

Supplementary Material

Integrating 12 Spatial and Single Cell Technologies to Characterise Tumour Neighbourhoods and Cellular Interactions in three Skin Cancer Types

P. Prakrithi^{1,2#}, Laura F. Grice^{1,2,3#}, Feng Zhang^{1,2#}, Levi Hockey^{1,2}, Samuel X. Tan⁴, Xiao Tan^{1,2}, Zherui Xiong^{1,2}, Onkar Mulay^{1,2}, Andrew Causer^{1,2}, Andrew Newman^{1,2}, Duy Pham¹, Guiyan Ni¹, Kelvin Tuong⁵, Xinnan Jin^{1,2}, Eunju Kim^{3,4}, Minh Tran¹, Hani Vu^{1,2}, Nicholas M. Muller⁴, Emily E. Killingbeck⁶, Mark T. Gregory⁶, Siok Min Teoh¹, Tuan Vo¹, Min Zhang⁷, Maria Teresa Landi⁸, Kevin M. Brown⁸, Mark M. Iles⁹, Zachary Reitz⁶, Katharina Devitt¹⁰, Liuliu Pan⁶, Arutha Kulasinghe⁶, Yung-Ching Kao⁴, Michael Leon⁶, Sarah R. Murphy⁶, Hiromi Sato⁶, Jazmina Gonzalez Cruz¹⁰, Snehlata Kumari¹⁰, Hung N. Luu¹¹, Sarah E. Warren⁶, Chris McMillan^{12, 13}, Joakim Henricson^{14,15}, Chris Anderson^{14,15}, David Muller^{12,13}, Arun Everest-Dass¹⁶, Blake O'Brien¹⁷, Mathias Seviiri¹⁶, Matthew H. Law^{18, 19, 20}, H. Peter Soyer⁴, Ian Frazer¹⁰, Youngmi Kim^{6,21}, Mitchell S. Stark⁴, Kiarash Khosrotehrani⁴, Quan Nguyen^{1,2}

¹ QIMR Berghofer, Immunology and Infection program, Brisbane, QLD 4006, Australia

² Division of Genetics and Genomics, Institute for Molecular Bioscience, The University of Queensland, QLD 4072, Australia

³ School of Biomedical Sciences, The University of Queensland, QLD 4072, Australia

⁴ Frazer Institute, Dermatology Research Centre, Experimental Dermatology Group, The University of Queensland, Brisbane, Australia

⁵ Ian Frazer Centre for Children's Immunotherapy Research, Child Health Research Centre, Faculty of Medicine, The University of Queensland, Brisbane, Queensland, Australia

⁶ Bruker Spatial Biology, NanoString® Technologies, Seattle WA 98109, USA

⁷ School of Agriculture and Food Sciences, The University of Queensland, QLD 4072, Australia

⁸ Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

⁹ Leeds Institute for Data Analytics, University of Leeds, Leeds, UK

¹⁰ Frazer Institute, The University of Queensland, Brisbane, QLD 4072, Australia

¹¹ UPMC Hillman Cancer Center & School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA

¹² School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, QLD 4072, Australia

¹³ Australian Infectious Diseases Research Centre, School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, QLD 4072, Australia

¹⁴ Department of Emergency Medicine in Linköping, and Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden

¹⁵ Department of Dermatology and venereology in Östergötland, and Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden

¹⁶ Institute for Biomedicine and Glycomics, Griffith University, Gold Coast, QLD 4215

¹⁷ Sullivan Nicolaides Pathology, Bowen Hills, Qld 4006, Australia

¹⁸ QIMR Berghofer, Population Health program, Brisbane, QLD 4006, Australia

¹⁹ School of Biomedical Sciences, Faculty of Health, Queensland University of Technology, Brisbane, QLD, 4001, Australia

²⁰ School of Biomedical Sciences, Faculty of Medicine, University of Queensland, Brisbane, QLD, 4072, Australia

²¹ Pfizer, 21621 30th Dr SE, Bothell, WA 98021

#co-first authors

Correspondence: quan.nguyen@qimrberghofer.edu.au

50 **Table of Contents**

51	Supplementary Figures	2
52	List of Supplementary Tables	32
53	Supplementary Methods	39
54	Supplementary Notes	54
55	Note S1: High-throughput profiling of cSCC protein and RNA with GeoMx	54
56	Note S2: Polaris for a targeted study of interactions via PD-1 and PD-L1	56
57	References	58
58		

59 **Supplementary Figures**

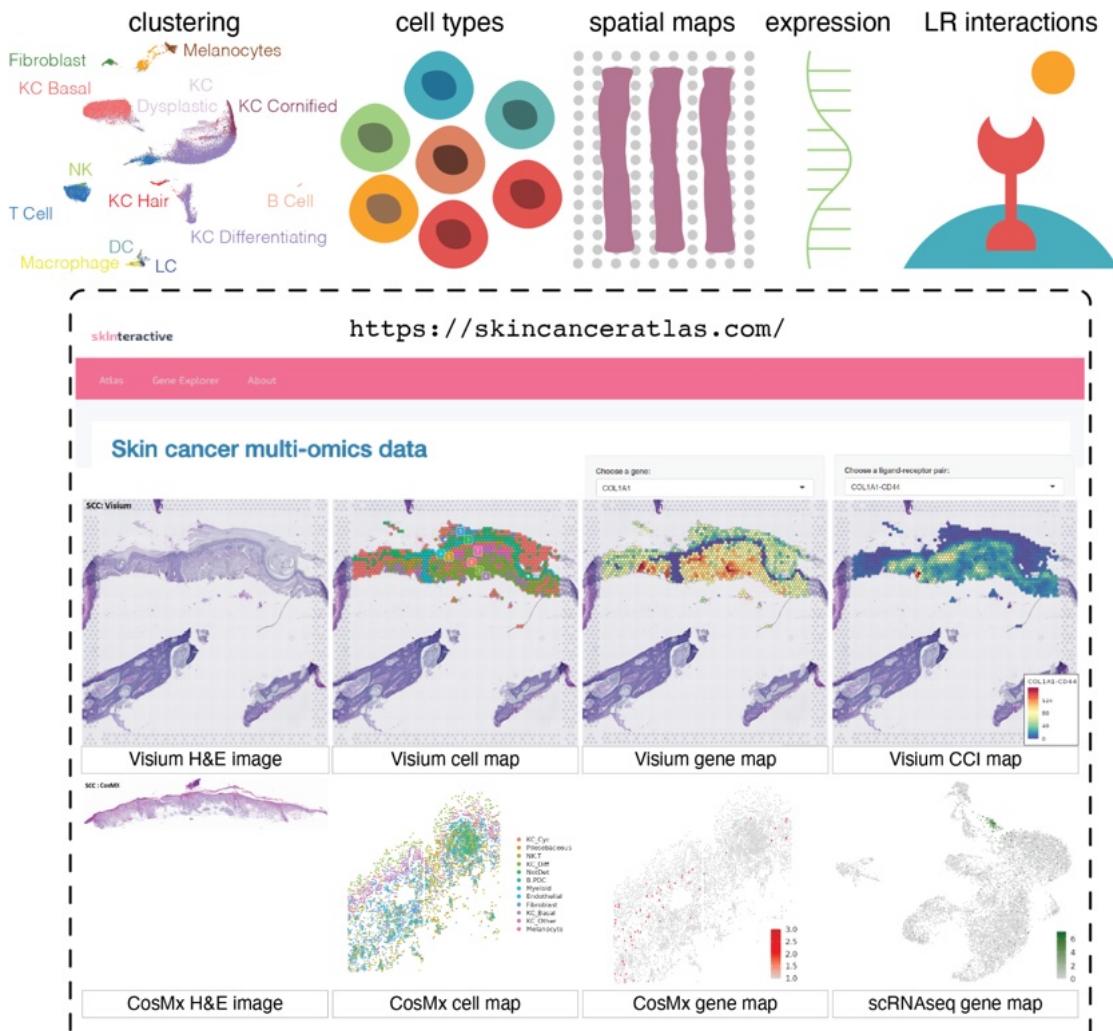
60

61

62

63

sklteractive: A spatial single-cell multi-omics atlas and interactome for skin cancer



64

65

66

67

68

69

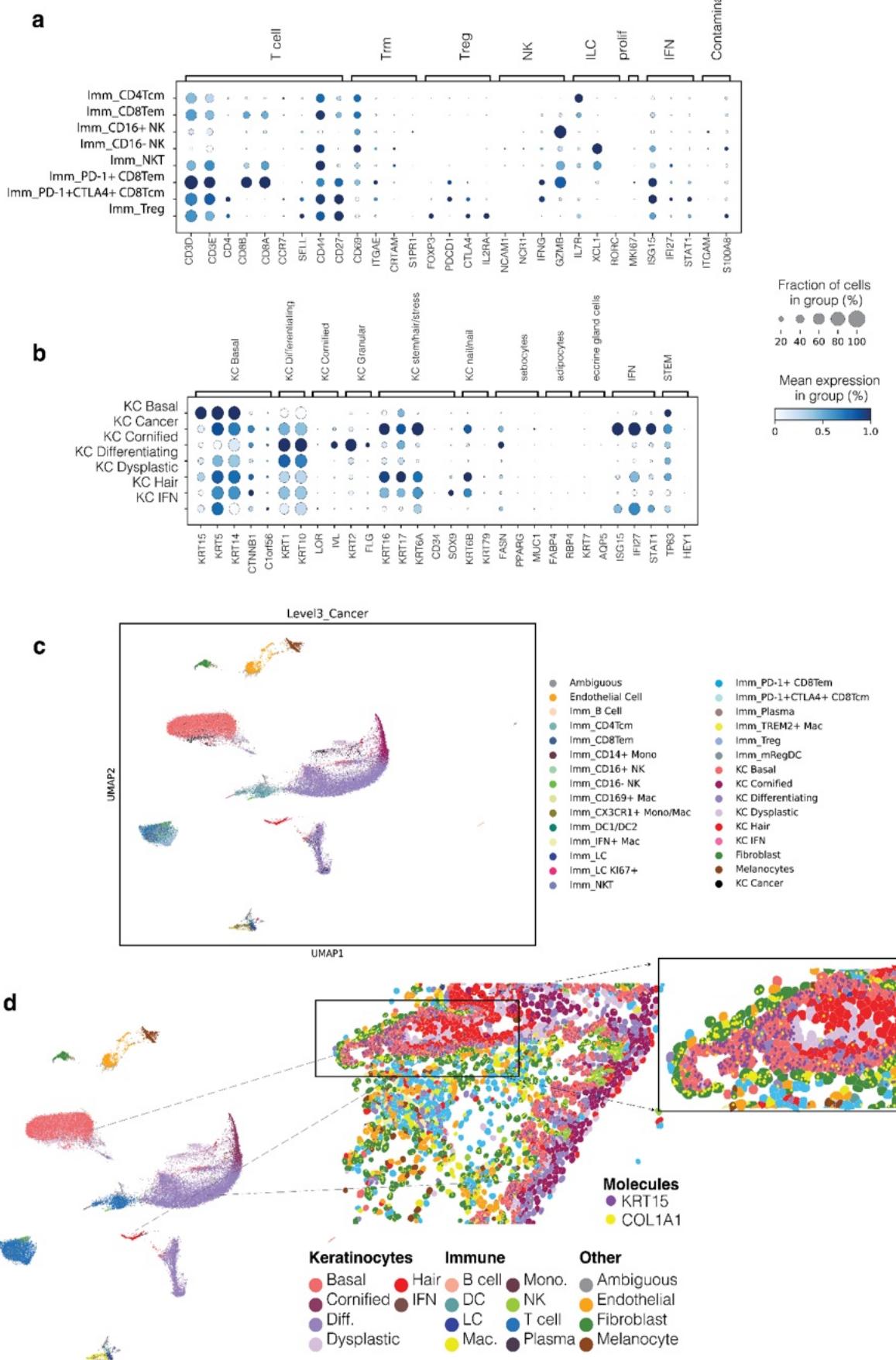
70

71

72

Figure S1. The skincanceratlas database allows interactive analysis of genes, cells, interactions across the multi-omics skin cancer data resource.

The skincanceratlas function allows users to visualise clustering, cell type and spatial data for cSCC, BCC and melanoma. The Gene Explorer function (bottom) allows visualisation of gene expression and LR interactions at single cell and/or spatial resolution from data generated with Visium, CosMx and scRNASeq for all three skin cancer types.



73

Fig S2 - Annotation of KC and Immune cells in BCC-cSCC.
(a) Marker expression for the level 3 annotation of immune cells.

74

75

76 (b) Marker expression for the level 3 annotation of KC cells.
77 (c) UMAP displaying the overall level 3 annotation of all 30 cell types.
78 (d) Visualisation of the single cell (sub)clusters that are mapped to the tissue using the CosMX data.
79 scRNAseq data was used as the reference to transfer cell annotation labels to CosMX cells using the
80 RCTD deconvolution method. The KC basal, KC hair, KC differentiating, KC cornified clearly form
81 distinct layers in the cSCC skin sample B18. The cell types defined in scRNAseq data clearly mapped
82 to distinct anatomical layers of the skin, providing cross-platform evidence for correct cell type annotation
83 of the scRNAseq reference data, which did not have a spatial context.

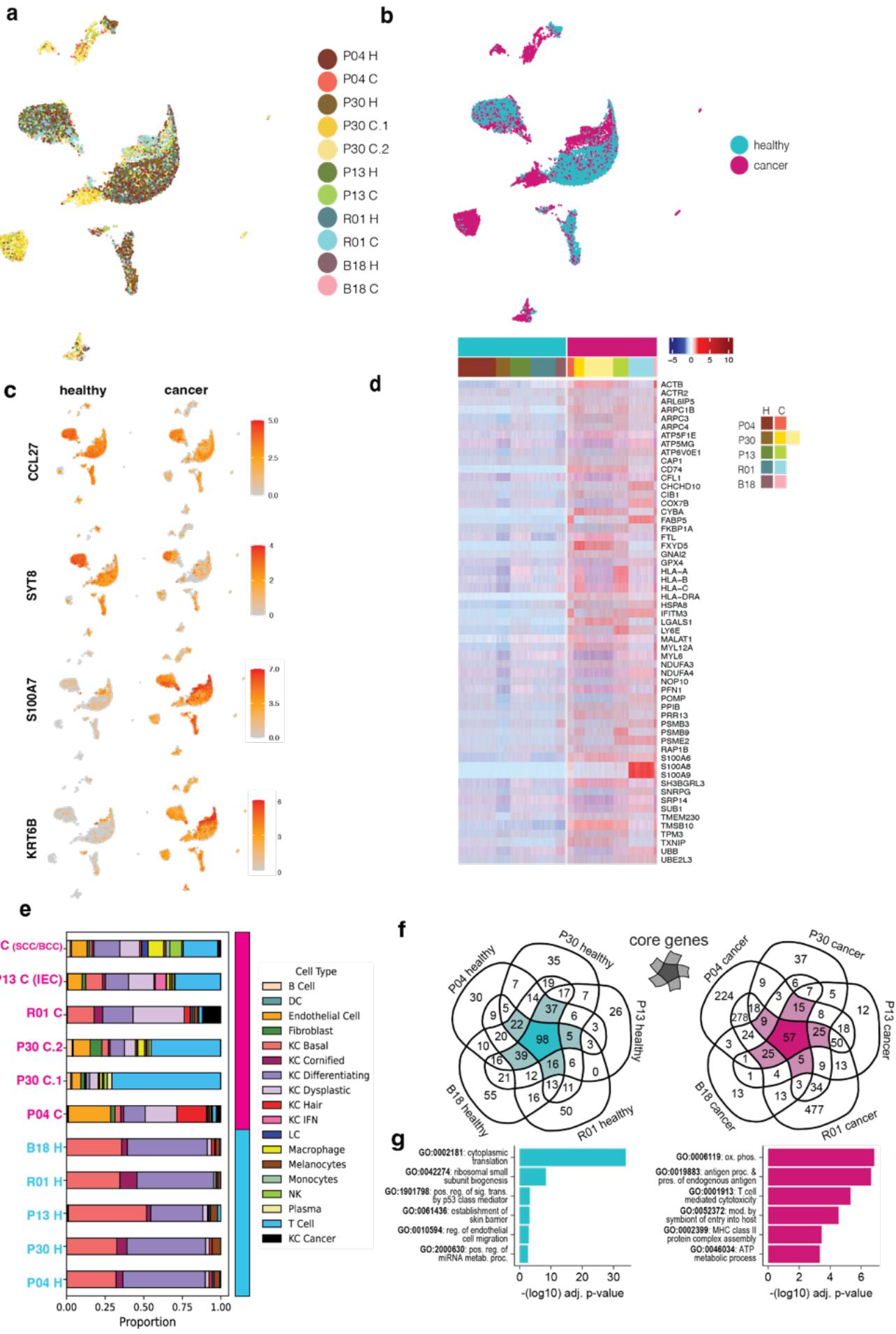
84

85

86

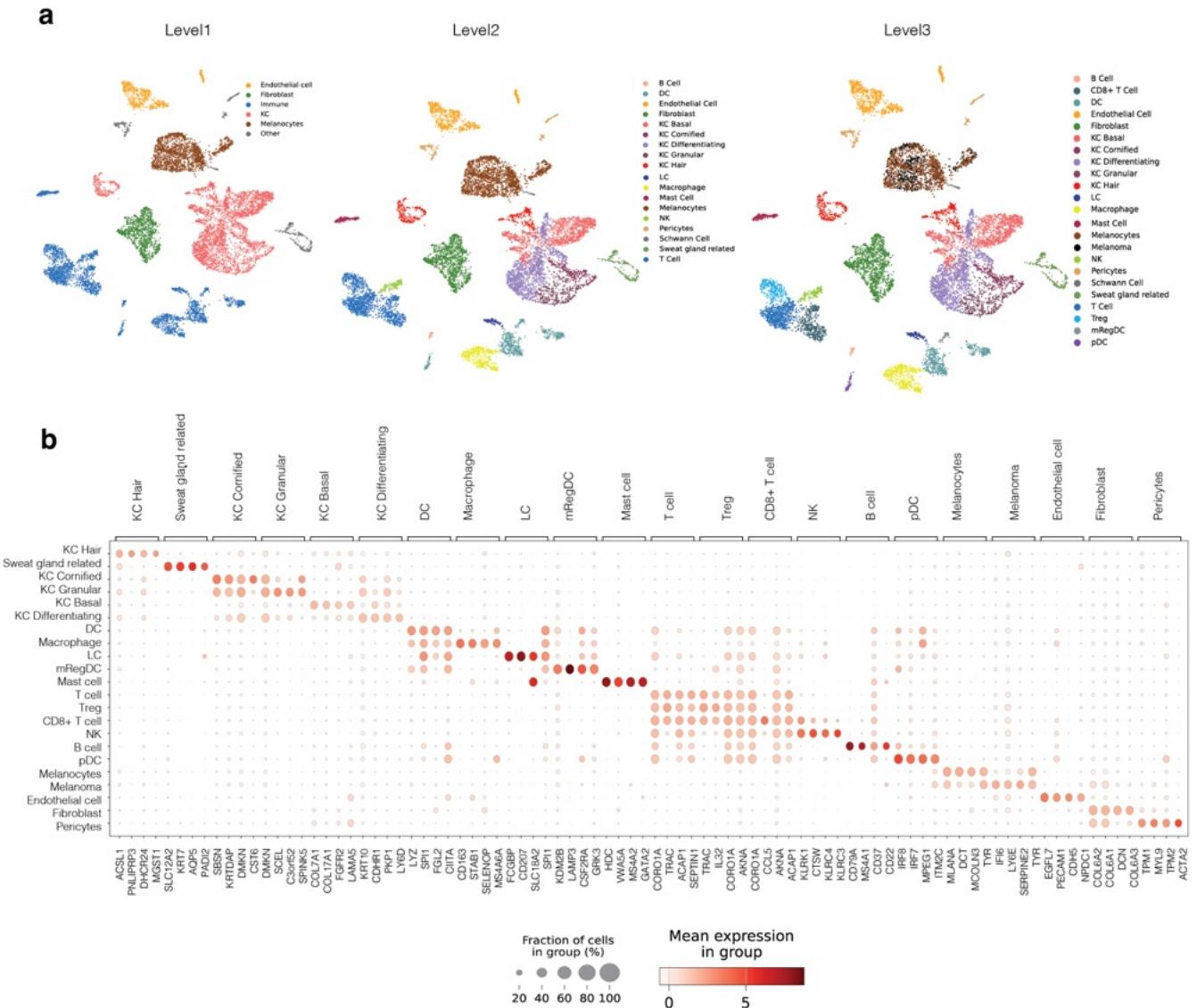
87

88



90 **Figure S3. A single-cell atlas of cSCC cell types**
91 (a-b) UMAP plots showing the integration of ~46,000 cells from 11 samples of five cSCC patients,
92 showing cell distribution by patient sample (a) and cancer status (b).
93 (c) Expression of selected genes that are amongst the top 10 most differentially expressed genes
94 between cancer and healthy cells when compared across the entire dataset. Each column shows only
95 the healthy (left) or cancer (right) cells split from the full UMAP plot shown in (a-b).
96 (d) Heatmap showing expression in healthy and cancer samples of the 57 core cancer genes shared
97 between all five patients (as in Panel F).
98 (e) Bar plot showing the distribution of each cell type across eleven samples.
99 (f) Genes overexpressed in healthy (left) or cancer (right) cells across individual patients. Core suites of
100 genes that are detected across four or more of the five patients are highlighted in the centre of each
101 diagram; suites contain a total of 217 healthy and 136 cancer genes respectively.
102 (g) The top six non-redundant gene ontology terms enriched in the two gene suites, ranked by -log10
103 adjusted P-value.

104
105
106
107
108
109
110



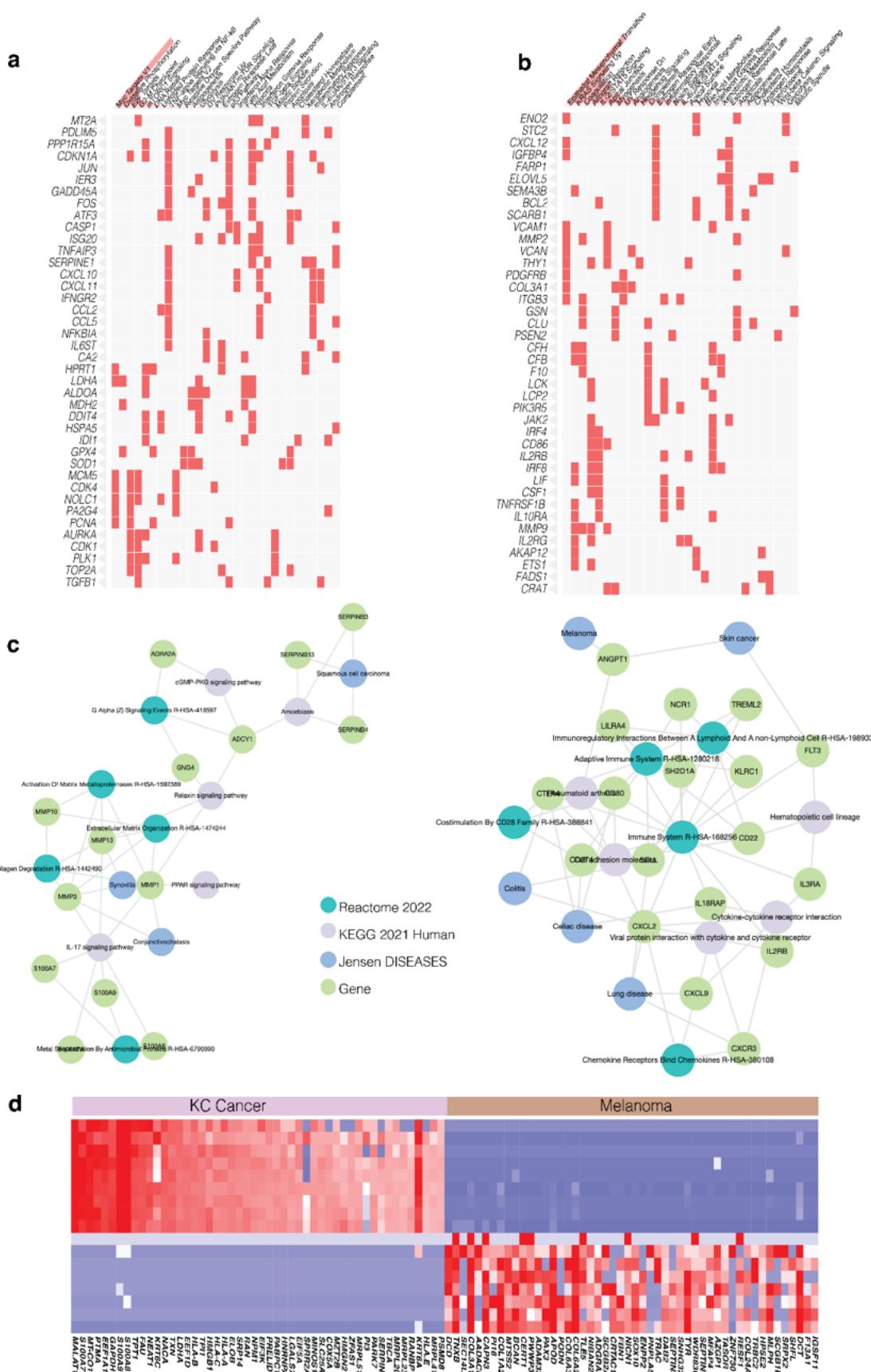
111

112

113 Fig S4 - Annotation of melanoma samples

114 (a) UMAP shows the annotation of all the three annotation levels from common to specific cell types.
 115 UMAP showing Level 3 annotation also displays melanoma cells, which were not defined by Level 1
 116 and Level 2 annotation.

117 (b) Heatmap showing marker expression for the level 3 annotation.



118
119
120

Fig S5 - Pathway analysis for gene markers of cSCC-BCC and melanoma

(a) Heatmap showing enriched pathways for 3267 genes KC Cancer Vs melanoma, tested against the

121 MSigDB. All DE genes upregulated in Cancer cells in cSCC-BCC compared to normal KC cells (left)
122 and in melanomas compared to normal melanocytes were used. Pathway names are shown in columns
123 and genes in these pathways are shown in rows. The red bars indicate -log(P-value) and the terms are
124 sequentially ordered with the most significant pathway on the left.

125 (b) MSigDB pathways for 2713 genes up in Melanoma Vs KC Cancer.

126 (c) Integrated pathway analysis, using KEGG and Reactome databases for genes higher in cancer KCs
127 compared to the normal KCs in cSCC-BCC samples observed in scRNA-seq (left; 176 KC Cancer vs
128 normal KCs) and for genes higher in malignant melanocytes compared to the normal melanocytes in
129 snRNAseq Melanoma samples (68 genes up in Melanoma Vs normal melanocytes).

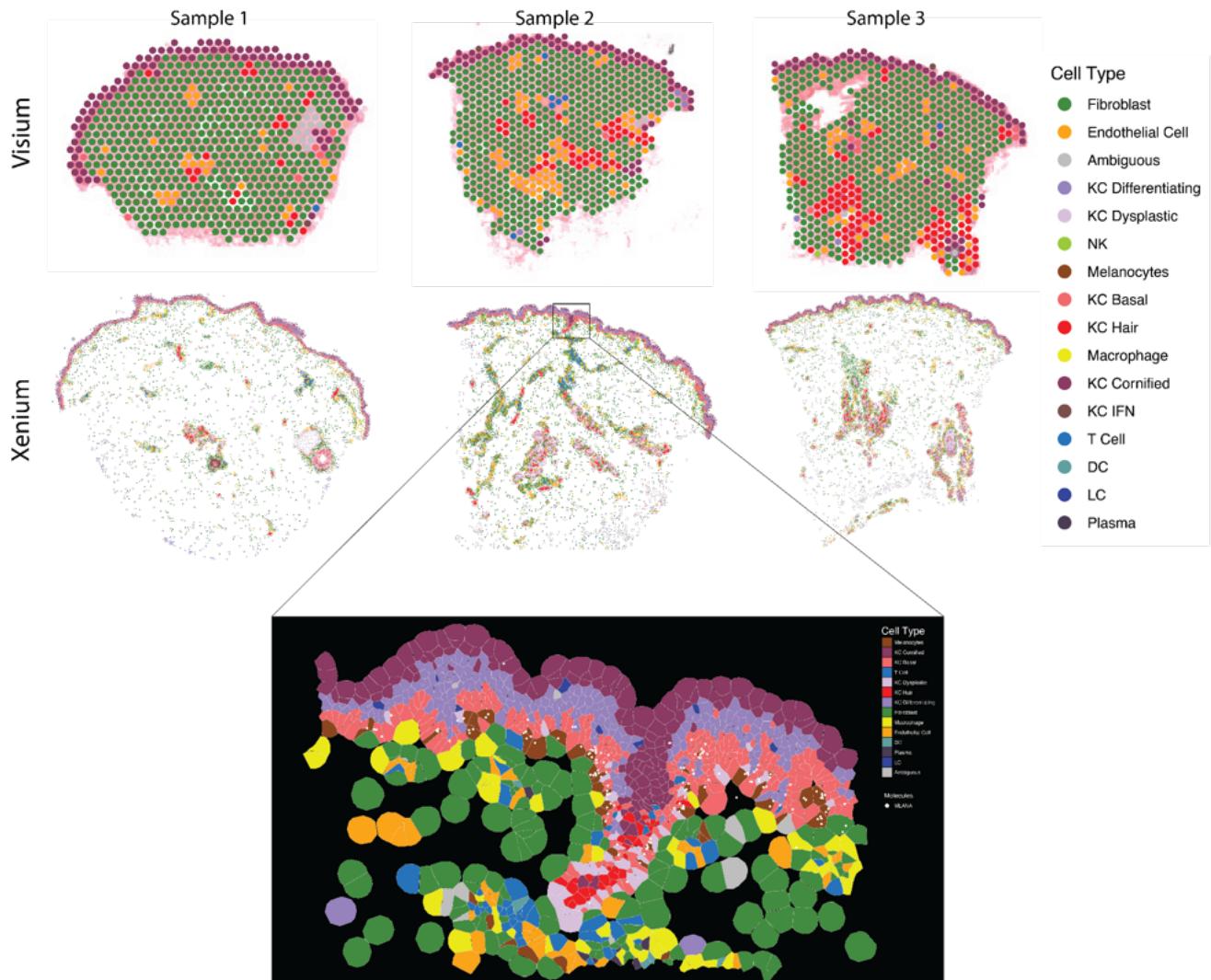
130 (d) DE genes from comparing KC cells and Melanocytes.

131

132

133

134



135

136

137

Figure S6. Cell type annotation for the Visium and Xenium data of samples from three normal control donors.

138

139

140

The donors did not have skin cancer. The top and middle panels show Visium and Xenium data of adjacent tissue sections, respectively, and a zoom-in view of a Xenium FOV is shown at the bottom panel. The legend colors indicate cell types.

141

142

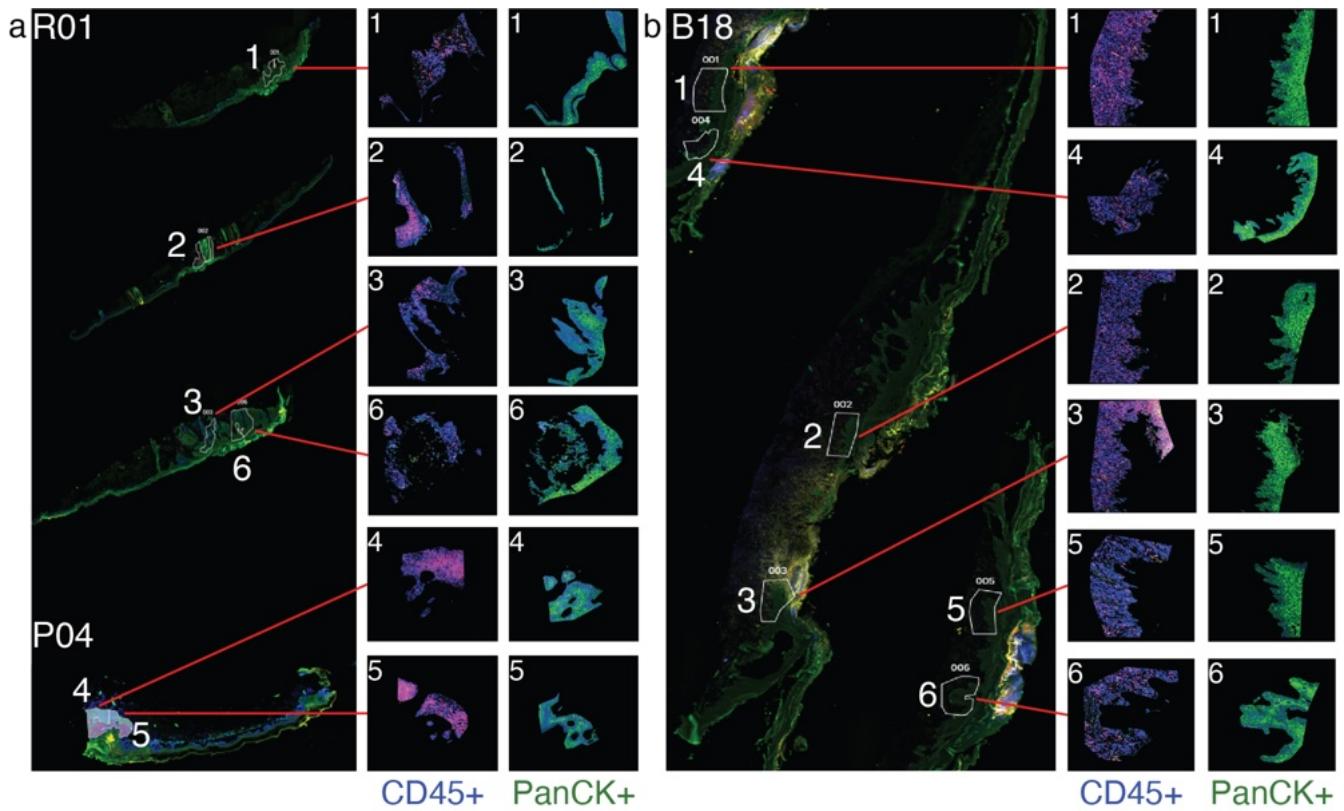


Figure S7. Captured regions of interest for GeoMx protein quantification

Morphology imaging of cSCC biopsies for GeoMx protein data, from patients R01 (a, top 3 tissues), P04 (a, bottom tissue) and B18 (b). Boxes show zoomed-in regions of interest (ROIs) selected for marker quantification, connected to their location in the full tissue image by red lines. Morphology markers show DNA (blue, DAPI), CD3 (red), CD45 (yellow) and PanCK (green). ROIs were segmented by either CD45 or PanCK expression.

143
144
145
146
147
148
149
150

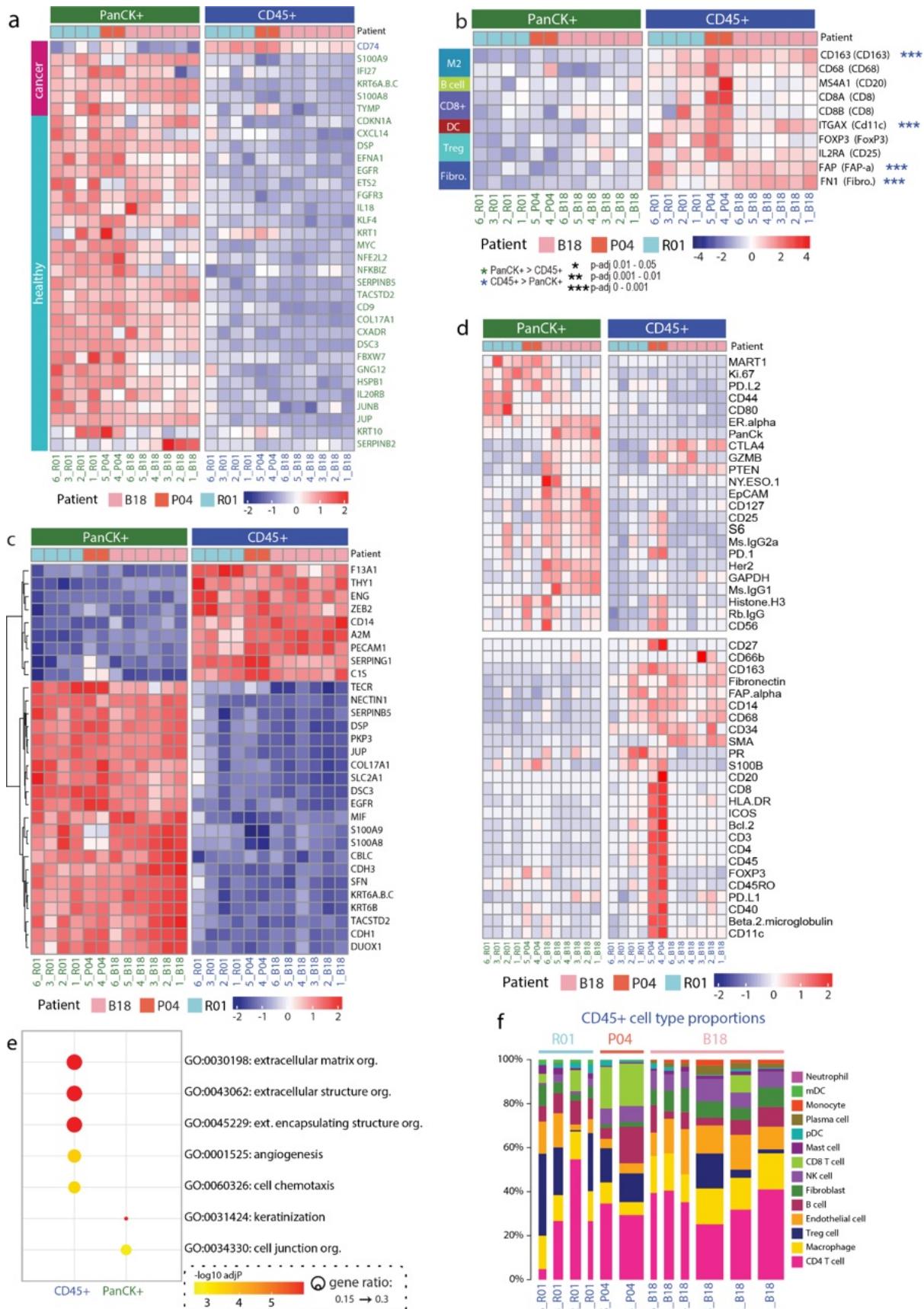
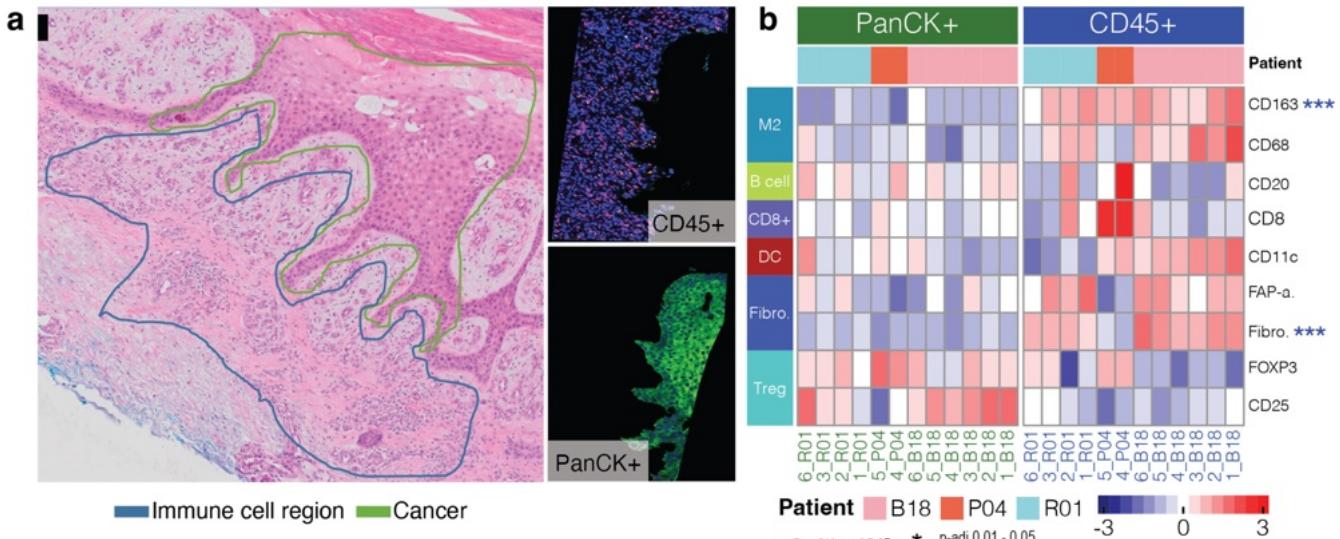


Figure S8. Analysis of GeoMx gene markers and cell type compositions using GeoMx RNA and protein assays.

(a) Heatmap of 33 genes which were differentially expressed between PanCK+ and CD45+ segments within the GeoMx RNA dataset, and were also identified as members of the core cancer- or healthy-associated gene suites. Column annotations separate samples by GeoMx segment (PanCK+ and

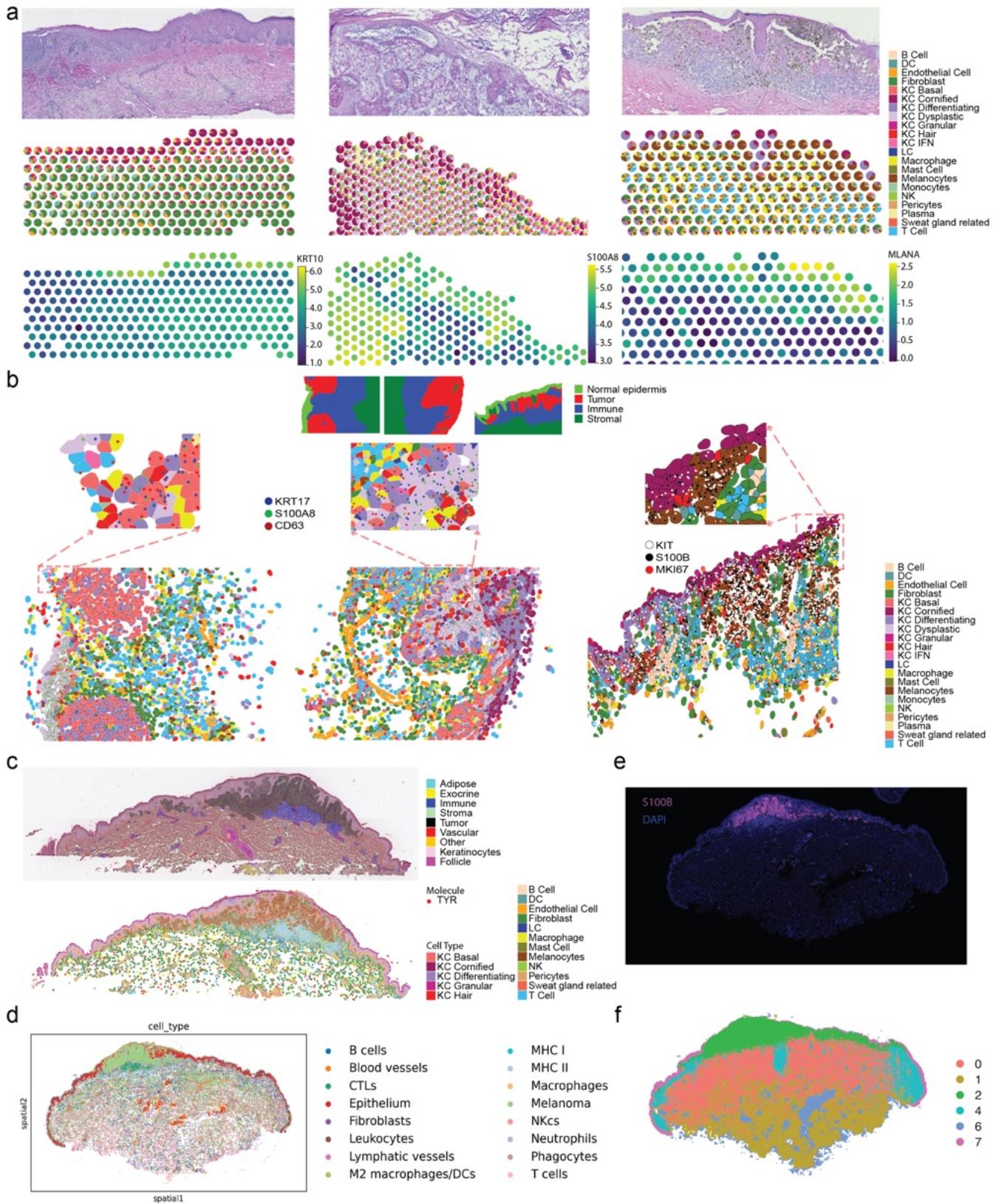
157 CD45+ segments) and patient, while row annotation indicates whether genes were present in the
158 cancer- or healthy-associated gene suites from **Fig S3f**. Gene label colour indicates whether the gene
159 was upregulated in the CD45+ (blue) or PanCK+ (green) segment in the GeoMx analysis.
160 (b) Validation of the presence of key cell types from our scRNASeq atlas, using RNA markers captured
161 by GeoMx; this plot is complementary to **Fig S9n** which shows the same comparison for equivalent
162 protein markers. Selected markers are CD163 and CD68 (M2 markers), MS4A1 (B cell marker), CD8A
163 and CD8B (CD8+ T cell markers), ITGAX (dendritic cell marker), FOXP3 and FOXP3 (Treg markers),
164 and FAP and FN1 (fibroblast markers). Corresponding protein markers for each gene are given in
165 brackets. Top annotation bars separate samples by segment (PanCK+ and CD45+ segments) and
166 patient. Marker association with particular cell types was taken from NanoString product information.
167 Asterisks indicate markers that were differentially expressed between CD45+ and PanCK+ segments;
168 markers that were not found to be differentially expressed are included as their expression provides
169 evidence for the presence of these cell types.
170 (c) Top 30 differentially expressed genes between PanCK+ and CD45+ segments within the GeoMx
171 RNA dataset. A total of 267 differentially expressed genes were identified between segments.
172 Differentially expressed genes were sorted by FDR and the top 30 genes were selected regardless of
173 whether they were enriched in the PanCK+ or CD45+ segments.
174 (d) Expression of all the 48 proteins (Immuno-oncology panel) in GeoMX proteomics data. The heatmap
175 shows normalised data across FOVs.
176 (e) Top 5 GO enrichment results per segment from an analysis of the top 100 differentially expressed
177 genes associated with the CD45+ and PanCK+ segments of the GeoMx RNA data. Differentially
178 expressed genes were sorted by absolute fold-change values and the top 100 genes were selected from
179 each of the CD45+ and PanCK+ segments.
180 (f) Inferred cell type proportions of ROIs from patients B18, R01 and P04 in the CD45+ segment of the
181 RNA GeoMx data. RNA expression information was used to perform cell type deconvolution for the
182 segmented ROIs. Bar width is proportional to the number of nuclei per ROI. Top annotations indicate
183 the sample of origin for each ROI.



187
 188 **Figure S9. Spatial mapping of cancer cell types within cSCC skin cancer as measured by GeoMx**
 189 **spatial transcriptomics**

190 (a) Left, H&E stained image of cSCC tissue from patient B18, with pathologist's annotation indicating
 191 cancer (green) and immune cell (blue) tissue regions. This annotation was used in combination with
 192 morphology marker staining to define regions of interest (ROIs) for GeoMx spatial proteomics. Right,
 193 selected ROIs corresponding to the highlighted regions (ROI 002, patient B18), segmented based on
 194 CD45 (top) and PanCK (bottom) markers.

195 (b) Validation of the presence of key cell types from our scRNASeq atlas, using protein markers captured
 196 by GeoMx. Selected markers are CD163 and CD68 (M2 markers), CD20 (B cell marker), CD8 (CD8+ T
 197 cell marker), CD11c (dendritic cell marker), FOXP3 and CD25 (Treg markers), FAP alpha and
 198 Fibronectin (fibroblast markers). Top annotation bars separate samples by segment (PanCK+ and
 199 CD45+ segments) and patient. Marker association with particular cell types was taken from Nanostring
 200 product information. FAP-a. = FAP-alpha, Fibro. = Fibronectin. Asterisks indicate markers that were
 201 differentially expressed between CD45+ and PanCK+ segments; markers that were not found to be
 202 differentially expressed are included as their expression provides evidence for the presence of these
 203 cell types.



222

223 **Fig S10** Integrative analysis of spatial multiomics atlas for melanoma - Cross-modality spatial mapping
224 across three cancer types225 (a) Visium mapping for BCC (left), cSCC (middle), and Melanoma (right), with Hematoxylin and Eosin
226 (H&E) staining (top), cell type annotation (middle), and marker expression levels (bottom).227 (b) Spatial localisation of marker gene expression in CosMx data from BCC (left), cSCC (middle), and
228 melanoma (right). Cell type annotations and marker expression are shown together (bottom), with
229 selected regions highlighted in enlarged panels (middle). Pathological annotations are provided in the

230 top thumbnails.
 231
 232 (c) Pathological annotation of the Xenium melanoma sample (48974-2B) on H&E image (top), and
 233 corresponding cell type annotation and TYR marker gene expression shown for the same region
 234 (bottom).
 235 (d) Cell type annotation for CODEX data on melanoma tissue
 236 (e) Immunofluorescent staining of S100A8 protein marker corresponds to the tissue region containing
 237 melanoma cells in Panel e.
 238 (f) Cell type annotation for spatial Glycomics data on an adjacent tissue section from the same block as
 239 shown in Panels d-e.
 240
 241

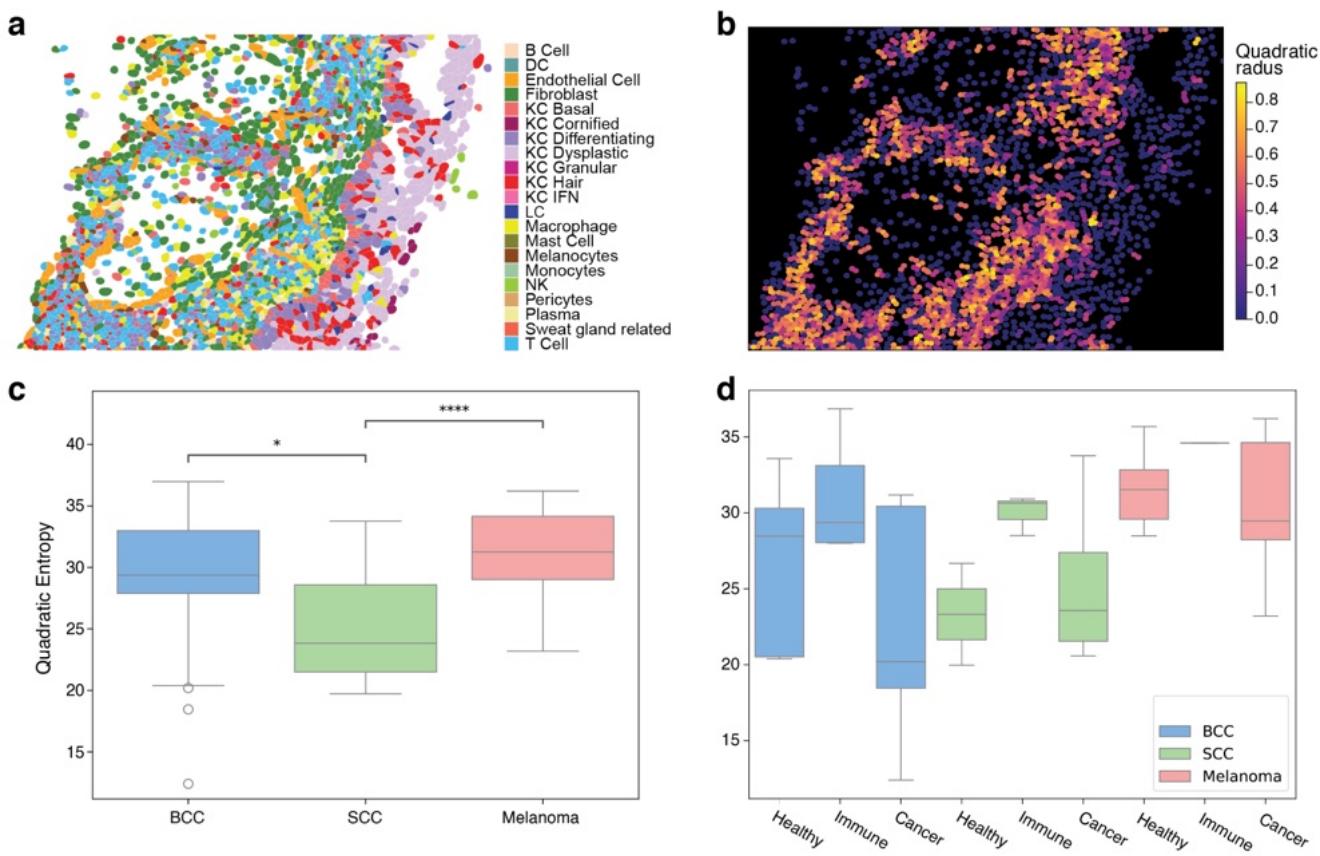
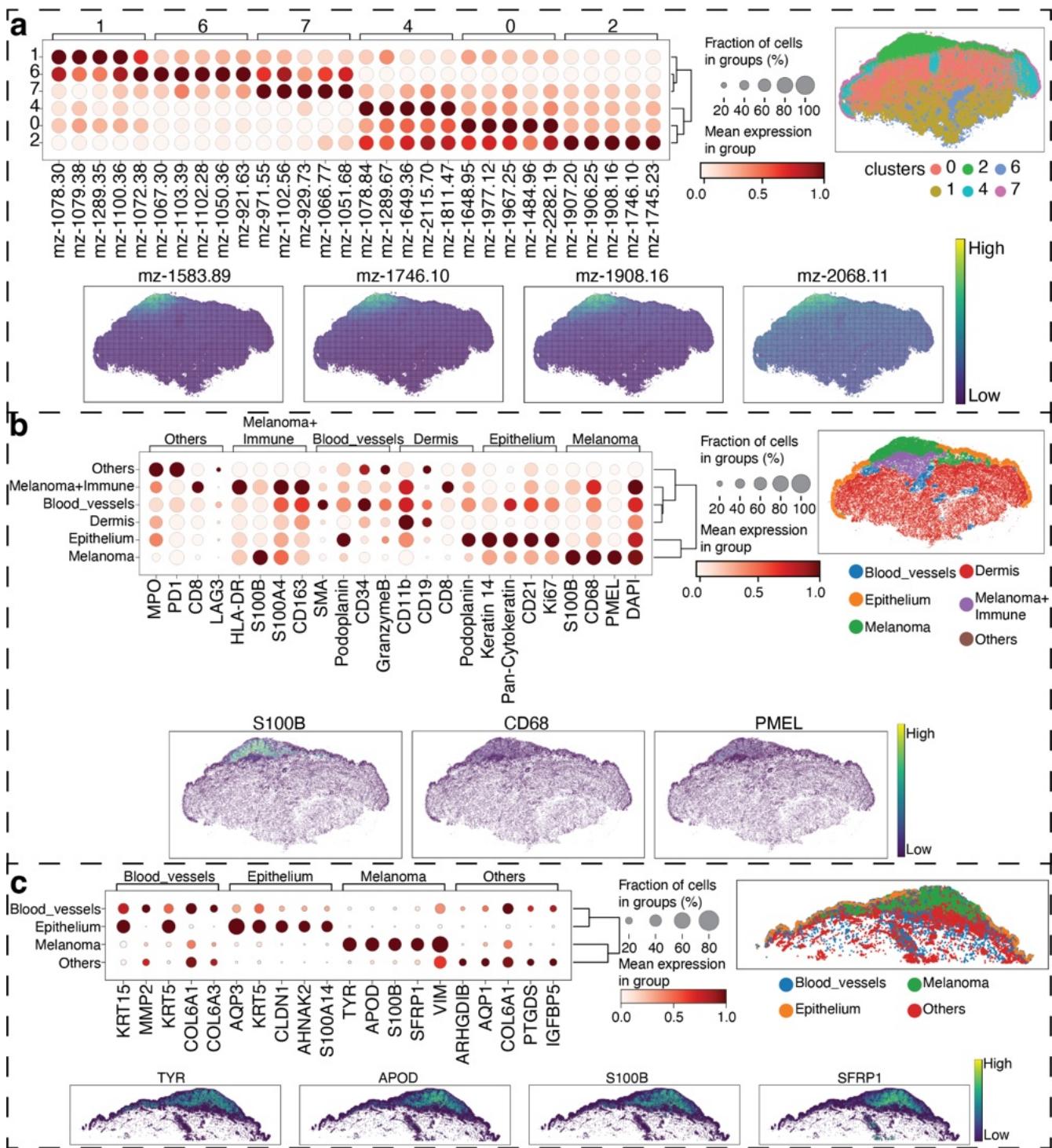


Figure S11: Spatial cellular heterogeneity across cancer types using CosMx

242
 243 (a) Spatial cell type annotations of cSCC patient B18 FOV 15, the FOV with the highest overall
 244 heterogeneity score for this sample.
 245
 246 (b) Spataial cell heterogeneity plot for the FOV shown in Panel a. Cells are coloured by heterogeneity
 247 score (measured by cell neighbourhood analysis), highlighting regions of the FOV containing diverse
 248 cell types in close proximity.
 249
 250 (c) Boxplot of heterogeneity scores across three cancer types, indicating the lowest overall
 251 heterogeneity score for cSCC. Asterisks indicate significant p-values from paired Wilcoxon test.
 252
 253 (d) Boxplot of heterogeneity scores across tissue regions of each cancer type. Each FOV was
 254 designated as being either healthy, immune or cancer by a pathologist. Wilcoxon ranked sum test was
 255 applied.



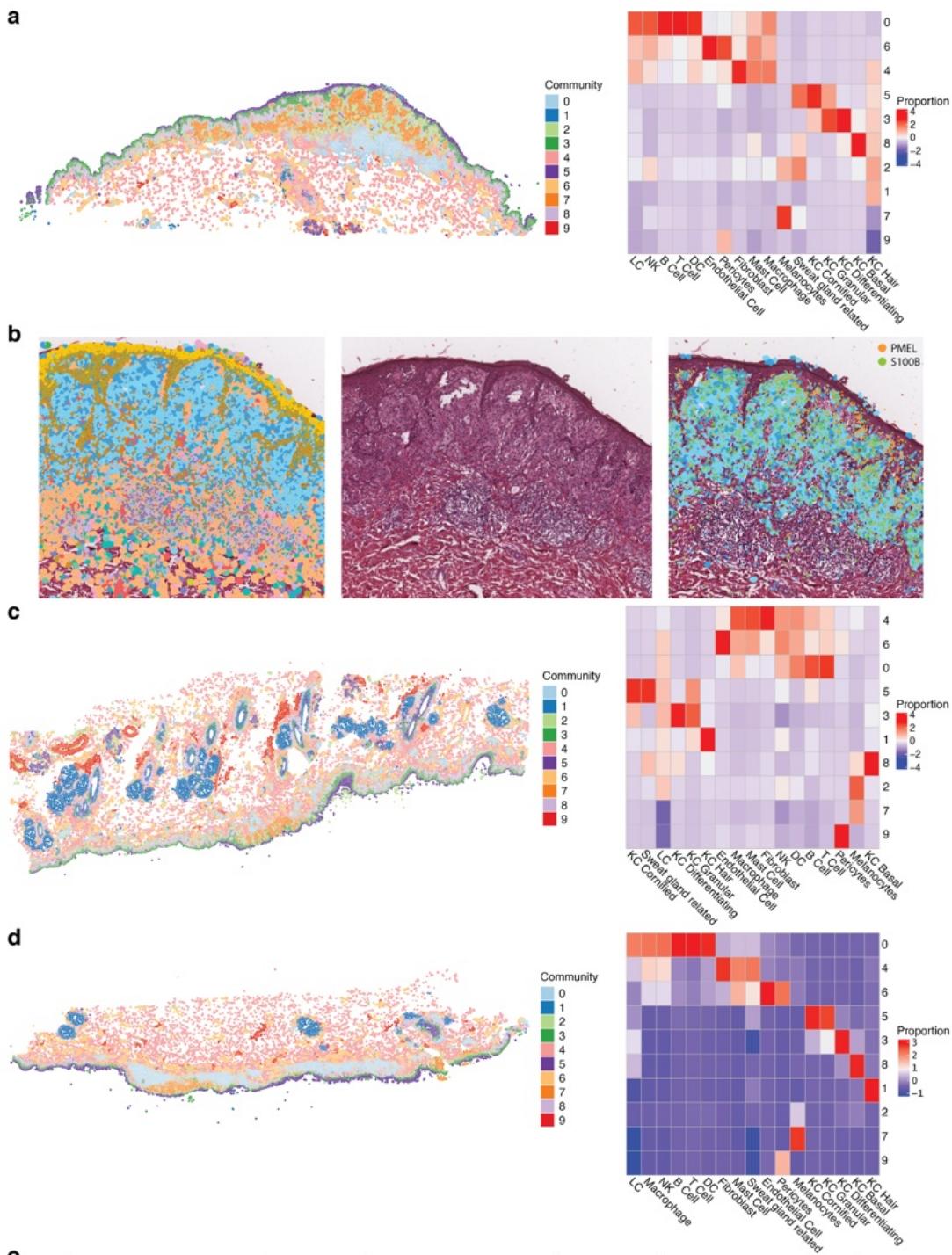
256

257 S12. Integrative analysis of spatial transcriptomics, proteomics, and glycomics data for 258 melanoma community.

259 (a) Joint pathway analysis using DE genes/proteins of melanocyte communities in Xenium and CODEX
260 data, and glycans of melanocytes in MALDI data, to map to KEGG metabolic pathways. The X-axis
261 shows enriched genes/proteins in Xenium and CODEX data, while the Y-axis shows enriched glycans.
262 (b-c) Cell type proportions of the community identified in CODEX data (b) and Xenium data (c). The
263 melanoma community in both datasets is enriched with melanocytes.
264 (d) Clustering of the melanoma spatial glycomics sample, with the identification of the melanoma-
265 enriched cluster 2. A heatmap of top markers for each cluster is on the left, and a tissue plot with cluster
266 label is shown on the right. The expression for four representative metabolites is shown at the bottom
267 on the panel.

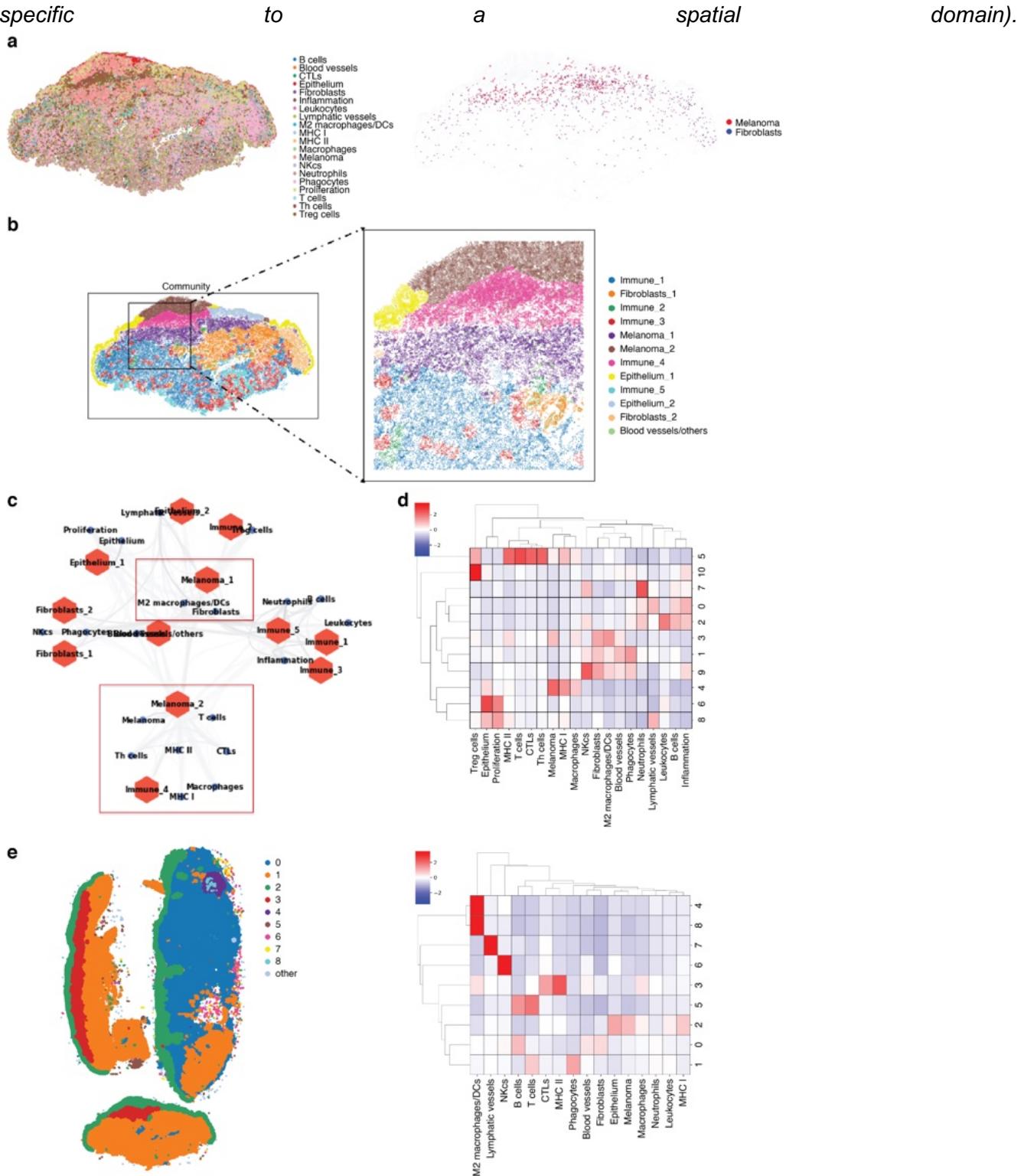
268 (e) Annotation for CODEX data from the adjacent tissue section, showing the melanoma cluster, with a
269 heatmap of protein markers for each cell type, the tissue plot showing all key cell types, and the the
270 expression of the three proteins at the bottom.
271

272 (f) Similar to d) and e) the annotation of the melanoma adjacent tissue section for the Xenium data is
273 shown.
274



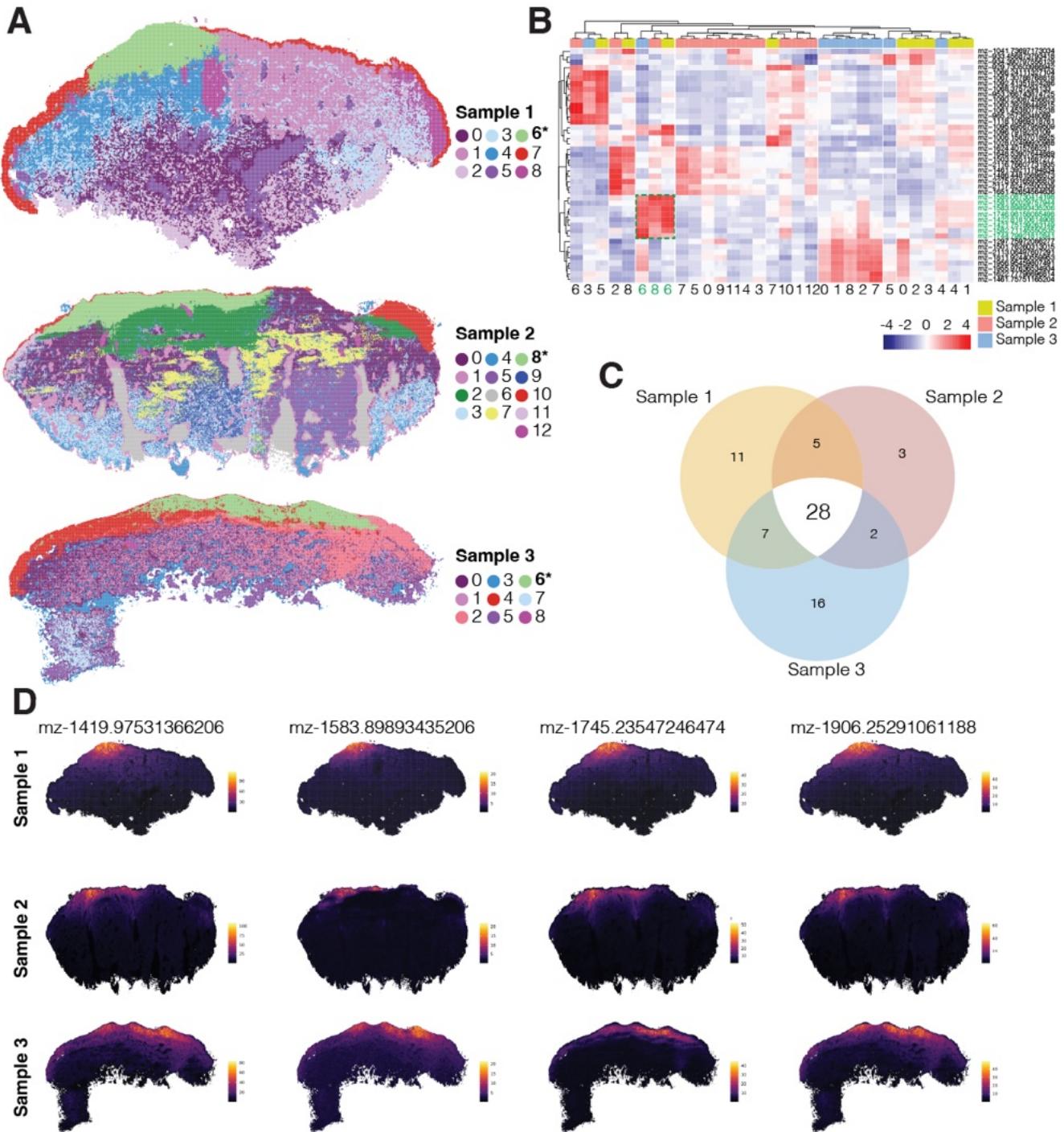
275
276 **Fig S13: A spatial community mapping of three melanoma samples for Xenium data**
277 (a) Community mapping for one sample based on cell type composition. The heatmap shows cell type
278 composition per community.
279 (b) Expression of marker genes at single cell resolution, showing cell type (left), H&E image, and
280 expression.

281 (c) Communities and cell type composition for two other melanoma samples.
 282 (d) Cell-cell interaction analysis of the melanoma community. The line shows co-occurrence score with
 283 melanocytes, where a high score indicates the cells are closely distributed with melanocytes across the
 284 whole tissue. The scores change when the distance to search for neighbouring cells increases (i.e., less
 285 specific spatial domain).



286
 287 **Fig S14: A spatial community mapping of three melanoma samples for CODEX data**
 288 (a) A detailed cell type annotation and an example of colocalization between two cell types (Melanoma
 289 and Fibroblasts) are shown.
 290 (b) Community mapping with a zoom-in view of communities, highlighting two melanoma communities
 291 (shown as melanoma-1 and melanoma-2).

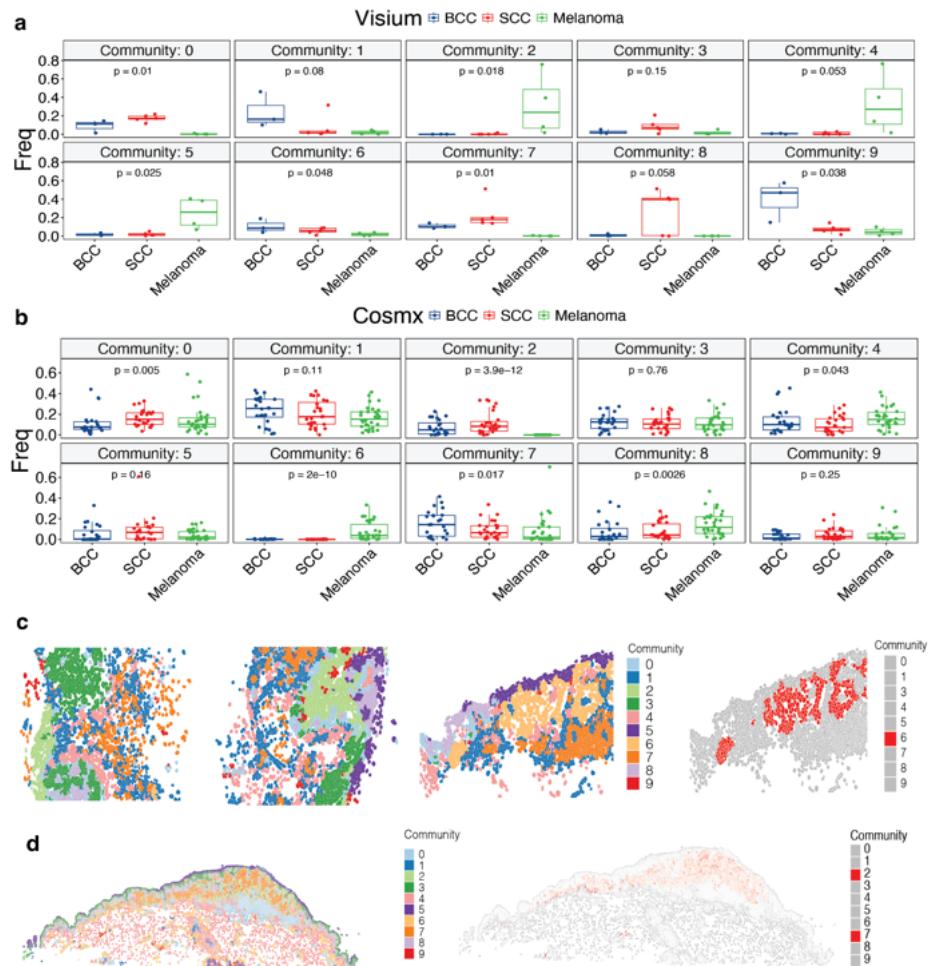
292 (c) Community network analysis, with most similar communities connected in a nearest neighbourhood
 293 (d) Examples of cell-type composition for each community of all three melanoma samples
 294 (e) Spatial distribution of communities, defined from merged analysis of all three samples.
 295



296
 297 **Figure S15. Analysis of spatial glycomics data.**

298 (a) Melanoma community analysis using spatial glycomics data. The colors show clustering results.
 299 Identification of a melanoma cluster (cluster 6 in samples 1 and 3, cluster 8 in sample 2; indicated with
 300 an asterisk) is based on histology and relative location compared to clusters from CODEX data.
 301 (b) Heatmap showing the expression of top metabolite markers (named by the mass m/z values) for
 302 each cluster identified in A. Top annotation bar shows sample of origin. The green box and associated
 303 green labels highlight the metabolites enriched across the three melanoma clusters.
 304 (c) Venn diagram highlighting the similarities and differences between markers identified for the
 305 melanoma cluster in each sample.

306 (d) Expression of the top four metabolite markers shared between samples. The top metabolites are the
307 most differentially expressed compared to all other clusters.
308
309



310
311 **Figure S16. Integrative analysis across different individuals and technologies to build a robust**
312 **spatial community atlas.**

313 (a-b) Boxplots showing the proportion of each community deriving from BCC, cSCC and melanoma in
314 Visium (a) and CosMx (b) data. Wilcoxon ranked sum test was applied.

315 (c) Spatial distribution of ten communities in CosMx data in representative FOVs from BCC (patient B18,
316 FOV 13, left), cSCC (patient B18, FOV 14, second left) and melanoma (patient 48974-2B, FOV 12,
317 second right). Pathological annotations for these FOVs are given in **Fig S10b**. Spatial location of cells
318 from CosMx_6, a member of the melanoma meta-community (Fig 5), is shown next to the original
319 distribution of all communities (right).

320 (d) Spatial distribution of ten communities in Xenium melanoma sample (right) relative to the location of
321 cells from Community Xenium_2 from the melanoma meta-community (left).

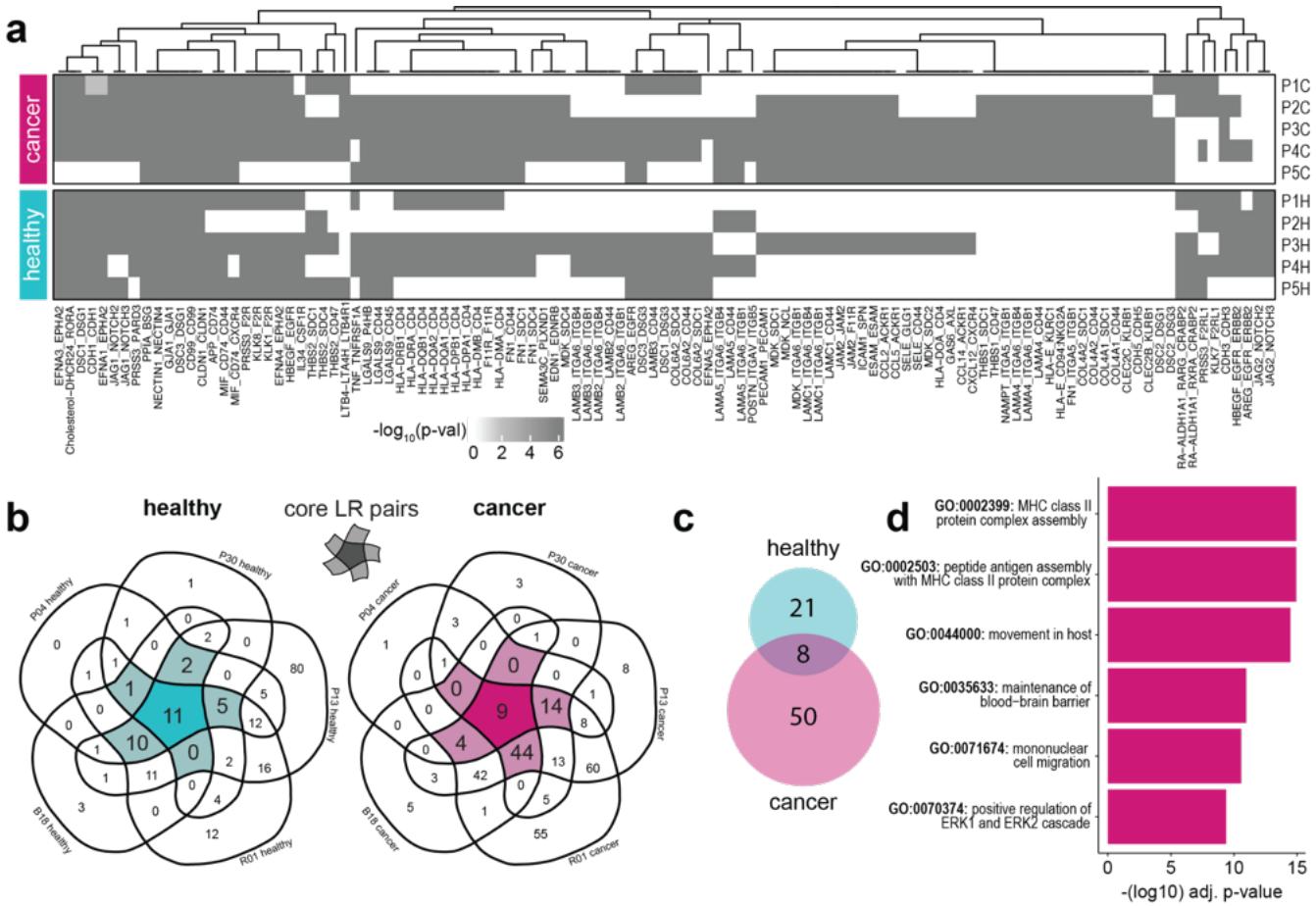


Figure S17. Cell-cell interaction analysis of scRNASeq data for cSCC

(a) Heatmap indicating significant LR pairs across samples. 312 significant pairs were detected from the entire suite of expressed genes in each sample. For visualisation purposes, only those LR pairs predicted in at least 4 of 10 samples are shown, with their corresponding $-\log_{10}(p\text{val})$ values. Non-significant LR pairs in each sample are white. LR communication was calculated between pairs of Level 2 cell types; duplicate LR pairs in each sample (i.e. predicted for multiple cell types within one sample) were collapsed, with only the highest $-\log_{10}(p\text{val})$ score shown.

(b) Patient-specific LR predictions in healthy (left) or cancer (right) cells across patients. The core suites of LR pairs that are shared across four or more of the five patients are highlighted in the centre of each Venn diagram.

(c) Overlap between core LR suites in cancer and healthy patients.

(d) The top six non-redundant GO terms that are enriched in the cancer-specific suite of LR genes, ranked by -log₁₀ adjusted P-value. GO analysis was performed against a background gene universe of all LR pairs that are present in the CellChat database and are also expressed in our dataset.

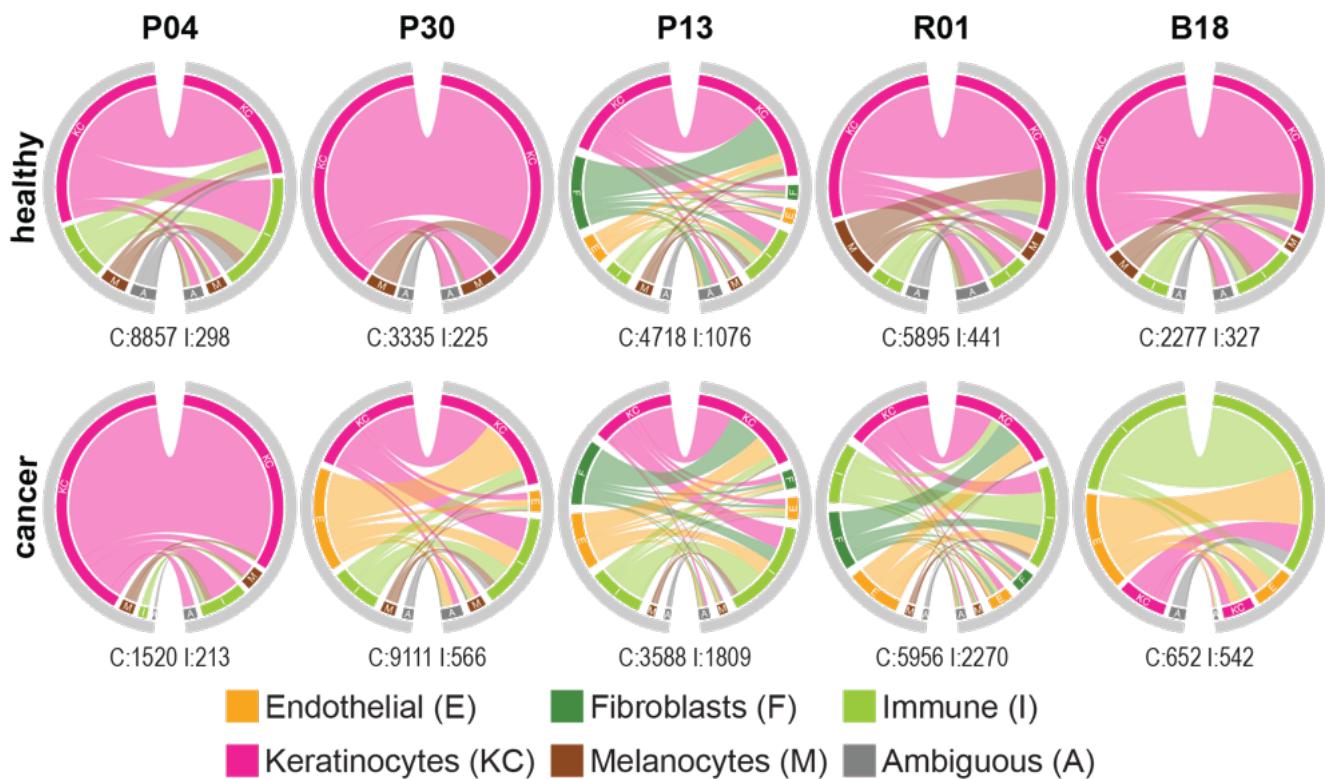
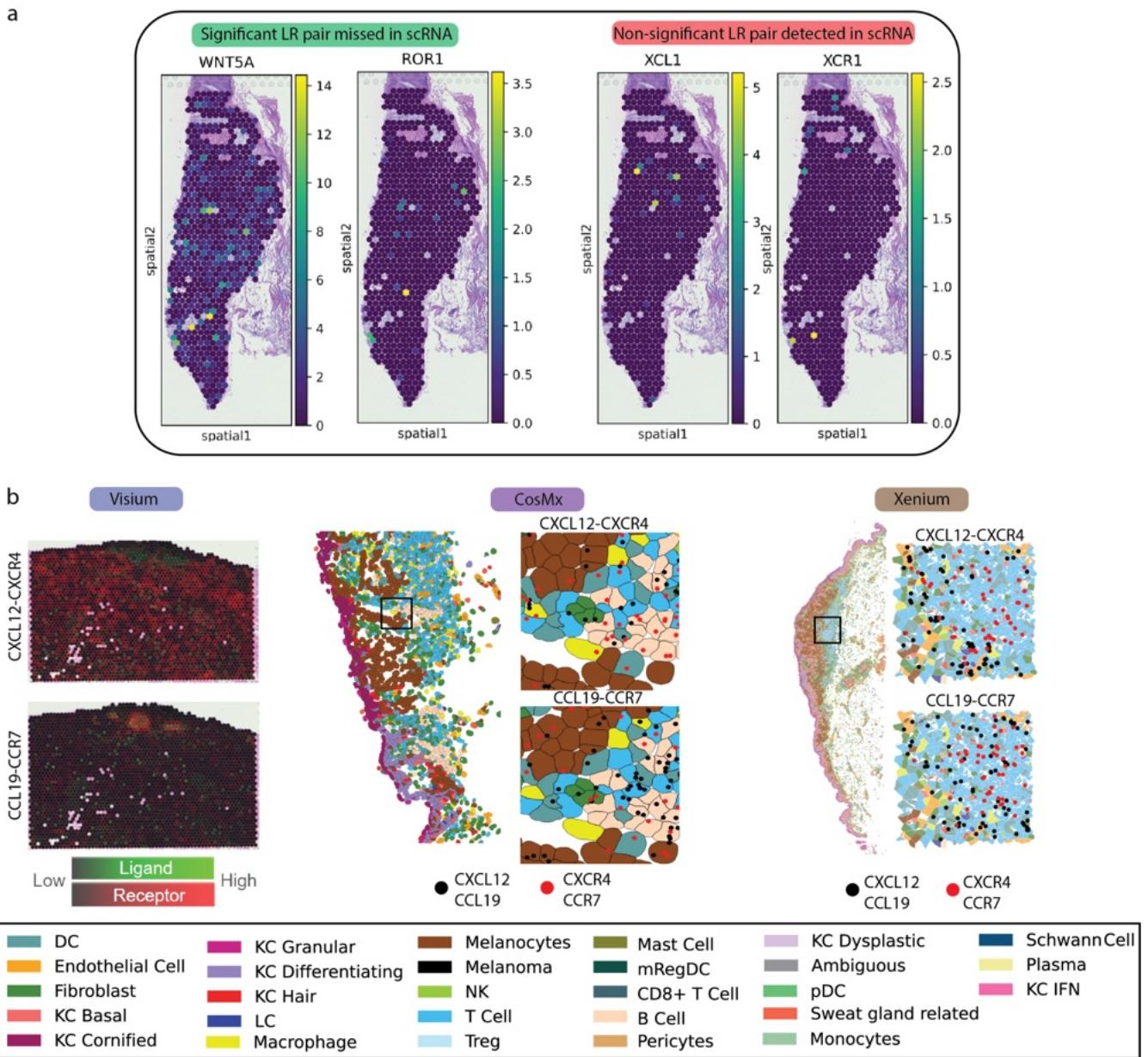


Figure S18. CCI analysis in cSCC samples using scRNASeq

Changes in cell type participation in cellular crosstalk between ligands and receptors in healthy and cancer samples from different patients. Circle diagrams show cells expressing ligands on the left, and receptors on the right. Connecting line colours indicate sender cell identity, and line widths indicate the number of unique LR connections predicted between pairs of cell types. Colours around the edge of the ring indicate sender and receiver cells' identities. Numbers below each plot indicate the number of cells (C) and significant interactions (I) for each sample. Interactions were predicted for pairs of Level 2 cell types but summarised to broader cell type categories for visualisation.



356
357
358
359 **Figure S19. Comparing LR detection between scRNAseq data without and different spatial**
360 **transcriptomics platforms with spatial context.**

361 (a) Examples of cases mis-detected by scRNAseq (WNT5A-ROR1) or detected by scRNAseq but not
362 colocalize in spatial data, suggesting potential false positive.
363 (b) Consistent detection of LR pairs by Visium, CosMX and Xenium data

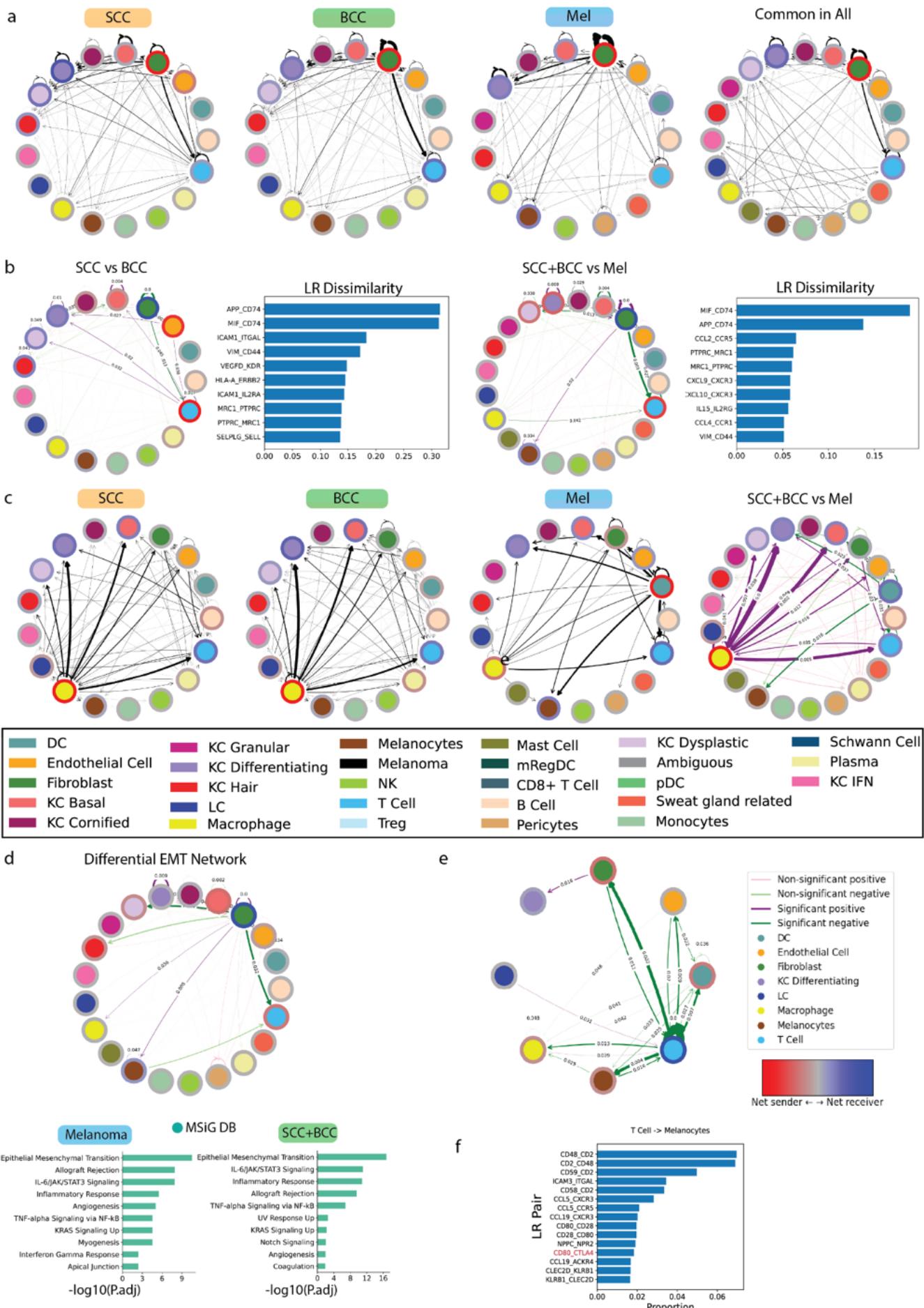
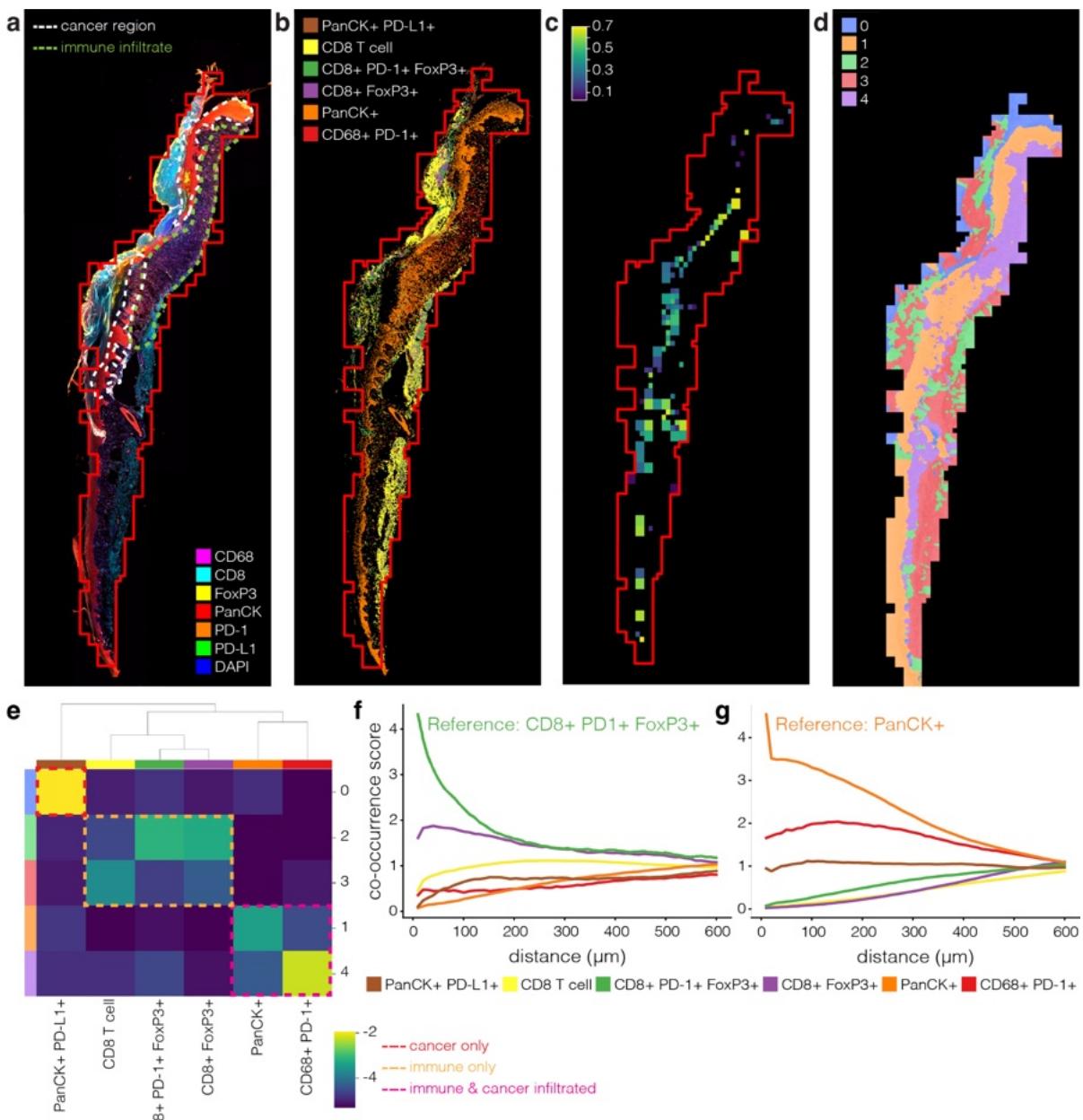


Figure S20. Comparing interactions across the three cancer types

372 (a) MMCCI integrated interactions for each cancer type and across multiple samples and spatial
373 transcriptomics platforms. From left to right, circos interaction plots for BCC, cSCC and Melanoma. The
374 last plot on the right is the integrated interactions for all cancer types for all samples.
375 (b) Comparing differential interactions at cell type level for BCC vs cSCC and for the integrated BCC-
376 cSCC vs Melanoma using all L-R pairs. For cSCC vs BCC comparison, purple arrows show more
377 interactions in BCC and green arrows indicating more in cSCC. For cSCC-BCC vs Melanoma, the purple
378 arrows show more in melanoma and green arrows show more in cSCC-BCC.
379 (c) Differential interactions at cell type level using MIF-CD74 pair as the top different L-R pair shown in
380 panel b. The difference indicates which cell type pairs are more actively interacting by considering the
381 MIF-CD74 pair.
382 (d) Differential CCI network plots for LR pairs in the EMT pathway found from MSigDb pathway analysis.
383 The pathway-centric interaction analysis shows specific cell type pairs where the EMT pathway was
384 most significantly different. Pathway enrichment analysis from MSigDb on upregulated LR pairs common
385 across CosMx and Visium 1) Melanoma and 2) cSCC+BCC.
386 (e) Interactions between cell-types in Melanoma compared to cSCC only.
387 (f) Top LR pairs interacting between T cells and melanocytes.

391



392

393

394

Figure S21. Spatial cell type and community mapping at the protein level by targeted multispectral imaging (Polaris)

395

396

397

398

399

400

401

402

403

404

405

406

407

408

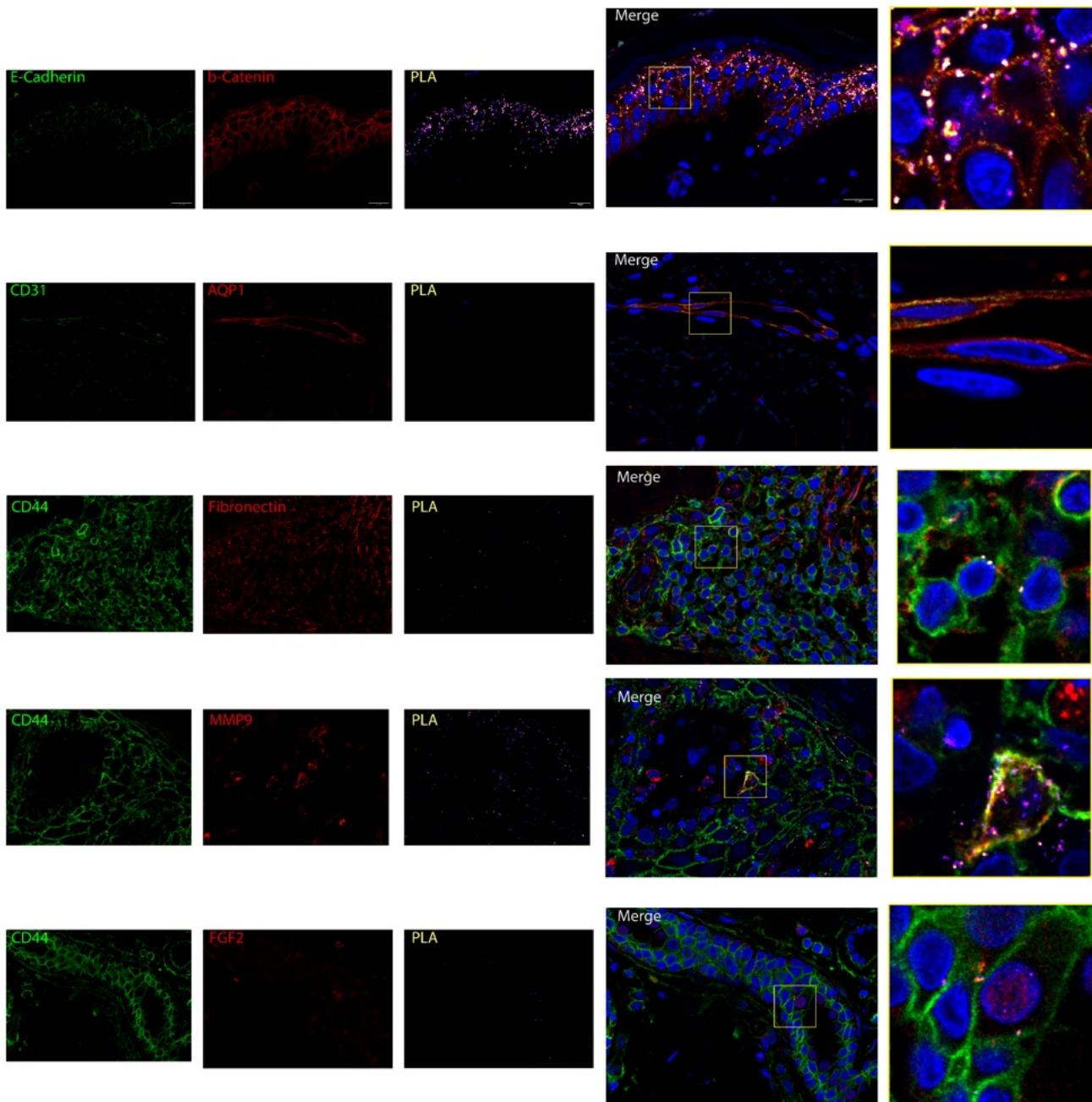
(a) Polaris protein imaging data from cSCC patient B18 with pathological annotation of cancer and immune regions, based on tissue morphology. A panel of six antibodies was used to profile protein markers and major cell types of interest including CD8 (T cell marker), FoxP3 (T-regulatory cell marker), CD68 (monocyte/macrophage marker), PanCK (epithelial cell marker), PD-1 and PD-L1 (immune checkpoint inhibitor, with the former expressed on immune cells and the latter on tumour cells); DAPI was also included as a nuclear stain. White and green dashed lines outline the cancer and immune regions of the tissue, as annotated by a pathologist. The red box indicates the tissue border as determined later during STRISH analysis, shown here to aid comparison with subsequent panels.

(b) Cell type classification based on clustering and gating of single-cell resolution protein expression signals.

(c) Mapping of cancer-immune cell co-localisation using STRISH, identifying image tiles containing double-positive CD8+ PD-1+ immune cells and PanCK+ PD-L1+ cancer cells. The identified colocalisation zone correlates with the cancer infiltrate region identified in Panel a.

(d) Voronoi diagram mapping the spatial distribution of cell communities, defined by further clustering of

409 *cell type classifications.*
 410 (e) *Cell type composition of the five cell communities defined in Panel d, revealing higher-level groups*
 411 *of communities representing cancer only (red box), immune cells only (orange box), and cancer-immune*
 412 *infiltrate (pink box). Heatmap colours represent the neighbourhood enrichment score for each*
 413 *community-cell type combination.*
 414 (f-g) *Cell type co-occurrence measurements of CD8+ PD-1+ FoxP3+ immune cells (f) and PanCK+*
 415 *cancer cells (g) with either the same or other cell types in the dataset. Each line plots the co-occurrence*
 416 *score between the reference cell type and the test cell type calculated over increasing spatial distances.*
 417
 418



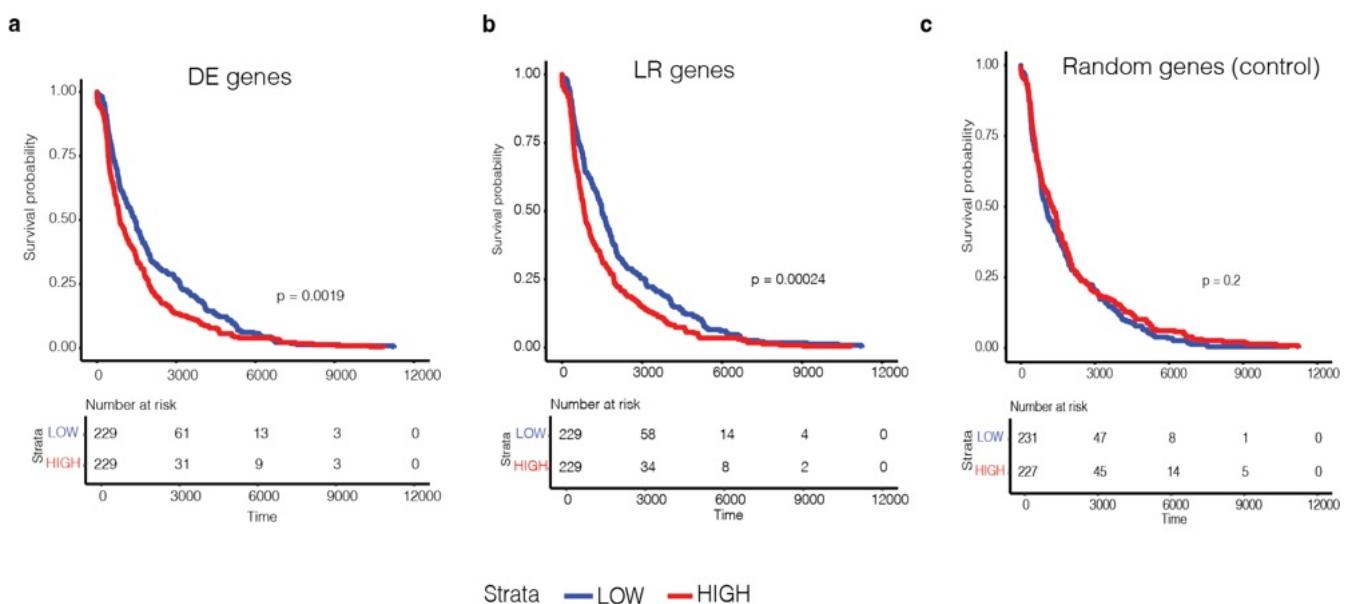
419
 420 **Figure S22. Experimental validation of ligand-receptor interaction using proximal ligation assay**
 421 **for melanoma samples.** In each row, fluorescence signal for a single antibody and PLA is shown,
 422 followed by a merge image and a Zoom-in window highlighting the interaction on the cell membrane. A
 423 positive PLA signal is visible within two proteins in a proximity less than 20 nm. From top row to bottom
 424 row, a positive control E-Cadherin and b-Catenin, a negative control for CD31 and AQP1, and the

425 interaction results for the three pairs CD44-Fibronectin, CD44-MMP9, and CD44-FGF2 ligand-receptor
426 pairs.

427

428

429



430
431 **Figure S23. Prognostic value of melanoma specific genes identified with scRNA-seq and spatial**
432 **transcriptomics.** The plots indicate survival curves of patients from the TCGA SKCM dataset for groups

433 stratified as high and low expressing patients for

434 (a) genes upregulated in malignant melanocytes compared to normal melanocytes identified by DE

435 analysis with edgeR

436 (b) LR genes upregulated in Melanoma compared to SCC and BCC

437 (c) a random set of genes excluding the genes used in a and b. The melanoma specific genes identified

438 in our study are associated with poor prognosis where the high expressing patients indicate significantly

439 poor survival while the random genes show no significant stratification.

440 **List of Supplementary Tables**

441 **Table S1:** Summary of samples used in this study

442

443 **Table S2:** Pre-oligonucleotide-conjugated-antibodies and complementary reporters

444 **Table S3:** Table of Primary antibodies used in PLA assay

445

446
447**Table S1. 24 patient samples to generate the multiomics data resource, suitable for addressing multiple questions.**

Patient#	Cancer type/s	Age	Sex	Lesions	scRNASE q/snRNA seq	Visium			CosMx			Polaris	RNAScope			GeoMx (RNA)	GeoMx (protein)	Codex	Xenium
						cSCC and Mel	cSCC	BCC	Mel	cSCC	BCC	Mel	cSCC	cSCC	BCC	cSCC	cSCC	Mel	Mel
B18	cSCC + BCC [†]	76	F	1x leg BCC 1x face IEC 1x leg IEC	✓*†	✓	✓ [^]		✓	✓		✓	✓	✓ [§]	✓	✓			
E15	cSCC + BCC	54	F	1x leg BCC 2x leg IEC 1x leg cSCC		✓	✓						✓ [§]	✓ [§]					
F21	cSCC + BCC	76	M	1x face BCC 1x leg cSCC		✓	✓						✓ [§]	✓ [§]					
P30	cSCC	65	M	2x arm IEC	✓*	✓			✓										
P13	cSCC	75	M	Leg	✓*	✓			✓										

P04	cSCC	72	M	Arm	✓*							✓			✓	✓			
R01	cSCC	73	M	Neck	✓*							✓			✓	✓			
D12	BCC	82	M	Face							✓								
6747-085P	Mel	89	M	Chest				✓				✓							
21031-08TB	Mel	89	M	Face				✓				✓							
48974-2B	Mel	79	F	Back				✓				✓							
66487-1A	Mel	53	F	Arm				✓				✓							
6475-07FC	Mel	53	F	Abdomen													✓	✓	
9474-06BR	Mel	55	M	Back													✓	✓	
23346-10SP	Mel	70	M	Back													✓	✓	
53023-07BR	Mel	81	M	Face														✓	
98594-09PY	Mel	65	F	Back														✓	
30037-07BR	Mel	76	M	Scalp														✓	

health y skin 1	Health y ctrl	27	M	Forear m		✓														✓
health y skin 2	Health y ctrl	41	M	Forear m		✓														✓
health y skin 3	Health y ctrl	45	M	Forear m		✓														✓
MPS1 3	Mel malign antl	28	F	Right upper arm	✓															
MPS4 2	Mel interm ediate	35	M	Back	✓															
MPS4 3	Mel benign	37	F	Left upper back	✓															
# B18 = P5, P30 = P2, P13 = P3, P04 = P1, R01 = P4 in Supplementary Tables																				
* Healthy and cancer biopsies																				
† Cancer biopsy from patient B18 contained pooled cSCC and BCC cells																				
^ 3x biological replicates																				
§ Tran et. al. 2022																				

448

449

450

Table S2. Pre-oligonucleotide-conjugated-antibodies and complementary reporters

Antibody	Clone no.	Reporter
CD4	AKYP0048	RX003-Cy5
HLA-A	AKYP0078	RX004-AF750
Podoplanin	AKYP0007	RX023-ATTO550
CD68	AKYP0050	RX015-Cy5
CD20	AKYP0049	RX007-AF750
CD8	AKYP0049	RX026-ATTO550
CD11c	118/A5	RX024-Cy5
IDO1	AKYP0084	RX027-AF750
CD31	AKYP0047	RX001-AF750
GranzymeB	AKYP0086	RX041-ATTO550
CD3e	AKYP0062	RX045-Cy5
Vimentin	AKYP0082	RX022-AF750
LAG3	D2G4O	RX055-ATTO550
PD-L1	AKYP0097	RX043-AF647
CD34	AKYP0080	RX025-AF750
Foxp3	AKYP0101	RX031-AF647
SMA	AKYP0081	RX013-AF750
CD19	RM332	RX028-ATTO550
CD45	AKYP0074	RX021-AF750
MPO	AKYP0113	RX098-ATTO550
PD1	AKYP0070	RX046-AF647
CD21	AKYP0061	RX032-ATTO550
CD11b	RM290	RX030-Cy5
S100B	E7C3A	RX020-ATTO550
HLA-DR	AKYP0063	RX033-AF750

PMEL	HMB-45	RX085-ATTO550
CD163	AKYP0114	RX016-AF647
S100A4	S100A4	RX064-AF750
Keratin 14	AKYP0064	RX002-ATTO550
HIF1A	AKYP0100	RX062-AF647
Ki67	AKYP0052	RX047-ATTO550
SOX10	SP275	RX006-Cy5
Pan-Cytokeratin	AKYP0053	RX019-AF750

452

453

Table S3. Table of Primary antibodies used in PLA assay

Antibody	Clone	Source	Target	Supplier	t no
FGF-2	147	Rabbit	Fibroblast growth factor-1	tegy)	sc-25287
Fibronectin	Polyclonal	Rabbit	Fibronectin	Abcam	ab2413
MMP-9	54	Rabbit	el B)	Abcam	ab76003
CD44	156-3C11	Mouse	Type I transmembrane glycoprotein	Cell Signalling (NEB)	3570
CD31	JC/70A	Mouse	PECAM-1	Dako	MO823
β-Catenin	E-247	Rabbit	Beta Catenin	Abcam	ab32572

E-Cadherin	NCH-38	Mouse	E-Cadherin	Dako	M3612
AQP1	H-55	Rabbit	Aquaporin	Bio-Strategy	SC-20810

454

455 The Navenibody M1 and R2 (secondary antibodies conjugated with PLA probes) are part of the NaveniFlex Tissue MR Atto647N (Navinci,
 456 Swiden, Cat# 60026).

457 Supplementary Methods

458
459 **Integrating 12 Spatial Technologies to Characterise Tumour**
460 **Neighbourhoods and Cellular Interactions in three Skin Cancer Types**
461
462 P. Prakrithi^{1,2#}, Laura F. Grice^{1,3#}, Feng Zhang^{1,2#}, Levi Hockey^{1,2}, Samuel X. Tan⁴, Xiao Tan^{1,2}, Zherui
463 Xiong^{1,2}, Onkar Mulay^{1,2}, Andrew Causer^{1,2}, Andrew Newman^{1,2}, Duy Pham¹, Guiyan Ni¹, Kelvin Tuong⁵,
464 Xinnan Jin^{1,2}, Eunju Kim^{3,4}, Minh Tran¹, Hani Vu^{1,2}, Nicholas M. Muller⁴, Emily E. Killingbeck⁶, Mark T.
465 Gregory⁶, Siok Min Teoh¹, Tuan Vo¹, Min Zhang⁷, Zachary Reitz⁶ Katharina Devitt⁸, Liuliu Pan⁶, Arutha
466 Kulasinghe⁵, Yung-Ching Kao⁴, Michael Leon⁶, Sarah R. Murphy⁶, Hiromi Sato⁶, Jazmina Gonzalez
467 Cruz⁸, Snehlata Kumari⁸, Hung N. Luu⁹, Sarah E. Warren⁶, Chris McMillan^{10,11}, Joakim Henricson^{12,13},
468 Chris Anderson^{12,13}, David Muller^{10,11}, Arun Everest-Dass¹⁴, Blake Obrien¹⁵, H. Peter Soyer⁴, Ian Frazer⁷,
469 Youngmi Kim⁶, Mitchell S. Stark⁴, Kiarash Khosrotehrani⁴, Quan Nguyen^{1,2}

Supplementary Methods

470 **1. Single-cell RNA sequencing, data pre-processing and annotation**
471

472 **Sample preparation**

473 10x Genomics Chromium single-cell library preparation and sequencing was performed according to
474 the manufacturer's instructions, using the Single Cell 3' Library, Gel Bead and Multiplex Kit (version 2,
475 PN-120233; 10x Genomics). Cell numbers were optimised to capture approximately 3,000 cells per
476 reaction. Single-cell transcriptome libraries were sequenced on an Illumina NextSeq500 machine, using
477 a 150-cycle High Output reagent kit (NextSeq500/550 version 2, FC-404-2002; Illumina) with the
478 following reads and indices: read 2 (98 bp), i7 index (8 bp), read 1 (cell barcodes and UMI; 26 bp).

479 **scRNAseq data processing (cSCC-BCC)**

480 The BCL sequencing output file was converted to a FASTQ file using bcl2fastq/2.17. Read mapping was
481 performed using CellRanger/3.0.2 against the *Homo sapiens* GRCh38p10 reference genome. Median
482 absolute deviation (MAD) filtering was performed using scater v1.14.6 to remove outlier cells with low
483 library size and/or gene counts (>3 MADs below the median value), or high mitochondrial or ribosomal
484 gene percentage (>3 MADs above the median value). Doublet scoring was performed using scds v1.2.0
485 and cells were removed if they were predicted as doublets by at least two of the three included prediction
486 methods (bcds, cxds, or hybrid), and expressed more than 3,000 genes. After QC filtering, the 48,226
487 remaining cells in P30 were randomly subsampled to 3,335 cells due to its much larger depth compared
488 to other samples. This number of cells was selected as the other four healthy samples contained an
489 average of 3,335.25 cells each. Samples were processed individually using Seurat version 5.0 before
490 being integrated using Seurat canonical correlation analysis (CCA). Following integration,
491 dimensionality reduction (using the first 20 principal components), clustering and sub-clustering, and
492 marker prediction were all performed in Seurat v5.0 (ref). Different cluster resolutions were tested and
493 assessed for cluster stability using Clustree v0.4.2.

494 **Cell type annotation BCC-cSCC**

495 We performed cell type annotation at three granular levels, which we call "Level 1", "Level 2" and "Level
496 3" annotation. For the broader Level 1 annotation, the integrated set of 11 samples was clustered using
497 Seurat (resolution 0.4, 20 principal components (PCs))¹. Seurat's FindAllMarkers function was used to
498 find the top differentially expressed genes for each cluster. Cluster annotation was performed based on
499 manual comparison of the top gene markers with the literature. To determine more detailed annotations,
500 each major cell type (endothelial cells, fibroblasts, immune cells, keratinocytes and melanocytes) was
501 extracted and re-clustered in turn for our Level 2 classification. Optimal cluster resolutions were selected
502 using clustree. Cell subtypes were identified by comparing marker genes, detected using Seurat's
503 FindAllMarkers function, with the literature.

504 **Annotation of aneuploid cells using inferred CNV profiles**

Supplementary Methods

505 We applied two top-performing tools, CopyKAT v1.1.0 and InferCNV v1.19.1, to add complementary
506 information to prioritise malignant cells. In both cases, the raw count matrices were used as the input,
507 using the default parameters and the reference free option (no normal cell IDs were provided). CopyKAT
508 predicted ploidy status for each cell barcode, including Aneuploid, Diploid and Not Defined (for uncertain
509 cells, often associated with low quality data). To define a set of ‘confident aneuploid’ cells, we integrated
510 results from CopyKat and InferCNV. Although InferCNV does not predict aneuploid cells and is mainly
511 used to infer CNV profiles to further stratify the known malignant cells into subclones, we made use of
512 its quality filtering analysis results. InferCNV excludes cells of low quality and the ones with CNV profiles
513 that are likely to be normal and retains only the cells likely to be aneuploid. Finally, we called the
514 ‘Aneuploid’ cells that passed the selection criteria of both tools. The inferred CNV results were then
515 combined with DE gene module score analysis to finally classify keratinocytes or melanocytes as cancer
516 cells.

517 Single Nucleus RNA Sequencing (snRNAseq) of melanoma samples

518 snRNAseq was performed using nuclei dissociated from two 25-µm FFPE melanoma tissue sections.
519 Nuclei dissociation was carried out using the gentleMACS Octo Dissociator (Miltenyi Biotec) with the
520 default program 37C_FFPE_1. The dissociated nuclei were purified by flow cytometry based on
521 Ethidium Homodimer-1 staining to exclude keratin debris. Library preparation was performed following
522 the Chromium Fixed RNA Profiling user guide (10x Genomics). Single-nucleus emulsions (GEMs) were
523 generated using the Chromium X system (10x Genomics). Libraries were constructed from the captured
524 nuclei by probe ligation, probe extension, amplification, and indexing. Library was sequenced on a
525 NextSeq 2000 platform (Illumina) with a dual-indexed sequencing run following read configuration:
526 Read1, 28 cycles; i7 index, 10 cycles; i5 index, 10 cycles; and Read2, 90 cycles.

527 snRNAseq data processing (melanoma samples)

528 The reads were mapped to the reference Human Transcriptome Probeset V1.0.1 GRCh38 (2020-A)
529 using CellRanger (V7.0), implementing the cellranger multi pipeline. The filtered cell count matrix from
530 CellRanger output was used for downstream pre-processing, integration and clustering analysis.
531 Doublets and cells with more than 7000 genes or over 20,000 reads were removed. Harmony was used
532 to integrate the datasets from the three patients (n_neighbors=30, n_pcs=75). Leiden clustering was then
533 applied with the default resolution 1, resulting in 33 clusters. Similar to the annotation pipeline applied
534 for the cSCC-BCC data, the melanoma clusters were annotated at three levels, based on gene markers
535 most differentially expressed from the Wilcoxon test comparing each cluster to all the other clusters. A
536 list of known skin gene markers were also used to check for specificity of marker genes to corresponding
537 clusters. All melanocytes defined from the level 1 clustering were further subclustered, forming seven
538 melanocyte subclusters. The immune cells were also subclustered into eight level 2 clusters, and a total
539 of 19 clusters at the level 3 clustering.

Supplementary Methods

To define melanoma, only melanocytes from the patient with definite invasive melanoma (MPS13) were used to find highly-likely cancer cells. Two selection criteria include that the melanocytes had to have a high “melanoma score” and that the cell had an aneuploid genotype as predicted by inferCNV and CopyCAT. To compute a module score we selected a list of genes upregulated in the melanoma sample compared to the benign sample using both edgeR pseudobulking and scanpy non-parametric test. A density plot was produced to inform the cutoff selection. As the majority of the cells were with a score >80th percentile cut-off. To increase confri inferred ‘Aneuploid’ by the CNV analysis and with a high module score by both the aforementioned methods are labelled as malignant melanocytes (red) as shown in the UMAP.

scRNASeq cancer vs healthy comparison

Global transcriptional differences between healthy and cancer cells, either across the whole dataset or for each patient individually, were determined using Seurat’s FindAllMarkers function (with parameters only.pos = TRUE, min.pct = 0.25, logfc.threshold = 0.25 and with adjusted p-value threshold ≤ 0.05). Comparisons between patients were performed using the Venn diagram tool (<http://bioinformatics.psb.ugent.be/webtools/Venn/>). Differentially expressed genes detected in ≥ 4 patients for either cancer or healthy samples were defined as “core gene suites”. Each gene suite was tested for GO enrichment using Cluster Profiler⁸¹. Specifically, the ontology enrichment was performed using the enrichGO function (parameters OrgDb = "org.Hs.eg.db", ont = "BP", readable = TRUE, pvalueCutoff = 0.01, pAdjustMethod = "BH" against a gene universe of all genes that were expressed in our dataset), results were filtered using gsfilter (max = 200) and simplified using the simplify function (cutoff = 0.5). Where multiple GO terms matched identical sets of genes of interest, they were filtered to keep the GO term with the lowest adjusted p-value. The top six remaining enriched GO terms were categorised (**Fig 2d**) according to their relevant ancestor GO term that sat one level below Biological Process (GO:0008150) in QuickGO (<https://www.ebi.ac.uk/QuickGO/>); where two ancestor terms were possible, the most specific term was selected manually. Scaled expression of the 39 core cancer genes was visualised in a heatmap using Complex Heatmap⁸². Differential cell type abundance was calculated using a modified EdgeR pipeline⁸³, with cell type counts for each patient used as input and with design matrix “~patient + disease”.

2. Visium spatial transcriptomics and data pre-processing

Sample preparation

Visium spatial transcriptomics data was generated for twelve samples, five cSCC (patients B18, E15, F21, P30 and P13), three BCC (B18, E15 and F21, i.e. cSCC lesions were also taken from all three patients), and four melanoma (6747-085P, 21031-08TB, 48974-2B and 66487-1A). FFPE samples (cSCC P30 and P13, BCC B18, and Melanoma 66487, 48974, 21031, and 6747) and fresh-frozen samples (cSCC E15, F21 and B18, and BCC E15 and F21) were used for data generation.

Supplementary Methods

575 For fresh-frozen samples, tissue optimisation was performed as per the Visium Spatial Tissue
576 Optimization User Guide (CG000238 Rev A, 10x Genomics), and a tissue permeabilisation time of 18-
577 25 min was selected. For data generation, tissue was cryosectioned at 10 µm thickness and transferred
578 to chilled Visium Spatial Gene Expression Slides (2000233, 10x Genomics, USA), and allowed to adhere
579 by warming the back of the slide. Tissue sections were dried for 1 min at 37°C, fixed in chilled 100%
580 methanol for 30 min, and stained with hematoxylin for 5 min and then eosin for 2 minutes, as per the
581 Visium Spatial Gene Expression User Guide (CG000239 Rev A, 10x Genomics). Brightfield histology
582 images were captured using a 10x objective on an Axio Z1 slide scanner (Zeiss). Brightfield images
583 were exported as high-resolution tiff files using Zen software. This H&E staining and imaging protocol
584 was used to stain all skin sections for histopathological annotation in this study. Libraries were
585 sequenced on a NextSeq500 (Illumina) machine using a High Output 150 cycle kit (Illumina) at the
586 University of Queensland Sequencing Facility, with the read structure as: Read1 - 28bp, Index1 - 10bp,
587 Index2 - 10bp, Read2 - 120bp. The sequencing data was converted to FASTQ format as described
588 above for scRNASeq data and the data was processed as described below. For FFPE samples, the
589 FFPE blocks were sectioned at 5µm by rotary microtome and the sections were processed for spatial
590 sequencing library preparation following the Visium Spatial Gene Expression for FFPE User Guide
591 (CG000407, CG000408, CG000409 - 10x Genomics). The imaging and sequencing were similar to the
592 protocol for the fresh-frozen samples as described above.

593 **Data processing**

594 Reads were trimmed by cutadapt/1.8.3 to remove sequences from poly-A tails and template-switching-
595 oligos. SpaceRanger V1.0 was used to map FASTQ reads to the human reference genome (version
596 GRCh38-3.0.0). After read mapping, spots were removed if they expressed mitochondrial genes more
597 than 20 percent. Thirteen mitochondrial and 103 ribosomal genes were removed from the samples that
598 were originally fresh-frozen; these genes were not captured using the probe-based FFPE library
599 preparation protocol. In total, we captured 14,990 transcriptional capture spots and 36,485 unique genes
600 across all samples. Replicate samples from each of the three cancer types were integrated using Seurat
601 CCI ¹ with 3000 variable genes, and further processed using SCTtransform ¹. Uniform Manifold
602 Approximation and Projection (UMAP) plots were generated based on the first 20 PCs of cSCC and
603 BCC, and the first 30 PCs for melanoma. Clustering was performed on each integrated cancer dataset
604 separately, using resolutions of 0.5 for cSCC, 0.3 for BCC and melanoma.

605 **Cell type annotation using spot deconvolution**

606 We performed Visium spot deconvolution using the Robust Cell Type Decomposition (RCTD) method ⁴,
607 using raw gene counts and with the Level 1 cell type annotations from our single-cell cSCC atlas used
608 as a reference. RCTD was run in doublet mode “full”, and RCTD scores were weighted to a maximum
609 value of 1, to indicate the proportion of each spot predicted to be taken up by a given cell type.
610 Neighbourhood cell type composition was determined by calculating the average RCTD score for each
611 cell type in each cluster.

Supplementary Methods

612 **CCI analysis**

613 Visium CCI analysis was performed for each Visium dataset separately using stLearn⁵. For each of the
614 twelve samples in turn, raw Visium data was input into stLearn, and spots expressing ≥20%
615 mitochondrial reads, and genes expressed in <3 spots were filtered. Data were normalised using
616 stLearn's normalize_total function. CCI analysis was performed at the spot level for each sample
617 separately, using stLearn's st.tl.cci.run function with parameters min_spots = 20 and n_pairs = 10,000.

618 **3. GeoMx data generation and pre-processing**

619 **Sample processing**

620 Formalin-fixed skin samples from three cSCC patients (B18, R01 and P04) were profiled by NanoString
621 GeoMx DSP RNA and protein assays. For each assay, 24 regions of interest (ROIs) of variable size
622 (approximately 300-1,500 cells/ROI) were selected based on tissue morphology from adjacent H&E
623 images, representing 12 pairs of adjacent cancer and immune tissue regions (two pairs from patient
624 P04, six from B18 and four from R01) (**Fig S3**).

625 Sample preparation for both assays followed standard procedure. Briefly, freshly-cut FFPE sections of
626 5 µm thickness were placed onto a glass slide. After baking for 1 h at 60°C, slides were processed on
627 a Leica automation platform, with three major processing steps: 1) slide baking and dewax, 2) antigen
628 retrieval for 20 min at 100°C, and 3) 1ug/ml Proteinase K treatment for 15 min. Slides were then removed
629 from the instrument and were incubated with the relevant probe cocktail (i.e. the GeoMx Cancer
630 Transcriptome Atlas assay (CTA) assay probe cocktail for the RNA assay, or the GeoMx protein
631 immuno-oncology (IO) cocktail incubation for the protein assay) overnight. The following day, slides
632 were washed and applied with morphology marker incubation before loading onto the GeoMx machine
633 for processing.

634 For the RNA assay, cell segmentation markers used were CD3E (red), PTPRC (yellow), KTR5 (green)
635 and DAPI (blue). For gene expression quantification, we used the CTA probe panel, of which the majority
636 of genes are related to cancer biology (i.e. global immune response, microenvironment immune
637 activities, tumour reactivity to treatment, tumour inflammation signature and cancer metastasis), but also
638 includes 31 endogenous control housekeeping genes and one negative control spike-in. We captured
639 RNA quantification information for 1,825 genes using this panel.

640 For the protein assay, equivalent cell segmentation morphology markers were used as in the RNA
641 assay, namely CD3 (red), CD45 (yellow), PanCk (green), DAPI (blue). The IO protein panel includes a
642 total of 48 protein markers, including three positive controls (Histone H3, GAPDH, and S6) and three
643 negative controls (Rb IgG, Ms IgG1, and Ms IgG2a).

644 **Differential expression analysis (protein and RNA data)**

Supplementary Methods

Differential gene expression analysis was performed for RNA and protein data separately using the edgeR pipeline ⁷. Analysis was performed using log Counts Per Million (logCPM)-transformed expression values. Prior to analysis of RNA data, five lowly-expressed genes were filtered using the function *filterByExpr*, leaving 1,820 markers for downstream analysis. No proteins were filtered prior to analysis, leaving 48 protein markers for downstream analysis. Library size normalisation was performed using the calcNormFactors function. Differential expression analysis between CD45+ and PanCK+ regions was performed using the cpm.DGEList function with the design matrix “~patient + group” to compare between-group differences while accounting for between-patient differences. We used the *glmQLFit* quasi-likelihood negative binomial model to fit a linear model to the expression data. Genes/proteins were categorised as differentially expressed if they received a false detection rate (FDR) ≤ 0.05 and \log_2 fold change >1.0 . GO enrichment analysis of differentially expressed genes from the RNA assay was performed as described above for scRNASeq comparisons. GO analysis was not performed for protein data due to the small number of captured markers available to serve as the background universe for statistical testing.

659 **Spatial Deconvolution (RNA data)**

We used the R package SpatialDecon⁸ to assess the abundance of different cell types across 12 CD45+ ROIs. SpatialDecon estimates cell type abundance within spatially-resolved gene expression datasets. SpatialDecon takes as input the normalised gene expression data (here, the logCPM-normalised GeoMx data). It also uses the function derive_GeoMx_background to calculate expected background counts for the data using the normalised gene expression matrix. Cell types were assigned based on an inbuilt reference expression profile of expected cell types, the SafeTME matrix included with the SpatialDecon package. This reference contains the expression profiles of 906 genes over 18 cell types, and was purposely designed to exclude genes commonly expressed by cancer cells ⁸. For this reason, we did not perform deconvolution for PanCK+ segments as the reference profiles available for GeoMx data are biased towards immune cells, and specifically exclude the cancer-associated genes expected to be enriched in the PanCK+ fraction.

671 **4. CosMx data collection and pre-processing**

672 **Sample preparation**

Nanostring CosMx data was generated for nine samples, three cSCC (patients B18, P30 and P13), two BCC (patients B18 and D12), and four melanoma (patients 6747-085P, 21031-08TB, 48974-2B, 66487-1A). FFPE tissue sections were prepared as previously described ⁹. Briefly, 5 μ m tissue sections on VWR Superfrost Plus Micro slides (Cat. No. 48311-703) were baked overnight at 60°C to improve tissue adherence to the slides, then prepared for *in situ* hybridisation (ISH) by deparaffinisation and heat-induced epitope retrieval (HIER) at 100°C for 15 min. For all samples except cSCC Patient P30, HIER

Supplementary Methods

679 was performed in a pressure cooker using ER2 epitope retrieval buffer (Leica Biosystems product,
680 EDTA-based, pH 9.0). HIER for Patient P30 was performed on a Leica Biosystems Bond RX automated
681 tissue handler using ER1 epitope retrieval buffer (Leica Biosystems product, citrate-based, pH 6.0).

682 Following HIER, tissue sections were digested with Proteinase K diluted in ACD Protease Plus at 40°C
683 for 30 min (3 µg/ml Proteinase K for all samples except Patient P30, which was at 5 µg/ml). The tissue
684 sections were washed twice with diethyl pyrocarbonate (DEPC)-treated water (DEPC H₂O) and
685 incubated in 1:3,000 diluted† fiducials (1:5,000 for Patient P30) (Bangs Laboratory) in 2X SSCT (2X
686 saline sodium citrate, 0.001% Tween-20) solution for 5 min at room temperature in the dark. Excess
687 fiducials were removed by rinsing the slides in 1X PBS, before tissue sections were fixed with 10%
688 neutral buffered formalin (NBF) for 5 min at room temperature. Fixed samples were rinsed serially with
689 Tris-glycine buffer (0.1M glycine, 0.1M Tris-base in DEPC H₂O) and 1X PBS for 5 min each before
690 blocking with 100 mM N-succinimidyl (acetylthio) acetate (NHS-acetate, ThermoFisher) in NHS-acetate
691 buffer (0.1M NaP, 0.1% Tween PH 8 in DEPC H₂O) for 15 min at room temperature. The sections were
692 then rinsed with 2X saline sodium citrate (SSC) for 5 min and an Adhesive SecureSeal Hybridization
693 Chamber (Grace Bio-Labs) was placed over the tissue.

694 NanoString ISH probes were prepared by incubation at 95°C for 2 min and immediately placed on ice,
695 and the ISH probe mix (1 nM ISH probe, 1X Buffer R, 0.1 U/µL SUPERase•In [Thermofisher] in DEPC
696 H₂O) was pipetted into the hybridisation chamber. The chamber was sealed to prevent evaporation, and
697 hybridisation was performed at 37°C overnight. The tissue sections were then washed twice in 50%
698 formamide (VWR) in 2X SSC at 37°C for 25 min, washed twice with 2X SSC for 2 min at room
699 temperature, and blocked with 100 mM NHS-acetate in the dark for 15 min. In preparation for loading
700 onto the CosMx SMI instrument, a custom-made flow cell was attached to the slide.

701 **CosMx SMI instrument run**

702 RNA target readout on the CosMx SMI instrument was performed following the published protocol ⁹.
703 Briefly, the assembled flow cell was loaded onto the instrument and washed with Reporter Wash Buffer
704 to remove residual air bubbles. A preview scan of the entire flow cell was taken, and 15-25 FOVs were
705 selected to match areas of interest identified by H&E staining of adjacent serial sections. For some
706 tissues we selected closely-adjacent FOVs to capture almost the entire tissue biopsy, while for other
707 tissues we selected FOVs covering a smaller fraction of the biopsy. RNA readout began by flowing 100
708 µl of Reporter Pool 1 into the flow cell and incubation for 15 min. Reporter Wash Buffer (1 mL) was
709 flowed into the flow cell to wash away unbound reporter probes, and Imaging Buffer was added to the
710 flow cell for imaging. Nine Z-stack images (0.8 µm step size) were acquired for each FOV, and
711 photocleavable linkers on the fluorophores of the reporter probes were released by UV illumination and
712 washed away with Strip Wash buffer. The fluidic and imaging procedure was repeated for the 16 reporter

Supplementary Methods

713 pools, and the 16 rounds of reporter hybridisation-imaging were repeated multiple times to increase
714 RNA detection sensitivity.

715 After RNA readout, tissue samples were incubated with a 4-fluorophore-conjugated antibody cocktail
716 against CD298, S100b/PMEL17 (PanCK for Patient P30), CD45, and CD3 proteins and DAPI stain in
717 the CosMx SMI instrument for 1 h. After unbound antibodies and the DAPI stain were rinsed with
718 Reporter Wash Buffer and Imaging Buffer was added to the flow cell, nine Z-stack images for five
719 channels (four antibodies and DAPI) were captured.

720 ***Primary data processing***

721 CosMx data was processed as described previously⁹. Briefly, the image processing comprises three
722 main steps: registration, features detection, and localisation. 3D rigid image registration was performed
723 with the use of fiducial markers embedded within the tissue sample. Each subsequent image stack was
724 registered using phase correlation and individual channels were aligned through a pre-calibrated affine
725 transformation.

726 Diffraction-limited features were identified that represented the fluorescence response from a single
727 molecule. A 2D Laplacian of Gaussian (LoG) filter was applied to remove background and enhance the
728 encoded reporter signatures. Potential reporter locations were identified as local maxima using a 3D
729 nearest neighbour search. Sub-pixel localisation for each feature was obtained by fitting a 2D polynomial
730 to the maxima in the X, Y, and Z axes.

731 ***Secondary analysis and decoding***

732 The XYZ locations of all individual reporter-binding events were used to determine the presence of
733 individual transcripts. Briefly, each unique location with at least one reporter-binding event is considered
734 a ‘seed’, and all neighbouring locations to each seed with at least one reporter-binding event were
735 determined. All possible four reporter combinations of unique reporter probes in a seed’s
736 neighbourhood, such that at least one of the four reporter-binding events was present at the seed
737 location, were then matched with gene-specific barcodes to detect the presence of a gene in a seed’s
738 neighbourhood.

739 ***Cell segmentation algorithm***

740 NanoString’s cell segmentation pipeline is a combination of image preprocessing and machine learning
741 techniques⁹. Briefly, the pipeline takes tissue images stained with both nuclear and membrane markers
742 to perform rescaling, normalisation, image deconvolution, and boundary enhancement. Image
743 subtraction was performed between the nuclear and the membrane channels to enhance the contrast
744 between adjacent cells while reducing auto-fluorescence signal. The preprocessed images were fed into
745 pre-trained Cellpose neural network models¹⁰ for both nuclear and cytoplasm modes of segmentation.

Supplementary Methods

746 Results from two segmentation tasks were combined to select the best results from each mode by
747 analysing intersection and union between all segmented cells.

748 Single-cell quality control and filtering

749 Single-cell expression profiles were derived by counting the transcripts of each gene that fell within the
750 area assigned to a cell by the segmentation algorithm. Cells with fewer than 20 total transcripts or that
751 have abnormally large cell area, which likely represent doublets, were omitted from the analysis. A
752 normalised expression profile was defined for each cell by dividing its raw counts vector by its total
753 counts. A separate UMAP projection was computed for each cancer type.

754 Cell type annotation

755 We used our scRNASeq data to construct a reference profile matrix of cell types for CosMx annotation.
756 This profile listed the average expression of each of the 960 genes in the CosMx panel for each of the
757 cell types in our scRNASeq Level 1 cSCC annotations. This reference matrix was then used to initialise
758 an Expectation-Maximisation (EM) algorithm that alternated between assigning cell types and updating
759 the cluster profiles. The EM algorithm was built around the following likelihood model:

760
$$Y_{ij} \sim NegBinom(mu = b_j + s_{j,k(j)}x_{ik(j)}, size = 10)$$

761 Where i indexes genes, j indexes cells, k indexes cell types, Y_{ij} is the raw counts of gene i in cell j , b_j
762 is the expected per-gene background counts in cell j , $X_{ik(j)}$ is the expected expression level of gene i
763 in cell type $k(j)$, and s_j is a scaling factor corresponding to the total counts in cell j . b , X and s were pre-
764 defined where b is calculated from a direct measurement of the background observation rate using
765 included CosMx control target observations, X represents the cell type expression profile measurements
766 derived from scRNASeq data, and s_j is calculated as a cell's total counts divided by the total counts in
767 the profile under consideration.

768

769 In the M-step, cluster assignments $k(j)$ are updated by assigning each cell j to the cluster k under which
770 it achieves highest likelihood. In the E-step, each column k of the matrix X is estimated by taking the
771 maximum likelihood estimate calculated by applying the above likelihood model to all cells assigned to
772 cell type k . To preserve the influence of the original reference profiles, the cell type expression profiles
773 estimated in the E-step are averaged with the original reference profiles before proceeding to the M-
774 step. The E-step and M-step are iterated until convergence.

775

776 5. Cell type co-localisation and CCI with Polaris spatial proteomics

777 Sample processing

Supplementary Methods

Polaris data was generated for three cSCC samples (patients B18, P04 and R01). Multispectral analysis of FFPE tissue utilised the MOTiF PD-1/PD-L1 kit (Akoya Biosciences, Cat. No. OP-000001). Staining with Leica BOND RX (Leica Biosystems) and imaging with Vectra Polaris (Akoya Biosciences) was performed at The Walter and Eliza Hall Institute (WEHI) Histology core facility as per kit manufacturer instructions. Briefly, tissue was stained through cycles of incubation with primary antibody, anti-IgG polymer HRP and covalent labelling with Opal TSA fluorophores, followed by HIER to remove bound antibodies prior to subsequent antibody cycles. Target markers included CD8 (Opal 480), PD-L1 (Opal 520), PD-1 (Opal 620), FoxP3 (Opal 570), CD68 (Opal 780, PanCK (Opal 690) and spectral DAPI DNA stain. Whole slide multispectral scanning was performed on the Vectra Polaris instrument using automatically adjusted exposure settings. Image tiles were spectrally unmixed in InForm (Akoya Biosciences), then restitched in QuPath software¹¹.

789

790 **Data processing**

791 We applied StarDist cell detection ¹² to segment cell boundaries and measure protein marker signals
792 within these boundaries, transforming the imaging data into non-imaging, single-cell protein expression
793 counts (i.e. the mean signal intensity for each protein). To remove artificially high background from
794 fluorescent intensities (outliers), the maximum value of each marker was capped at the 95th percentile.
795 Leiden cell type clustering was implemented in Scanpy ¹³ using a standard single-cell analysis pipeline.
796 Resulting clusters were assessed to determine cell type identity. Cells showing very low expression of
797 all six protein markers were classified as “unidentified” and removed from downstream analysis.

798

799 Using the cell clustering information, we sought to identify the subregions where cancer and immune
800 cells co-localised and possibly interacted. Spatial information at the subcellular level can facilitate a
801 lower level of spatial analysis, which uses the cell geometry and density to assist the interaction analysis.
802 To unbiasedly detect zones of cancer and immune cell co-localisation, and potential interaction, we
803 applied a scanning window strategy called STRISH ¹⁴ to identify regions containing double-positive
804 CD8+ PD-1+ immune cells in the immediate vicinity of double-positive PanCK+ PD-L1+ cancerous
805 epithelial cells. STRISH divides the image into tiles and applies a scanning window strategy to identify
806 zones of co-localisation. Specifically, STRISH starts by dividing the image into four non-overlapping
807 tiles, each one-fourth of the size of the original slide scan, and gradually splits these large tiles into
808 smaller windows, until a given cell count threshold is met (here, <100 cells per window). Results are
809 then scored to measure the density of cells’ expressed markers (e.g. specific ligands and receptors),
810 normalised across the windows with a positive score, and used to plot a heatmap showing the co-
811 localisation scores of the two markers or cell types being tested.

812

813 We also calculated co-occurrence scores between cell types using a conditional probability strategy.
814 The samples were scanned and stitched into large whole-slide images (>2000 µm in length) while
815 preserving the original form of the biopsy. Cell type co-occurrence was calculated between pairs of cell

816 *Supplementary Methods*

817 types across increasing distance intervals (0 to 600 µm). Note that because the skin biopsies are wide
818 but narrow, the cell type co-occurrence scores are expected to diverge after converging. The function
819 to calculate the conditional probability of cell type co-occurrence at different distance intervals was
820 adapted from the co-occurrence analysis method implemented in the Squidpy package¹⁵.

821 **6. RNAscope for multiplexed RNA in-situ hybridisation and data pre-processing**

822 **Sample processing**

823 RNA *in situ* hybridisation was performed using RNAscope for three cSCC and three BCC samples from
824 the same patients (patients B18, E15 and F21). Two consecutive sections of 10 µm thickness were
825 taken from OCT-embedded tissue blocks for the assay and for a negative control, respectively. Slides
826 were stained with a mixture of five RNA probes, namely THY1 (ADV430611T2), IL34 (ADV313011T3),
827 CSF1R (ADV310811T4), CD207 (ADV809521T7), and ITGAM (ADV555091T8) designed by ACD (Cat.
828 No. 324110). The negative control slide was stained with the RNAscope HiPlex 12 Negative Control
829 Probe included in the kit. After RNA hybridisation, specific signal was amplified with high efficiency using
830 RNAscope HiPlex Amp reagents. Tissue images were captured by an Axio Z1 slide scanner (Zeiss)
831 after adjusting fluorescent intensity. The imaging was performed in two rounds with identical microscopy
832 parameters for each round. The first round captured nuclei (DAPI), THY1 (Cy3), IL34 (Cy5) and CSF1R
833 (Cy7), while the second captured nuclei, CD208 (Cy5) and ITGAM (Cy7). Images from the two rounds
834 were processed, stitched, and post-processed by ZEN software.

835 **Data processing**

836 We sought to map interactions between the LR pair IL34_CSF1R using RNAscope data. First, similar
837 to the Polaris data processing methods described above, we performed cell segmentation across the
838 images using starDist¹². Next, we converted imaging data to a quantitative RNA expression matrix by
839 determining the mean intensity of RNA fluorescent signal within each cell boundary. We performed data
840 pre-processing to remove outlier signals and false cell detection events. Subsequently, we applied
841 STRISH to detect cell co-localisation, by splitting each scanned image into multiple non-overlapping
842 windows and quantified the IL34 and CSF1R co-localisation within every window. Each window
843 expressing both CSF1R and IL34 were statistically tested against the random combination of two pairs
844 of markers available for the same window and across all IL34_CSF1R-positive windows. Permutation
845 tests were performed testing different random LR pairs besides the target CSF1R and IL34 to determine
846 the most significant ones.

847

848 **7. CODEX (Phenocycler-Fusion, Akoya Biosciences)**

849 Single-cell spatial phenotyping of the Melanoma FFPE slides was performed in collaboration with Akoya
850 Biosciences (The Spatial Biology Company, Marlborough, MA, USA) on the Phenocycler®-Fusion (PCF)
851 platform. The tissue slides were stained with a 33-plex antibody panel, including markers for immune
852 checkpoints, immune cell lineages, activation states, tissue structure, and metabolism, in a single step.

Supplementary Methods

853 Then, on the Phenocycler® -Fusion, combinations of three antibodies were visualized by utilising by utilising
854 iterative fluorescent-reporter addition and imaging cycles. Commercially available oligo-conjugated
855 antibodies were obtained from Akoya Biosciences (The Spatial Biology Company, Marlborough, MA,
856 USA). The details of the antibodies are shown in the **Supplementary Table S2**.

857

858 PCF (Massachusetts, US) immunostaining was performed according to the PCF user manual and
859 previously published method (Goltsev et al. 2018, Black et al. 2021). First, 5 µm tissue sections were
860 baked for 12 hours at 55°C on a slide warmer (Premiere XH-2004) and dewaxed through two rounds of
861 5-minute incubation in Histochoice Clearing Agent (VWR, PN# H103-4L). Subsequently, the FFPE
862 sections were dehydrated by sequentially incubating in 100% ethanol (Sigma Aldrich, PN# 79317-16GA-
863 PB), 90% ethanol, 70% ethanol, 50% ethanol and 30% ethanol each twice for 5 minutes. Slides were
864 then rinsed three times in distilled water for 5 minutes to ensure no carryover of ethanol before target
865 retrieval. Next, heat-mediated epitope retrieval was performed by incubating slides in a Coplin jar
866 containing Tris EDTA solution (pH 9; Dako, #S2367) in a pressure cooker for 20 minutes. After 20
867 minutes, the Coplin jar was removed from the pressure cooker and left to equilibrate at room temperature
868 for at least 30 minutes. Slides were then rinsed through two rounds of distilled water for 2 minutes and
869 stored in Hydration Buffer from the PhenoCycler Staining Kit (Akoya Biosciences, MA, #7000008) until
870 ready for staining.

871

872 After the pretreatment, the FFPE sections were stained with Autostainer (Parhelia). Samples were first
873 allowed to equilibrate to room temperature in Staining Buffer from the PhenoCycler Staining Kit for 20-
874 30 minutes, followed by incubation in a pre-blocking solution made of N, J, and S blockers in the Staining
875 Buffer from the PhenoCycler Staining Kit (Akoya Biosciences, MA, #7000008). Then the FFPE sections
876 were stained with an antibody cocktail with optimal dilutions of each antibody and N, J, S, and G blockers
877 for 3 hours at room temperature. After washing with Staining buffer and post-fixation with 1.6 % PFA for
878 10 minutes, samples were fixed in methanol (Sigma Aldrich, MO, #34860) for 5 minutes. Lastly,
879 PhenoCycler Fixative reagent was applied for fixation for 25 minutes, and slides were stored in Storage
880 Buffer from the PhenoCycler Staining Kit at 4°C until ready to image.

881

882 Before imaging, the antibody-stained slides were equilibrated at room temperature in 1X CODEX Buffer
883 for 10 minutes. Next, the flow cell (Akoya Biosciences, #240204) was assembled onto the slide as
884 described in the user manual. Then, the slides were incubated in 1X CODEX buffer again for 10 minutes
885 to ensure secure sealing of the flow cell to the slide.

886

887 The reporter, fluorescently labeled–oligonucleotide probes (fluorescent-reporters) that corresponded to
888 DNA-barcoded antibodies for multiple cycle immunostaining, was prepared as described in the
889 PhenoCycler Fusion User Guide. Reporter stock solution was prepared with 10X CODEX Buffer
890 (#7000001), Assay Reagent (#7000002) and Nuclear Stain (#7000003) from Akoya Biosciences.

Supplementary Methods

Individual tubes of 3 reporters/cycle were diluted in reporter stock solution. Subsequently, a 96 well plate was prepared with 1 well/cycle containing the corresponding working reporter solution for that cycle. Blank cycles containing only reporter stock solution without any added reporters were included as the first and last cycle in each run for subtraction of autofluorescence background. The PCF is an integrated ultrahigh plex cycling platform, including the PhenoCycler automated fluidics cycler and the Phenolmager Fusion imaging system which automates the entire process of reporter hybridization, imaging, and dehybridization to capture whole slide images of three markers (+DAPI for nuclear staining) at a time. First, the imaging protocol was set up using Fusion Experiment Designer software according to the manufacturer's protocol. Then, two blank cycles (only DAPI nuclear stain without any fluorescent reporters) at the first and the last cycle of the experiment were run to evaluate the level of autofluorescence and subtract background using Fusion software. After images of all cycles had been acquired, the final QPTIFF file containing a composite image of all markers was viewed using the QuPath software (<https://qupath.github.io>), where each channel can be turned on and off individually or collectively to reveal the spatial expression pattern of the marker(s) of interest.

8. Spatial glycomics

The tissue sections mounted on ITO were placed onto a heating plate at 65 °C for 60 min. Paraffin was removed by incubating the slides in xylene (2 x 2 min) and tissue sections were re-hydrated by different washes: 100% ethanol (1 x 2 min), 90% ethanol (1 x 2 min), 70% ethanol (1 x 2 min), 50% ethanol (1 x 2 min), water (1 x 2 min) and completely dried in the desiccator. The sialic acids were stabilised by methylamidation (PMID: 38953530). The derivatization was performed by incubating the tissue slides in derivatization solution: (1 M methylamide hydrochloride, 0.5 M methylmorphine and 50 mM PyAOP in DMSO) for 2 x 5 min at ambient temperature. After derivatization, tissue sections were washed in 100% ethanol (2 x 2 min), Carnoy solution (60% ethanol, 30% chloroform, 10% acetic acid, 1 x 5 min) and water (1 x 1 min). Slides were dried in a vacuum desiccator (~10 min) before antigen retrieval. Citraconic anhydride buffer (~10 mL, pH = 3) based antigen retrieval was carried out in a commercial vegetable steamer (Philips All-In-One Cooker, HD2237) for 20 min. Following the antigen retrieval, slides were cooled down for 5 min, rinsed in water and dried in the vacuum desiccator before enzymatic deglycosylation. PNGaseF (0.1 µg/µL in 25 mM Ammonium bicarbonate) was deposited using the HTX TM-Sprayer (HTX Technologies, USA), 15 layer passes at 25 µL/min, velocity of 1200 mm/min, crisscross pattern, 3.0 mm track spacing, nitrogen gas pressure was set to 10 psi. N-glycans were released in a 3-hour incubation at 37 °C in a humidity chamber protected from evaporation. After incubation the ITO slides were briefly dried in the vacuum desiccator and the matrix (10 mg/mL CHCA, 50% ACN and 0.1% TFA) was applied using HTX TM-Sprayer (5 layers, 0.1 mL/min, 80 °C, velocity of 1300 mm/min, crisscross pattern, 2.5 mm track spacing, nitrogen gas pressure was set to 10 psi).

MALDI MSI data were acquired using a rapifleX MALDI TOF/TOF mass spectrometer (Bruker, Daltonics) equipped with a Smartbeam 3D 10 kHz laser and operated in a positive-ion reflectron. The

Supplementary Methods

929 laser power was optimized at the start of each run and then held constant during the MALDI–MSI
930 experiment. At each sampling position, 250 shots were used to acquire data at 20 µm pixel resolution
931 for the *m/z* 920–3500 range. All of the methods used were internally calibrated using commonly present
932 neutral glycans across the mass range. Spectra were normalized by total ion count (TIC) normalization
933 of all data points, unless otherwise stated. Preprocessing, background reduction and visualisation of
934 data was performed using SCiLS Lab (Bruker, Daltonics, version 2025b core).

935

Supplementary Notes - Additional Results

Integrating 12 Spatial Technologies to Characterise Tumour Neighbourhoods and Cellular Interactions in three Skin Cancer Types

P. Prakrithi^{1,2#}, Laura F. Grice^{1,3#}, Feng Zhang^{1,2#}, Levi Hockey^{1,2}, Samuel X. Tan⁴, Xiao Tan^{1,2}, Zherui Xiong^{1,2}, Onkar Mulay^{1,2}, Andrew Causer^{1,2}, Andrew Newman^{1,2}, Duy Pham¹, Guiyan Ni¹, Kelvin Tuong⁵, Xinnan Jin^{1,2}, Eunju Kim^{3,4}, Minh Tran¹, Hani Vu^{1,2}, Nicholas M. Muller⁴, Emily E. Killingbeck⁶, Mark T. Gregory⁶, Siok Min Teoh¹, Tuan Vo¹, Min Zhang⁷, Zachary Reitz⁶, Katharina Devitt⁸, Liuliu Pan⁶, Arutha Kulasinghe⁵, Yung-Ching Kao⁴, Michael Leon⁶, Sarah R. Murphy⁶, Hiromi Sato⁶, Jazmina Gonzalez Cruz⁸, Snehlata Kumari⁸, Hung N. Luu⁹, Sarah E. Warren⁶, Chris McMillan^{10,11}, Joakim Henricson^{12,13}, Chris Anderson^{12,13}, David Muller^{10,11}, Arun Everest-Dass¹⁴, Blake Obrien¹⁵, H. Peter Soyer⁴, Ian Frazer⁷, Youngmi Kim⁶, Mitchell S. Stark⁴, Kiarash Khosrotehrani⁴, Quan Nguyen^{1,2}

Note S1: Confirmation of cell types, differentially expressed genes, and cancer markers using GeoMx protein and RNA assays

We generated GeoMx data from three cSCC patients (R01, P04 and B18) using both protein and RNA modalities, and captured data from CD45+ and PanCK+ tissue segments across a total of 12 paired regions of interest (ROIs) (**Fig S7**).

Differential expression analysis (protein data)

Differential protein expression analysis between CD45+ and PanCK+ segments (**Fig S8b**) revealed that 24 of 48 captured protein markers showed differential abundance between segments, with 10 genes enriched in CD45+ segments and 14 in PanCK+ segments (**Table S8**). As expected, PanCK was amongst the differentially expressed proteins enriched in the PanCK+ segment, and CD45 was enriched in the CD45+ segment. Numerous other cancer-associated markers were also upregulated in the PanCK+ segments, including oestrogen receptor alpha (ER.alpha), which has been associated with cutaneous cSCC progression ¹⁶ and is negatively associated with survival in oral cSCC ¹⁷, the cell proliferation-associated protein Ki.67, and the melanocyte and melanoma marker MART-1 ¹⁸ (whose encoding gene, MLANA, is mapped in melanoma tissue in **Fig S10a**). CD44 and CD80 were also upregulated in the PanCK+ segment; these markers have both been associated with tumour-initiating stem cells ¹⁹. Meanwhile, immune markers including monocyte/macrophage markers CD163 and CD14 and T cell markers CD3 and CD4 were all enriched in the CD45+ fraction (**Table S8**).

Differential gene expression analysis (RNA data)

We first visualised the RNA expression of key marker genes associated with rarer cell types from our Level 2 scRNASeq cSCC atlas (**Fig S3e**, **Fig S8b**). We observed a statistical upregulation in the CD45+ segments of M2 marker gene CD163 and fibronectin (**Fig S8b**) (plus fibroblast marker FAP, which was

Supplementary Results

differentially expressed in the RNA (Fig S8b) and protein (**Fig S9n**) data). Markers for the other tested cell types (except Tregs) followed similar qualitative trends to the protein data, appearing to be more highly expressed in the CD45+ segments, albeit without statistical significance. The presence of these markers in both the protein and RNA datasets lends further support for the existence of these cell types.

We also performed differential gene expression analysis of CD45+ and PanCK+ ROIs for the RNA data, and uncovered a total of 267 of 1,825 genes exhibiting differential expression between conditions ($\log FC > 1.5$ and $FDR \leq 0.05$), including 107 genes enriched in CD45+ segments and 160 genes enriched in PanCK+ segments (**Fig S8c**). Amongst the top 50 PanCK-associated genes were four KC marker genes (*KRT5*, *KRT6A.B.C*, *KRT6B*, and *KRT14*)²⁰, and two members of the S100 gene family (*S100A8* and *S100A9*), which have been linked with cSCC tumorigenesis²¹. The rest of the top 50 PanCK+ markers also include cSCC markers *NECTIN1*, *EGFR* and *COL17A1* (**Fig S8c**). Thirty-three of the genes that were differentially expressed between CD45+ and PanCK+ segments (**Fig S8a**) were also members of the core healthy or cSCC gene suites identified in our scRNASeq atlas (**Fig S3f**), including six cancer-associated genes and 27 healthy-associated genes. *CD74* was upregulated in the CD45+ segment in the GeoMx data and was detected in the core cSCC gene suite. All other genes were upregulated in the PanCK+ segment (**Fig S8a**). It should be noted that both the CD45+ and PanCK+ segments in the GeoMx dataset are derived from cSCC samples, while the scRNASeq core gene suites were calculated separately for healthy and cancer biopsies, but without regard to cell type.

GO enrichment analysis of the top 100 differentially expressed genes per segment (**Fig S8e; Table S16**) revealed that the PanCK fraction is enriched for GO terms associated with keratinisation (GO:0031424) and cell junction organisation (GO:0034330). Meanwhile, the CD45+ fraction is strongly associated with angiogenesis (GO:0001525), wound healing (GO:0042060), ECM-related genes (GO:0030198 and GO:0043062), including collagen genes *COL1A1*, *COL1A2*, *COL5A1* and *COL3A1*. These ROIs were also enriched for GO terms related to cell chemotaxis (GO:0060326) and immunological functions such as the humoral immune response (GO:0002920), consistent with these regions being selected on the basis of CD45 expression.

Cell-type Deconvolution analysis (RNA data)

We performed deconvolution using SpatialDecon⁸ to estimate the cell type composition of our twelve CD45+ ROIs (**Fig S8f; Table S17**). This analysis revealed that CD45+ ROIs have multiple types of both B cells (naive and memory B cells) and CD4 T cells (CD4 naive and memory T cells), plus T regulatory cells, macrophages, and endothelial cells consistently present across all ROIs in abundance. Some samples also show presence of CD8 T cell types (CD8 naive and memory T cells) and NK cells. As the SpatialDecon reference profile is designed primarily for detecting immune cells, and explicitly excludes

Supplementary Results

known cancer markers⁸, we did not perform a comparable deconvolution analysis for the PanCK+ tumour segment regions (see *Methods for further details*).

Note S2: Polaris for a targeted study of interactions via PD-1 and PD-L1

We used Polaris high-plex protein staining to perform a focussed analysis of cell co-localisation and CCI in spatial context for three cSCC patients (R01, P04 and B18). Based on combinations of an antibody panel of six proteins (**Fig S21a**), we were able to identify major cell types, including epithelial (PanCK+), innate immune (CD68+), and adaptive immune (CD8+, FoxP3+) cells (**Fig S21b**) from all tissue samples.

In a non-cancer state, the immune cell-bound PD-1 receptor acts as an immune checkpoint inhibitor which, when bound by its ligand PD-L1 (or PD-L2), dampens the immune response^{31,32}; this serves as an important feedback mechanism limiting autoimmunity³³. In cancer, PD-L1 expression is co-opted by cancer cells as an immune evasion mechanism, inducing immune suppression of activated T cells³⁴. For this reason, inhibitors of the PD-L1/PD-1 interaction have proved valuable for the treatment of metastatic skin and other cancers³². We therefore sought to identify immune cells expressing PD-1 and cancer cells expressing PD-L1 sitting in close proximity to one another. While this does not provide direct evidence of CCIs, it does suggest their cross-activity.

We employed our pipeline, STRISH¹⁴, to automatically and unbiasedly detect regions containing T-cells double positive for CD8 and PD-1 in the immediate vicinity of cancerous epithelial cells, which themselves are double positive for PanCK and PD-L1 (**Fig S21c**). Interestingly, for patient B18, we found that the PD-1+ immune cells and PD-L1+ epithelial cells mostly co-localised at the interface of cancer and immune infiltration, as defined by a pathologist's annotation (**Fig S21a**). Similar patterns were observed by applying STRISH pipeline to the tissue samples from patients R01 and P04 (data not shown).

We then investigated cell type co-localisation more broadly using spatial community analysis³⁵. We performed a clustering analysis to group tissue regions with similar local densities of various cell types, which for patient B18 resulted in the identification of five cell communities distributed across the tissue (**Fig S21d**). Quantitative assessment of cell composition in each community allowed us to deduce biologically interpretable features of these communities (**Fig S21e**). In particular, Community 0 has a very high density of PD-L1+ epithelial cells, while Communities 2 and 3 are enriched for three types of CD8+ T cells (**Fig S21e**). Visual inspection of the location of Communities 2 and 3 reveals their spatial proximity (**Fig S21d**), which may be explained by the tendency of cancerous epithelial cells to reside densely around the cancer nest, and for immune cells to sit under the epidermis layer³⁶. We also observed that Communities 1 and 4 group together (**Fig S21e**), and that these communities contain a

Supplementary Results

1045 mixture of both cancer and immune cell populations. The spatial distribution of the mixed cancer/immune
1046 communities (**Fig S21d**) aligns with the spatially-specific interaction observed between PD-1 and PD-
1047 L1 along the interface between cancer and immune cells (**Fig S21c**).

1048
1049 We investigated the immune-only (Communities 2 and 3) and cancer-immune (Communities 1 and 4)
1050 groups in greater detail by performing a cell type co-occurrence analysis (**Fig S21f-g**). Briefly, Ripley's
1051 K-function was used to summarise the co-occurrence of specific cell types with all other cell types across
1052 increasing distance intervals ¹⁵. When PD-1+ T cells were set as the reference cell type, these cells
1053 were found to co-occur with all three T cell types (i.e. with themselves, CD8+ T cells, and FoxP3+ CD8+
1054 T cells) at a significantly higher rate than a random distribution, across distances between 0 to 400 µm
1055 (**Fig S21f**). PanCK+ epithelial cells were found to significantly co-occur with themselves and with double-
1056 positive PD-1+ CD68+ immune. Thus, our image-based analysis pipeline has detected biologically-
1057 meaningful neighbourhoods of cell types in cancer tissues.

1058 **References**

- 1059 1. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
- 1060 2. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
- 1061 3. Welch, J. D. *et al.* Single-cell multi-omic integration compares and contrasts features of brain cell identity. *Cell* **177**, 1873–1887.e17 (2019).
- 1062 4. Cable, D. M. *et al.* Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* **40**, (2022).
- 1063 5. Pham, D. *et al.* stLearn: integrating spatial location, tissue morphology and gene expression to find cell types, cell-cell interactions and spatial trajectories within undissociated tissues. *bioRxiv* 2020.05.31.125658 (2020) doi:10.1101/2020.05.31.125658.
- 1064 6. Waskom, M. L. seaborn: statistical data visualization. *Journal of Open Source Software* **6**, 3021 (2021).
- 1065 7. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
- 1066 8. Danaher, P. *et al.* Advances in mixed cell deconvolution enable quantification of cell types in spatial transcriptomic data. *Nat. Commun.* **13**, 385 (2022).
- 1067 9. He, S. *et al.* High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. *Nat. Biotechnol.* 1–13 (2022).
- 1068 10. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).
- 1069 11. Bankhead, P. *et al.* QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).
- 1070 12. Schmidt, U., Weigert, M., Broaddus, C. & Myers, G. Cell Detection with Star-Convex Polygons. in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018* 265–273 (2018).
- 1071 13. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
- 1072 14. Tran, M. *et al.* A robust experimental and computational analysis framework at multiple resolutions, modalities and coverages. *Front. Immunol.* **13**, 911873 (2022).
- 1073 15. Palla, G. *et al.* Squidpy: a scalable framework for spatial omics analysis. *Nat. Methods* **19**, 171–178 (2022).
- 1074 16. Lan, J., Gao, X.-H. & Kaul, R. Estrogen receptor subtype agonist activation in human cutaneous squamous cell carcinoma cells modulates expression of CD55 and Cyclin D1. *EXCLI J.* **18**, 606–618 (2019).
- 1075 17. Doll, C. *et al.* Prognostic significance of estrogen receptor alpha in oral squamous cell carcinoma. *Cancers* **13**, (2021).

- 1094 18. Bhatt, M. D. *et al.* The accuracy of detecting melanoma on frozen section melanoma antigen
1095 recognized by T cells 1 (MART-1) stains and on permanent sections of previously frozen tissue: A
1096 prospective cohort study. *J. Am. Acad. Dermatol.* **84**, 1764–1766 (2021).
- 1097 19. Miao, Y. *et al.* Adaptive immune resistance emerges from tumor-initiating stem cells. *Cell* **177**,
1098 1172–1186.e14 (2019).
- 1099 20. Voiculescu, V. *et al.* From normal skin to squamous cell carcinoma: A quest for novel biomarkers.
1100 *Dis. Markers* **2016**, 4517492 (2016).
- 1101 21. Choi, D.-K. *et al.* Clinicopathological roles of S100A8 and S100A9 in cutaneous squamous cell
1102 carcinoma in vivo and in vitro. *Arch. Dermatol. Res.* **306**, 489–496 (2014).
- 1103 22. Arias-Mejias, S. M. *et al.* The role of integrins in melanoma: a review. *Int. J. Dermatol.* **59**, 525–534
1104 (2020).
- 1105 23. Xu, Z. *et al.* Integrin β 1 is a critical effector in promoting metastasis and chemo-resistance of
1106 esophageal squamous cell carcinoma. *Am. J. Cancer Res.* **7**, 531–542 (2017).
- 1107 24. Achen, M. G. *et al.* Vascular endothelial growth factor D (VEGF-D) is a ligand for the tyrosine
1108 kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4). *Proc. Natl. Acad. Sci. U. S. A.* **95**,
1109 548–553 (1998).
- 1110 25. Majumder, M. *et al.* Co-expression of α 9 β 1 integrin and VEGF-D confers lymphatic metastatic ability
1111 to a human breast cancer cell line MDA-MB-468LN. *PLoS One* **7**, e35094 (2012).
- 1112 26. Northcott, J. M., Dean, I. S., Mouw, J. K. & Weaver, V. M. Feeling stress: The mechanics of cancer
1113 progression and aggression. *Front Cell Dev Biol* **6**, 17 (2018).
- 1114 27. Huang, Y.-L. *et al.* Collagen-rich omentum is a premetastatic niche for integrin α 2-mediated
1115 peritoneal metastasis. *Elife* **9**, (2020).
- 1116 28. Collo, G. & Pepper, M. S. Endothelial cell integrin alpha5beta1 expression is modulated by
1117 cytokines and during migration in vitro. *J. Cell Sci.* **112** (Pt 4), 569–578 (1999).
- 1118 29. Wang, T. *et al.* Downregulation of miR-205 in migrating epithelial tongue facilitates skin wound re-
1119 epithelialization by derepressing ITGA5. *Biochim. Biophys. Acta* **1862**, 1443–1452 (2016).
- 1120 30. Budden, T. *et al.* Ultraviolet light-induced collagen degradation inhibits melanoma invasion. *Nat.*
1121 *Commun.* **12**, 2742 (2021).
- 1122 31. Ishida, Y., Agata, Y., Shibahara, K. & Honjo, T. Induced expression of PD-1, a novel member of the
1123 immunoglobulin gene superfamily, upon programmed cell death. *EMBO J.* **11**, 3887–3895 (1992).
- 1124 32. Tsai, K. K., Zarzoso, I. & Daud, A. I. PD-1 and PD-L1 antibodies for melanoma. *Hum. Vaccin.*
1125 *Immunother.* **10**, 3111–3116 (2014).
- 1126 33. Chen, D. S. & Mellman, I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* **39**,
1127 1–10 (2013).
- 1128 34. Shi, L., Chen, S., Yang, L. & Li, Y. The role of PD-1 and PD-L1 in T-cell immune suppression in
1129 patients with hematological malignancies. *J. Hematol. Oncol.* **6**, 74 (2013).
- 1130 35. Schürch, C. M. *et al.* Coordinated cellular neighborhoods orchestrate antitumoral immunity at the
1131 colorectal cancer invasive front. *Cell* **183**, 838 (2020).

1132 36. Wong, C. E. et al. Inflammation and Hras signaling control epithelial–mesenchymal transition during
1133 skin tumor progression. *Genes & Development* vol. 27 670–682 Preprint at
1134 <https://doi.org/10.1101/gad.210427.112> (2013).

1135