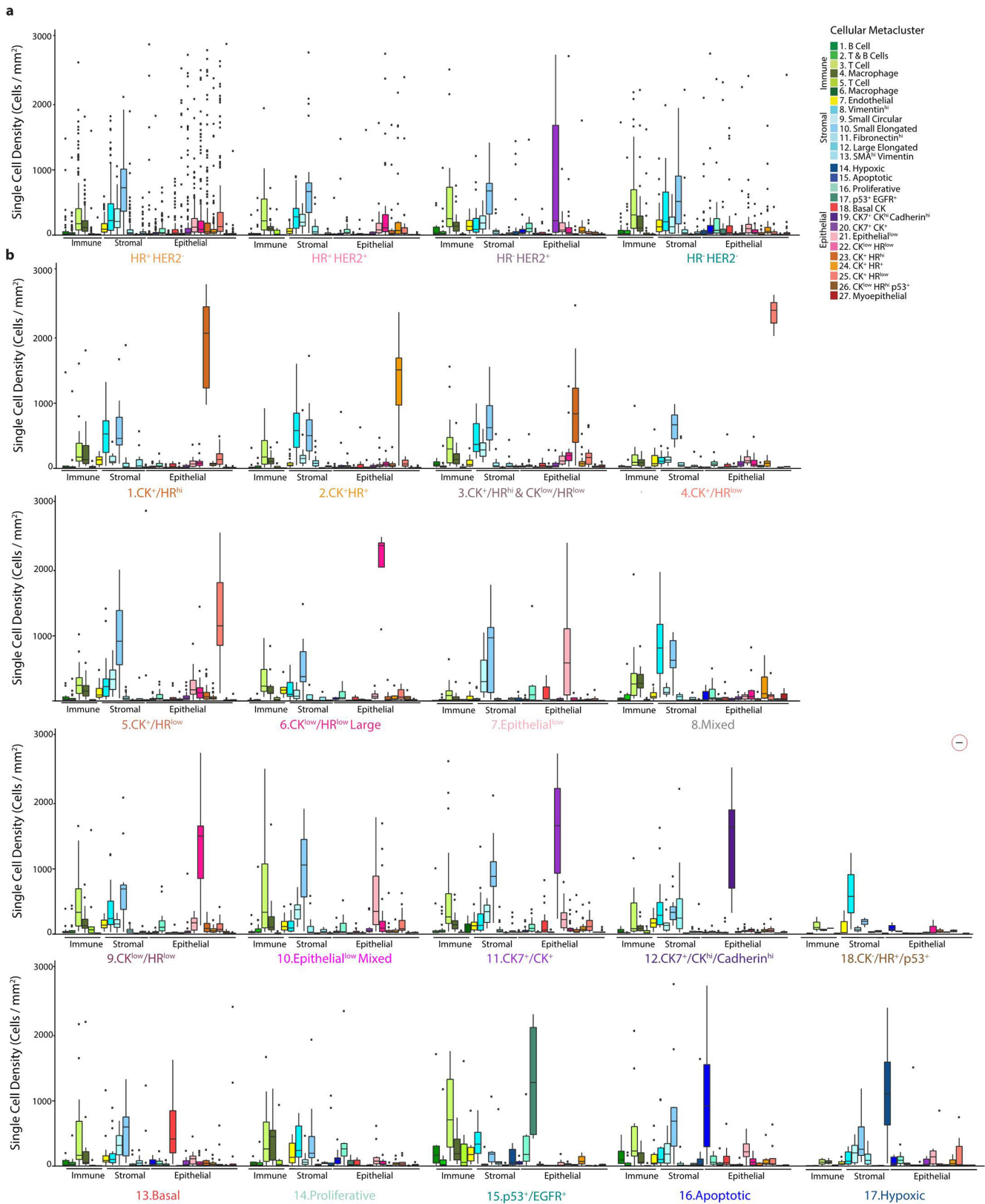
Extended Data Fig. 6 | Densities of single-cell phenotypes in different clinical subtypes and SCP patient subgroups. Box plots of cellular metacluster densities in patients of each clinical subtype (**a**) (HR⁺HER2⁻, n = 173; HR⁺HER2⁺, n = 29; HR⁻HER2⁺, n = 23; triple negative (HR⁻HER2⁻), n = 48) (**a**) and each SCP subgroup (**b**) (SCP1, n = 17; SCP2, n = 21; SCP3, n = 20; SCP4, n = 12; SCP5, n = 32; SCP6, n = 10; SCP7, n = 13; SCP8, n = 11; SCP9, n = 20; SCP10, n = 24; SCP11, n = 31; SCP12, n = 14; SCP13, n = 15; SCP14, n = 11; SCP15, n = 8; SCP16, n = 10; SCP17, n = 9; SCP18, n = 3). For box plots, centre line represents the median, box limits are the first and third quartiles, whiskers extend to 1.5 × interquartile range and the points beyond the whiskers are outliers.

Extended Data Fig. 7 | Stromal environments based on their composition of microenvironment communities and their distinct pairwise cell-type interactions. a, Hierarchical clustering of tumour cores ($n = 281$) according to stromal community content and splitting into corresponding stromal environments ($n = 11$). The stacked bar plot at the top indicates the average number of cells from each cellular metacluster present within each type of microenvironment community. **b**, The presence of significant ($P < 0.01$) cell–cell interactions (red) and cell–cell avoidances (blue) identified per image based on a permutation test (1,000 permutations). Black outlined regions indicate significant interactions that are enriched in images from the respective stromal environments ($P < 0.05$ (one-sided Fisher's exact test for enrichment, corrected for multiple testing). Colour bars on the right indicate the SCP subgroup, grade and clinical subtype of the tumour. Cell-type interactions along the top are indicated by the labelled cell type of interest and neighbouring cell.

Article
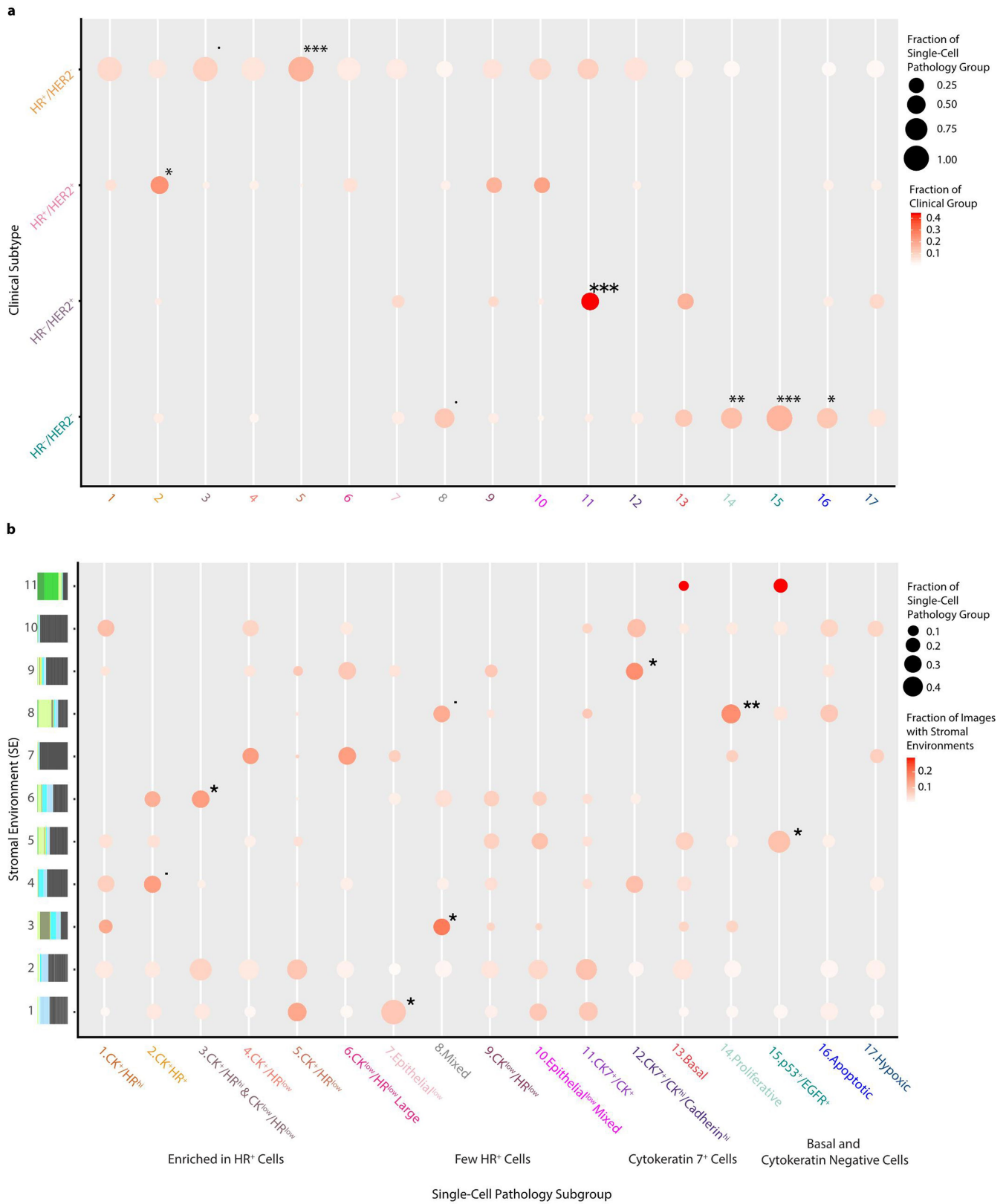
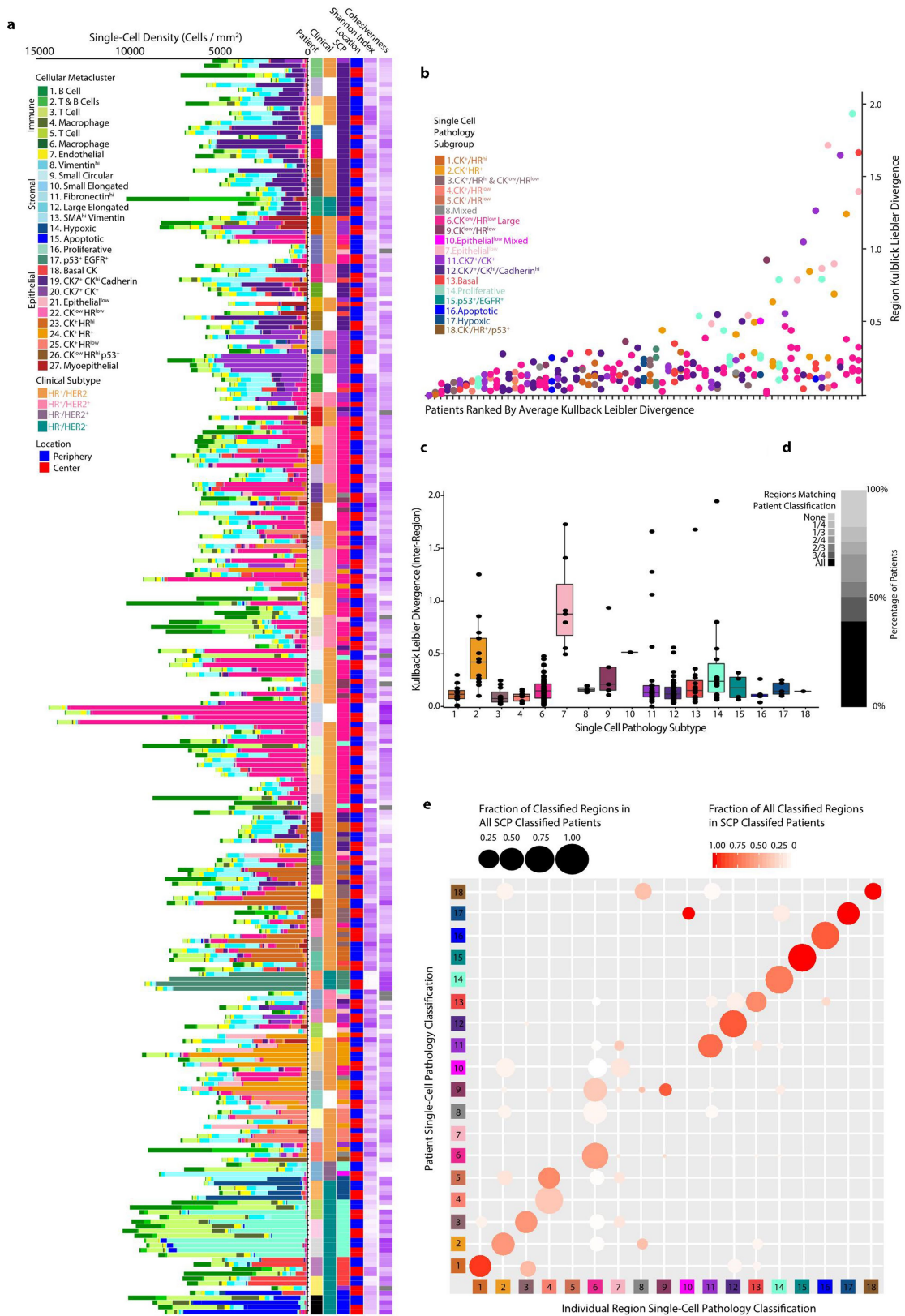**Extended Data Fig. 8** | See next page for caption.

Extended Data Fig. 8 | Comparisons and enrichments between classifications. a, **b**, Bubble plots visualizing the overlap between SCP subgroups of breast cancer (SCP1, $n=17$; SCP2, $n=21$; SCP3, $n=20$; SCP4, $n=12$; SCP5, $n=32$; SCP6, $n=10$; SCP7, $n=13$; SCP8, $n=11$; SCP9, $n=20$; SCP10, $n=24$; SCP11, $n=31$; SCP12, $n=14$; SCP13, $n=15$; SCP14, $n=11$; SCP15, $n=8$; SCP16, $n=10$; SCP17, $n=9$ (SCP18, $n=3$ excluded)) and clinical subtypes (**a**) (HR$^+$HER2$^-$, $n=173$; HR$^+$HER2$^+$, $n=29$; HR$^-$HER2$^+$, $n=23$; triple negative (HR$^-$HER2$^-$), $n=48$) and stromal environments (**b**) (SE1, $n=49$; SE2, $n=88$; SE3, $n=9$; SE4, $n=24$; SE5, $n=25$; SE6, $n=24$; SE7, $n=8$; SE8, $n=14$; SE9, $n=14$; SE10, $n=18$; SE11, $n=2$).

$\bullet P<0.1$, $*P<0.05$, $**P<0.01$, $***P<0.001$ (one-sided Fisher's exact test for enrichment.). Exact P values for highlighted pairs: SE1 and SCP7, $P=0.013$; SE3 and SCP8, $P=0.021$; SE5 and SCP15, $P=0.031$; SE6 and SCP3, $P=0.034$; SE8 and SCP14, $P=0.008$; SE8 and SCP8, $P=0.093$; SE9 and SCP12, $P=0.036$; HR$^+$HER2$^-$ and SCP3, $P=0.079$; HR$^+$HER2$^-$ and SCP5, $P=3.58\times10^{-4}$; HR$^+$HER2$^+$ and SCP2, $P=0.032$; HR$^-$HER2$^+$ and SCP11, $P=2.36\times10^{-4}$; HR$^-$HER2$^-$ and SCP8, $P=0.060$; HR$^-$HER2$^-$ and SCP14, $P=0.008$; HR$^-$HER2$^-$ and SCP15, $P=6.13\times10^{-6}$; HR$^-$HER2$^-$ and SCP16, $P=0.031$.

Article
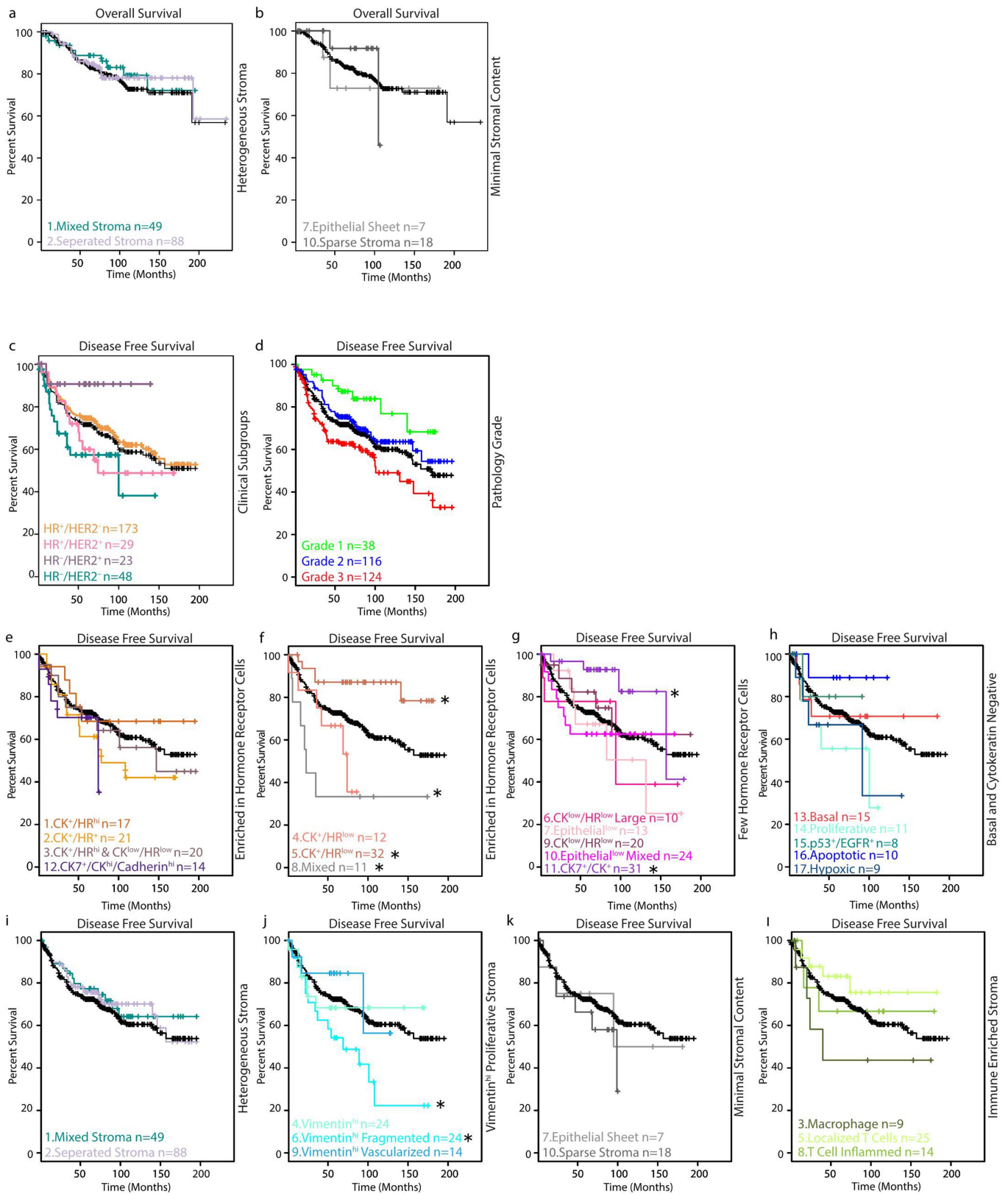
Extended Data Fig. 9 | Kaplan–Meier survival curves for overall and disease-free survival. a, **b**, Overall survival for stromal environments that are not shown in Fig. 4. **c–l**, Disease-free survival for each patient group on the basis of clinical subtype (**c**), grade (**d**), SCP subgroup (**e–h**) and stromal environment (**i–f**). *$P < 0.05$ compared to all other samples (two-sided log-rank test). For exact P values, see Supplementary Tables 5 and 7.

Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Multicore cohort regional heterogeneity analysis. Quantification of intratumour regional heterogeneity in the Zurich multicore cohort. **a**, Hierarchically clustered stacked bar plot of cell-type metacluster densities in each tumour, grouped by patient. Coloured columns (right) indicate patient, clinical subtype, SCP subgroup, location of core in the tumour, Shannon entropy (intracore heterogeneity) and tumour-specific cohesiveness score. **b**, Dot plot of the Kullback–Leibler divergence from the cell-type distribution of an individual tumour region to the average distribution of the patient, coloured according to the SCP subgroup classification per tumour region ($n = 263$ tumour cores), grouped by patient ($n = 71$ patients) and ordered by increasing average Kullback–Leibler divergence per patient. **c**, Box plots of the same Kullback–Leibler divergence of each region to the average cell-type distribution of the patient, grouped by tumour regions that are individually identified as the same SCP subgroup, independent of the patient. Box plots as in Extended Data Fig. 6. SCP1, $n = 12$; SCP2, $n = 13$; SCP3, $n = 11$; SCP4, $n = 10$; SCP6, $n = 76$; SCP7, $n = 7$; SCP8, $n = 3$; SCP9, $n = 5$; SCP10, $n = 1$; SCP11, $n = 26$; SCP12, $n = 51$; SCP13, $n = 15$; SCP14, $n = 18$; SCP15, $n = 4$; SCP16, $n = 5$; SCP17, $n = 5$; SCP18, $n = 1$). **d**, Bar indicating the percentage of patients ($n = 71$) with the indicated fraction of individually classified images that match the whole tumour classification. **e**, Bubble plot visualizing the variation in intratumour regions within patients of each SCP subgroup. Rows represent tumours of each SCP subgroup as identified by the combined analysis of all imaged regions. Columns represent tumour regions individually matched to a SCP subgroup. For each whole-tumour classification on the y axis, the size of the circle indicates the fraction of corresponding images individually classified as a SCP subgroup. For each image classification on the x axis, colour indicates the fraction of images within each tumour type.

nature research

Corresponding author(s):   Bernd Bodenmiller

Last updated by author(s):   Oct 24, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	*Provide a description of all commercial, open source and custom code used to collect the data in this study, specifying the version used OR state that no software was used.*
Data analysis	All custom code used for this study is available in the linked github repository. The repository contains code written in Matlab version 2017b and R version 3.5. For segmentation and single-cell feature extraction CellProfiler version 2.1.1, Ilastik version 1.1.9 and histoCAT version 1.74 were used. CATALYST version 1.5.6 was used for spillover compensation. The PhenoGraph 2.0 implementation in histoCAT was used for single-cell clustering. The R package pvclust 2.0 was used for multiscale bootstrap resampling of the hierarchy subtrees during metaclustering, in order to assess the uncertainty of each subtree. For clustering on patients the cytofkit R implementation of PhenoGraph (version 1.10.0) was used. The R package entropy (version 1.2.1) was used for the calculation of both Shannon entropy and the Kullback-Leibler divergence. The Louvain community detection algorithm (C implementation by Lefebvre and Guillaume, version 0.2, wrapped by Matlab as used by the implementation of PhenoGraph 2.0 used by histoCAT/Cyt) was applied to identify highly interconnected spatial subunits in the tissue graph. Kaplan-Meier survival curves and coxph survival regression models were generated using the R package survival (version 2.42-4).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The image, single-cell and patient data supporting the findings of this study were uploaded to figshare during resubmission (via the Manuscript Tracking System). All code that led to the results of this study will be available on https://github.com/BodenmillerGroup/SCPathology_publication upon publication.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	In this investigative, first of its kind study, 281 breast cancer patients were analyzed. A second cohort containing 344 images from 71 patients was also studied and matched to the original cohort. As it was not possible to estimate the number of cellular phenotypes present before analysis we were not able to specify sample size.
Data exclusions	No data was excluded from this study. All samples that were successfully stained and imaged were used in our data analysis.
Replication	All cell types, patient groups and interactions were identified in multiple patient samples including multiple samples from the same patient. Further, a second patient cohort, acquired and processed at an independent institute was subject to identical analysis for the validation of initial classifications.
Randomization	In the first cohort from the University Hospital Basel, patients were not selected for any clinical or histological features, therefore patients in the cohort are representative of the population. In the second cohort, from the University Hospital Zurich, patients were selected to be equally divided between the three pathology grades of breast cancer and within each grade half did and half did not have metastasis.
Blinding	All samples were stained simultaneously. Image acquisition order was distributed spatially and blinded to clinical data.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	All Antibody information is available in Supplemental Information Table #2
Validation	All antibodies were validated by immunofluorescence imaging prior to isotope-polymer conjugation. Antibodies were tested for cell type and inter-cell location specificity within positive control tissues including lymph nodes, breast tumors, and healthy breast tissue.

Human research participants

Population characteristics	All patient data is available as a .csv raw data file and summarized in Supplemental Information Table 1 and 9.
Recruitment	The specimens derived from patients diagnosed with primary breast cancer between 1991 and 2013 at the Institute of Pathology at the University Hospital Basel and between 2004 and 2013 at the Institute of Pathology and Molecular Pathology at the University Hospital Zurich.
Ethics oversight	Project to use samples from the University Hospital Basel were approved by the Ethikkommission Nordwest- und Zentralschweiz (EKNZ-2014-397) and for samples to use from the Institute of Pathology and Molecular Pathology, University Hospital Zurich were approved by the Ethikkommission Kanton Zürich (KEK-2012-553).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	Histopathological data was obtained from the individual pathology reports
Outcomes	Clinical and survival data were extracted from the hospital database or from the patients' attending physician