

1

2 **Integrating 12 Spatial and Single Cell Technologies to Characterise**

3 **Tumour Neighbourhoods and Cellular Interactions in three Skin Cancer**

4 **Types**

5

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

15

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

37 #co-first authors

38 Correspondence: quan.nguyen@qimrberghofer.edu.au

39

40

41

42

43

44

45

46

47

48

49

50

51

52 Abstract

53 Cutaneous squamous cell carcinoma (cSCC), basal cell carcinoma (BCC), and melanoma, the three
54 major types of skin cancer, account for over 70% of all cancer cases. Despite their prevalence, the skin
55 cancer microenvironment remains poorly characterized, both in the outer skin layer where the cancer
56 originates and at the deeper junctional and dermal layers into which it progresses. To address this, we
57 integrated 12 complementary spatial single-cell technologies to construct orthogonally-validated cell
58 signatures, spatial maps, and interactomes for cSCC, BCC, and melanoma. We comprehensively
59 compared and integrated these spatial methods and provided practical guidelines on experimental
60 design. Integrating four spatial transcriptomics platforms, we found keratinocyte cancer signatures,
61 including six consistently validated gene markers. Spatial integration of transcriptomics, proteomics, and
62 glycomics uncovered cancer communities enriched in melanocyte–fibroblast–T-cell colocalization with
63 altered tyrosine and pyrimidine metabolism. Ligand-receptor analysis across >700 cell-type
64 combinations and >1.5 million interactions highlighted key roles for CD44, integrins, and collagens, with
65 CD44-FGF2 emerging as a potential therapeutic target. We consistently found differential interactions
66 of melanocytes with fibroblasts and T-cells. We validated these interactions using Opal Polaris,
67 RNAscope, and Proximal Ligation Assay. To integrate population-scale data, genetic association
68 mapping in >500,000 individuals suggested SNPs enriched for spatial domains containing melanocytes,
69 dysplastic keratinocytes, and fibroblasts, shedding light on functional mechanisms linking genetic
70 heritability to cells within cancer tissue. This publicly available multiomics resource offers insights into
71 the initiation and progression of the most lethal skin cancer (melanoma) and the most common forms
72 (cSCC and BCC) and can be explored interactively at <https://skincanceratlas.com>.

73 Background

74 A comprehensive single-cell, spatial map detailing the cellular organization and molecular signatures of
75 the three most common skin cancers remains unavailable. The keratinocyte cancers basal cell
76 carcinoma (BCC; ~75-80%) and cutaneous squamous cell carcinoma (cSCC; ~20%), are the most
77 prevalent malignancies. Though they have low mortality rates, their high incidence results in substantial
78 public health burdens, with ~3.5 million cases treated annually in the U.S. at a cost of ~\$4.8 billion
79 (Rogers et al., 2015; Guy et al., 2015). About a third of skin cancer-related deaths in Australia are due
80 to keratinocyte cancer (Czarnecki, 2024). In contrast, melanomas account for less than 10% of
81 diagnosed skin cancers but exhibit a markedly higher case fatality rate, constituting the majority of all
82 skin cancer-related deaths (Siegel et al., 2022).

83 The single-cell architecture of skin cancer microenvironments remains largely unexplored. All three
84 major skin cancer types originate in the epidermis, with BCC and cSCC arising from keratinocytes (KCs)
85 and melanoma from melanocytes (**Fig 1a**). Despite the common environmental risk factors, such as
86 solar UV exposure and the many overlapping genetic and anatomical attributes, it remains unclear which
87 factors determine the initiation and progression of skin cancer and why most cSCCs and BCCs are less
88 invasive than melanoma (Wu et al., 2014; Landi et al., 2020; Seviiri et al., 2022). Available single-cell
89 datasets for melanoma are limited and produced by early protocols of the single cell technologies with
90 fewer cells (Tirosh et al., 2016; Zhang et al., 2022). These studies mostly described acral melanoma,
91 uveal melanoma, or *in vitro* melanoma cell lines (as reviewed by Lim and Rigos, 2024). There are no
92 single cell RNA sequencing (scRNASeq) datasets available for melanoma that have representative cell
93 types spanning from benign to melanocytic lesions and invasive melanoma. Similarly, no scRNASeq
94 datasets exist for matched non-cancer and cancer lesions from the same cSCC/BCC patients. Single
95 cell transcriptome profiles for ‘purified’ melanoma cells or just cancer keratinocytes from patient lesions

96 are not available, and so transcriptional profiles of these cancer cells from patient tissues have not been
97 well-characterised. Spatial data for skin cancer at single cell resolution, whole transcriptome coverage,
98 with multimodal measurements of RNA, proteins and metabolites is not available.

99 To understand skin cancer initiation, comprehensive investigation of the skin tissue microenvironment,
100 the cancer cells and their dynamic interactions with immune and stromal cells is needed. One common
101 assumption posits that the differential initiation of one type of skin cancer over another is a stochastic
102 process driven by the random acquisition of UV-induced mutations in susceptible genes in key cell types
103 (Owens et al., 2003; Ratushny et al., 2012). However, healthy skin also carries a high mutational burden,
104 suggesting that DNA mutation alone does not explain the heterogeneity in the initiation of the three
105 cancer types (Martincorena et al., 2015; Sini et al., 2018; Lichtenberger et al., 2021). A determining
106 factor for cancer initiation and progression other than intrinsic DNA mutations is the tumour
107 microenvironment. cSCC and BCC both derive from long-term epidermal residents, such as a subset of
108 basal KCs at the epidermal-dermal junction (Owens et al., 2003; Ratushny et al., 2012; Lichtenberger
109 et al., 2021). In melanoma, there are multiple paths through which melanocytes develop into melanoma
110 cells (Shain and Bastian, 2016). The combination of genetic association studies with the investigation
111 of tumour microenvironment could result in insights on cellular intrinsic and tissue environment drivers
112 of skin cancer types.
113

114 Studying cancer progression also requires investigation at single cell and spatial context. For example,
115 only about 1% of pre-cancerous actinic keratosis lesions progress to cutaneous cSCC, and the factors
116 that differentiate progressor from non-progressor lesions remain unknown (Werner et al., 2013).
117 Comparatively, benign naevi (i.e. moles) have a far lower melanoma transformation rate of ~1/200,000
118 (Tsao et al., 2003). Around 30% of melanomas are derived from benign naevi (Pampena et al., 2017)
119 but the majority arise *de novo* from isolated melanocytes (Marks et al., 1990; Sagebiel et al., 1993;
120 Bevona et al., 2003; Weatherhead et al., 2007). To date, it is mostly impossible to reliably predict the
121 transition from common or dysplastic nevi to melanomas, making regular surveillance the only effective
122 tool for early detection (Goodson et al., 2009). Skin cancers also differ significantly in metastatic
123 potential. BCCs often do not proliferate rapidly and rarely metastasise, while cSCCs are more
124 proliferative and a subset of cSCC (~5%) are highly metastatic (Weinberg et al., 2007).

125
126 To accelerate advances in treatment, beyond excision, drug therapies that can permanently cure these
127 skin cancers are highly desirable but are lacking and so methods to find new targets are needed. For
128 melanoma patients, a common treatment regimen includes targeted therapies (e.g. BRAF- and MEK-
129 inhibitors) and immunotherapies (e.g. anti-CTLA-4 and anti-PD-1). However, only 30-50% of advanced-
130 stage melanoma patients respond to immunotherapies (Sharma et al., 2015; George et al., 2017;
131 Herrscher et al., 2020), and these treatments can cause severe immune-related adverse events (Fink
132 et al., 2021). Ideally, such treatments should only be offered to patients who are likely to respond, but it
133 is not yet possible to confidently stratify patients. Similarly, only approximately 50% of advanced-stage
134 cSCC patients respond to immunotherapy treatment (García-Sancha et al., 2021). To better harness
135 the potential of immunotherapies for melanomas and other skin cancers, it is necessary to identify new
136 ligand-receptor (LR) targets as actionable (e.g. combinational therapies with PD-1/PD-L1). This, in turn,
137 requires an understanding of cancer-immune cell interactions underlying the basis of carcinogenesis.
138

139 Here, we applied the latest single-cell and spatial technologies to map skin cancer cells, cell-cell
140 interactions, and microenvironments. Specifically, we aimed to address keratinocyte cancers (cSCC and
141 BCC) tumour micro-environment that vastly differs from melanomas. We integrated 12 distinct yet
142 complementary technologies to comprehensively identify spatial, single-cell signatures and compare
143 spatial cell-cell interactions (CCIs) to find shared and unique interactions involved in initiation and

144 progression in the three skin cancer types. By harnessing the power of spatial multiomics, this study
145 has provided valuable data resources to understand gene signatures, cells, spatial communities, and
146 cell-cell interactions for the most common cancers and generated new understanding of cell types and
147 their activities that are distinct between the three skin cancer types.

148 Results

149 **Ultraplex, multimodal, multiplatform, single cell and spatial omics data resource for skin cancer**
150 To understand the cellular microenvironment of the skin, we leveraged data from 12 molecular
151 technologies to curate a spatial, single-cell atlas and interactome of healthy (non-sun-exposed) and
152 cancerous skin. Biopsies were collected from 24 skin donors, consisting of patients diagnosed with
153 cSCC ($n = 7$), BCC ($n = 4$), and melanoma ($n = 7$, including $n = 3$ for snRNASeq) and from non-cancer
154 donors ($n = 3$, Visium and Xenium) (**Table S1**). Additional healthy skin samples were collected from
155 non-sun exposed skin of cSCC patients ($n = 5$) (**Table S1**). The diagnosis and biopsy details are
156 described in the Methods section. Each biopsy was measured by up to 12 technologies: Chromium
157 single-cell RNA sequencing (scRNASeq), FLEX single nuclei sequencing (snRNASeq), whole
158 transcriptome Visium spatial transcriptomics, single-cell resolution Xenium spatial transcriptomics,
159 NanoString CosMx Spatial Molecular Imaging (CosMx), NanoString GeoMx Digital Spatial Profiling for
160 proteins (Immune-oncology panel, GeoMx), GeoMx cancer transcriptome atlas (GeoMX CTA), Opal
161 Multiplex Polaris protein assay, RNAScope RNA *in situ* hybridisation, Proximal Ligation Assay (PLA),
162 MALDI-TOF spatial glycomics, and CODEX spatial proteomics (**Fig 1**). In the following sections, we
163 present the results in three themes, including a single cell and spatial atlas (**Fig 1** to **Fig 4**), interaction
164 atlas (**Fig 5** to **Fig 7**) and integration with population genetics data (**Fig 8**).

165 A single-cell atlas of cSCC cancer representing matched healthy and cancer samples

166 We first sought to generate a high-quality single-cell cSCC atlas and compared transcriptomic shifts
167 between non-cancer and cancer skin and signatures of cancer cSCC cells. We performed scRNASeq
168 on 11 paired samples from healthy (non sun exposed) and cancer biopsies collected from five cSCC
169 patients. A total of 45,909 skin cells passed quality control. Initial clustering identified five major clusters
170 ("Level 1" annotations) comprising endothelial cells, fibroblasts, melanocytes, keratinocytes, and
171 immune cell types (**Fig 2a-c**, **Fig S2**). To capture the complexity of the immune microenvironment, a
172 second round of clustering and annotation was performed to find eight major immune subsets ("Level
173 2" annotations; **Fig 2b-c**) and further identified 19 immune cell (sub)types ("Level 3" annotations; **Fig**
174 **2b**, **Fig S2a,c**), representing six T cell subsets, four macrophages, two NK, two Langerhans cells (LC),
175 two DC, B cells, plasma cells and CD14+ monocytes. Next, we characterised KCs, the dominant cell
176 type within our cSCC scRNASeq dataset (70.6% of cells). We assigned the following labels based upon
177 abundant gene expression: Basal KCs (*KRT15+*, *KRT14+*), Differentiating KCs (*PKP1+*, *KRT10+*),
178 Dysplastic KCs (*S100A8+*, *S100A9+*), KC interferon (*IFI27+*), KC cornified (*SBSN+*, *KRT2+*, *DSC1+*)
179 and KC hair (combination of *KRT6B+*, *KRT17+*, *KRT16+*, *KRT5+*), (**Fig 2c**, **Fig 2Sb**). Our annotation
180 for scRNASeq is supported by mapping the cells to the skin tissues, where the cell labels are transferred
181 from scRNASeq to CosMX data (**Fig S2d**). The integration and comparison between scRNASeq and
182 CosMX data provided evidence for accurate cell type annotation based on expected distribution of cell
183 types to skin layers.

184 Integrating gene signatures and inferred copy number variation to identify cancer keratinocyte 185 (KC) cells

186 To date, little is known about transcriptional signatures of KC cancer cells at single cell resolution. We
187 developed a stringent pipeline to map KC cancer cells (**Fig 2d-g**). A cell was considered a cancer KC

189 cell if the cell had abnormal polyploidy based on CNV analysis (using consensus results from two
190 separate CNV inference methods), and had high cancer module scores as calculated for genes that
191 were upregulated in tumour compared to normal tissues (**Fig 2d-g**). This way, we identified a total of
192 745 KC cancer cells. Although we did not *a priori* restrict KC cancer cells to belong to a specific KC
193 subtype, the majority (82.6%) of these cells were classified among dysplastic KCs, further supporting
194 that they are most probably cancer cells from the intersection of three lines of evidence. This stringently
195 defined 'cancer' KC population enabled the analysis of their gene signatures as discussed later.
196

197 Overall shifts in cellular composition between non-cancer and cancer skin samples revealed by 198 scRNAseq data

199 Our single-cell cSCC atlas incorporated matched healthy and cancer samples from 5 patients and
200 comprised 20,827 (45%) cells from cancer biopsies and 25,082 cells (55%) from healthy samples (**Fig**
201 **S3a-b**). scRNAseq-defined differentially expressed genes between KC cancer samples and matched
202 healthy samples show expected gene markers such as *S100A7* and *KRT6B*, (**Fig S3c,d**). We compared
203 cell abundance for patient-matched cSCC and also for melanoma and found overall consistent changes
204 in skin cancer compared to non-cancer, with increased abundance of immune cells, fibroblast and
205 endothelial cells in cancer samples (**Fig 4a**). Fibroblast proportions were consistently higher in
206 melanoma and cSCC-BCC than in non-cancer tissue, corroborating the role of fibroblasts in cSCC
207 development (Schütz et al., 2023). The increased presence of both the lymphoid (i.e. T, B and NK cells)
208 and myeloid (i.e. monocytes, macrophages and dendritic cells) in malignant skin matched the elevated
209 expression of immune gene signature depicted by the core gene suite analysis (**Fig S3d,f-g**). T cells
210 displayed the highest differences in abundance between cancer and non-cancer biopsies (**Fig 4a**, **Fig**
211 **S3e**). For KC populations, the basal and differentiating KCs were more prominent in healthy skin,
212 whereas dysplastic and IFN KCs were enriched in cSCC samples, and no significant difference in the
213 proportion of KC hair and KC cornified across conditions.

214 215 A single-cell reference resource for melanocytic lesions from benign to dysplastic naevus and 216 invasive melanoma

217 Limited single-cell datasets for cutaneous melanoma were available (Tirosh et al., 2016), a mixed acral
218 and cutaneous melanoma study (Zhang et al., 2022), and uveal melanoma and *in vitro* melanoma cell
219 lines (as reviewed by Lim and Rigos, 2024). We opted to produce the latest and more comprehensive
220 reference of melanoma cell types using single-nuclei sequencing of formalin fixed tissues. Three patient
221 samples were selected and scored by 23 pathologists, with one sample defined as definitely invasive
222 melanoma (5591 cells), another as benign naevus (3250 cells) and one severely dysplastic naevus
223 sample (1906 cells). From 10,747 single nuclei, we identified 11 immune cell types and five KC types,
224 as well as endothelial, fibroblast, pericyte, Schwann cell, and sweat gland clusters (**Fig 3a-c**, **Fig S4**).
225 Importantly, we used multiple lines of evidence to distinguish melanoma cells from melanocytes (**Fig**
226 **3d-g**). Here, similar to the approach to define KC cancer cells in cSCC, we integrated CNV analysis,
227 module scores, and spatial mapping of melanocytes to identify 118 melanoma cells with high confidence
228 (**Fig 3d-g**). These cells were used to find signatures of melanomas as described below.
229

230 Changes in single cell transcriptional profiles between non-cancer and cancer specimens

231 Overall, we found more transcriptomic variation between cancer cells from different lineages (3257
232 genes higher in cSCC than in melanomas, and 2713 upregulated in melanocytic lesions than in cSCC)
233 than within a lineage, (176 genes upregulated in cancer KCs vs healthy KCs and 68 upregulated genes
234 in melanoma cells compared to melanocytes) (**Fig 4b**). The 3257 genes higher in cSCC compared to
235 melanomas were enriched for Myc targets, E2F pathways, G2M pathways, mTOR signaling and DNA
236 repair pathways, whereas the 2713 genes higher in melanocyte lesions were most enriched in KRAS

pathway, EMT pathway, IL-2/STAT5 pathway, and UV responses (**Fig S5**). As expected, we found SOX10 higher in melanoma, whereas PTCH2 upregulated in KC cancer.

To find conserved gene expression changes in cSCC, we defined core gene signatures, differentially expressed in all five cSCC patients, revealing 57 genes upregulated in cSCC and 98 genes higher in healthy skin (**Fig S3f**). The core cSCC genes were enriched for immunological process GO terms, such as T-cell mediated cytotoxicity and antigen processing (**Fig S3g**), while the healthy signature was enriched for homeostatic processes, including "establishment of skin barrier" (GO:0061436) with genes like *CLDN1*, *IL18*, *KLF4*, *KRT1*, *NFKBIZ*, and *TP63*, suggesting potential loss of normal balance between proliferation and differentiation and skin integrity in cSCC.

Below, we focus on cancer vs. non-cancer KC and melanomas vs. melanocytes to identify transcriptome-wide signatures unique to well-defined cancer cells relative to their non-cancer counterparts, an important yet understudied topic.

Distinct gene signatures differentiated cancer cSCC from normal KC at single cell resolution

Using the scRNA-seq dataset, we identified 176 genes significantly upregulated in KC cancer cells compared to KC non-cancer cells, with 169 of these not elevated in melanomas. These genes highlighted activation of extracellular matrix remodeling pathways, balancing matrix degradation through *MMP1*, *MMP3*, *MMP10*, *MMP12*, *MMP13*, *SERPINB3*, *SERPINB4*, *SERPINB13*, and *SERPINE2* (**Fig 4c, Fig S5**). Genes linked to differentiation, including *SOX2*, *EOMES*, *DOK6*, *WNT5A.AS1*, *INHBA*, *S100A2*, *S100A7*, *S100A8*, and *S100A9*, were also upregulated. Additionally, specific keratinocyte differentiation markers like *CSPG4*, *SPRR2A*, *FABP5*, and *KRT24* also showed increased expression. Inflammation-related genes, such as *NLRP7*, *TPSB2*, and *TPSAB1*, were enriched, alongside IL-17 pathway components (*MMP1*, *MMP3*, *MMP13*, *S100A7*, *S100A8*, and *S100A9*), suggesting a strong inflammatory signature in cancer KC.

Confirming the new gene signatures of cSCC cancer cells by spatial multiomics

We harnessed data from multiple spatial platforms to identify highly consistent differentially expressed genes upregulated in KC cancer cells compared to non-cancer KC cells. To establish a stringent baseline, we included samples from healthy donors without skin cancer (**Fig S6**). Combining scRNaseq and three spatial transcriptomics platforms - Visium, CosMX, and Xenium - we identified six consistently upregulated genes in KC cancer cells: *SOX2*, *LAMP3 (CD208)*, *CXCL10*, *CXCL9*, *CCL5*, and *UBE2C* (**Fig 4d**). Among these, *CXCL10* and *CXCL9* were the most significantly upregulated. The transcription factor *SOX2*, absent in normal epithelial cells, is essential for cancer-initiating cells in cSCC (Siegle et al., 2014). *CXCL10*, *CXCL9*, and *CCL5* are elevated in cSCC compared to normal skin, playing key roles in tumor progression and T cell infiltration, regulating immune balance in the tumor microenvironment (Tuong et al., 2019). *LAMP3*, a marker for mature dendritic cells, is elevated in cancer, facilitating antigen processing and T cell activation, while *UBE2C* is important for cell division. In addition, visual assessment of *SOX2* expression in UMAP plots (**Fig 2g** vs. **Fig 4e**) and in spatial tissue plots (**Fig 4f**) highlights its specificity to KC cancer cells.

Identified new gene signatures distinguishing melanomas from melanocytes

The melanoma samples analysed here belong to the common BRAF V600E subtype, which accounts for approximately 50% of all melanoma cases. By stringently defining melanoma cells as described above, we identified key marker genes (**Fig 4b-c**, **Fig S4**, **Fig S5**). Among 68 genes highly expressed in melanomas compared to melanocytes, we observed strong enrichment for signaling and cellular interactions. These included cytokine-cytokine interactions (*CXCL2*, *CXCL9*, *CXCR3*, *IL3RA*, *IL18RAP*, *IL2RB*) and immunoregulatory pathways (*NCR1*, *SELL*, *KLRC1*, *SH2D1A*, *CD22*, *LILRA4*, *TREML2*).

286 Markers of differentiation were also upregulated, such as SOX10 (melanocyte development), *FLT3* and
287 *IRF8* (DC differentiation). Genes associated with immune evasion, including *CTLA4*, *CD274 (PD-L1)*,
288 *CXCL9*, *CXCR3*, and *ITGAD (CD11d)*, were also elevated. Additionally, genes linked to melanoma
289 progression and invasion, such as *UBD (FAT10)*, and *STC2*, were upregulated, as were genes
290 associated with melanin synthesis such as *TYR* and *DCT*. In contrast, genes downregulated in
291 melanoma but maintained in normal melanocytes included *CST6* (Cystatin E/M), a known suppressor
292 of melanoma proliferation and migration (Xu et al., 2021).

293 294 **Spatially map cells in stroma and cancer immune compartments in cSCC**

295 We next applied GeoMX Whole Transcriptome Atlas (WTA - 1820 oncogenes) and GeoMX immune-
296 oncology protein panel (48 proteins) to identify cells in the Cancer (panCK+) and stromal-immune
297 regions (CD45+), (**Fig S7; Supplementary Note 1**). To confirm the presence of immune cell types, we
298 used samples from three cSCC patients for whom we also had parallel scRNASeq data (R01, P04 and
299 B18; **Fig S3**). The 48-protein GeoMx panel captured cell surface markers of immune subtypes as
300 identified in our Level 2 and some of the level 3 annotation of the scRNASeq cSCC atlas. This way, the
301 multiomics approach provided independent lines of evidence for confirming the cell types within a
302 sample. With GeoMx data we found M2 macrophages (CD163, CD68), B cells (CD20), CD8+ T cells
303 (CD8), CD4+ T cells (CD4), DCs (CD11c), fibroblasts (FAP-alpha, Fibronectin), and Treg cells (FOXP3,
304 CD25) were all captured in our samples (**Fig S8, Fig S9**), lending further supporting evidence to the
305 existence of these cell types seen in our scRNASeq data. The cell type deconvolution result for GeoMX
306 TCA data shows the heterogeneity between patients, while highlighting the high proportion of T cells
307 across all immune regions adjacent to cancer, including CD4+ T cell, CD8+ T cells, and Treg (**Fig S8f**).
308

309 310 **Detailed spatial map of 21 scRNASeq/snRNASeq defined cell types to transcriptome-wide Visium 311 data, GeoMX WTA, and panel-based, single-cell resolution CosMX data**

312 First, using cell-type signatures for cSCC, BCC and melanoma from scRNASeq/snRNASeq data, we
313 mapped these cell types to spatial Visium data, including 5x cSCC, 3x BCC and 4x melanoma samples
314 across 9x patients (**Fig S10a, Table S1**). We spatially mapped all Level 2 cell types from scRNASeq
315 data to our Visium data, which also matches the pathological/anatomical annotation (**Fig S10a**). The
316 expression of marker genes in the Visium was consistent with histological annotation (**Fig S10**). For
317 example, in melanoma Patient 48974, we observed dark-pigmented cells (i.e. melanophages) at the
318 skin surface of the biopsy (**Fig S10a, top right**). The corresponding Visium spots in this region were
319 predicted to contain a large proportion of melanocytes, based on our deconvolution results (**Fig S10a,**
320 *middle*). These spots also expressed the highest amount of melanocyte marker *MLANA* (**Fig S10a,**
321 *bottom*). This result demonstrates the concordance of multiple layers of information (such as tissue
322 morphology, expression of individual genes, and broader transcriptional signatures), establishing an
323 accurate and comprehensive data resource for the later analysis of gene markers and all interactions.
324

325 Beyond the resolution at the regions of interest (GeoMX WTA and GeoMX protein) or spot level (Visium),
326 we next investigated cell types by CosMx data (He et al., 2022) that provides spatial information of
327 131,804 cells (34.4% cSCC, 34.3% BCC, 31.3% melanoma; **Table S1**). By computational annotation,
328 these cells were assigned to 21 cell types that can all be mapped to the tissue, demonstrating that
329 spatial cell labelling was highly consistent with pathological annotations, but much more detailed (**Fig**
330 **S10b**). CosMx data mapped scRNASeq cell-type signatures (e.g., keratinocytes, melanocytes,
331 fibroblast and immune cells) to distinct layers in the skin, corroborating the accuracy of our scRNASeq
332 cell-type annotation (**Fig S3, Fig S10b**). The subcellular capture resolution of CosMx technology allowed
333 the visual confirmation of cellular location of individual RNA molecules. For instance, in cSCC and BCC,
334 we observed correspondence between the single-cell and single molecular localisation of RNA markers
S100A8 and *KRT17* within the KC cells (**Fig S10**). This accurate resource of CosMx data, with spatial

335 single cell expression of nearly 500 ligand/receptor genes, is valuable for downstream interaction and
336 community analyses as described later.

337 **Single-cell spatial heterogeneity analysis suggested a complex tumour community in**
338 **melanomas**

339 The spatial mapping of CosMX single cell data enabled the characterization of cellular diversity of the
340 cancer microenvironment across the three major skin cancer types. Rao's quadratic entropy score was
341 computed and visualised for each cell in a tissue heatmap (**Fig S11**). A high level of heterogeneity
342 correlated with the mixed distribution of diverse immune cell types, for example an FOV with B cells,
343 plasmacytoid dendritic cells (pDCs), myeloid and T cells (**Fig S11a, b**). Indeed, we observed the highest
344 heterogeneity scores in immune-rich FOVs (**Fig S11d**). To compare cell type heterogeneity across the
345 different skin cancer subtypes, we grouped scores across all FOVs by cancer type, including 30 FOVs
346 from four melanoma patients, 24 FOVs from two BCC patients, and 27 FOVs from three cSCC patients
347 (**Fig S11c, d**). We detected a significant increase in cell type heterogeneity score in the melanoma
348 samples compared to in cSCC cancer (**Fig S11c, d**). This observation may be explained by the trend
349 that melanomas do not adhere to each other as much as KC cancer cells, and so their neighbour cells
350 can be more diverse. However, we noted that heterogeneity assessment would require bigger sample
351 cohorts.

352
353 **Multimodal mapping of communities in melanoma samples with spatial RNA, protein and glycan**
354 **omics**

355 Next, we compared the annotation for the same melanoma tissue block using three single cell resolution
356 platforms representing three modalities, the Xenium (RNA), CODEX (protein) and mass cytometry
357 imaging (MALDI MSI for spatial glycomics), (**Fig S10, Fig S12**). The expression of molecular markers
358 is shown as dot plots in **Fig S12** and the clustering analysis for each modality is shown in **Fig S13-S15**.
359 For all three modalities, the melanoma/melanocytes could be distinctly identified, using label transfer
360 and clustering analysis. We observed the localisation of melanoma markers to the melanocyte layer, for
361 example S100B (CODEX protein), and S100B and MKI67 (CosMX), and TYR gene (Xenium), (**Fig**
362 **S10b-f**). With 260 genes in the skin cancer panel, the single cell Xenium data could define 17 cell types
363 where their locations matched the pathological annotation of the tumor and immune cells (**Fig S10c**).
364 All key KC cell types were mapped with the Xenium data. In contrast, the CODEX data could not map
365 KC cells, due to lack of protein markers for these cell types, but could clearly pinpoint additional immune
366 cell types such as the M2 Macrophages and Neutrophils (**Fig S10d**). Although the cell-type clustering
367 using glycomics is less defined, it is clear that melanoma/melanocytes exhibited unique metabolomic
368 signatures compared to other cell types in the remaining clusters (**Fig S10f, Fig S12a**).
369

370 **Multiplatform integrative analysis of spatial transcriptomics data identified robust tissue**
371 **microenvironments across biological replicates**

372 Spatial data enables mapping the spatial organisation of neighbouring cells within skin tissue, facilitating
373 the identification of functional tissue communities. A key challenge in spatial community analysis is
374 defining shared communities that consistently appear across samples and hold functional or phenotypic
375 relevance. To address this, we integrated three spatial transcriptomics datasets (Visium, CosMX, and
376 Xenium), consolidating neighbourhood information for each cell/spot into a shared matrix to identify
377 meta-communities, consisting of communities with similar cell type composition (**Fig 5a**, see Methods).
378 In particular, we identified a meta-community comprising Visium_2, Xenium_2, Xenium_7, and
379 CosMX_6, all enriched for melanocytes and including fibroblasts, basal KC, T cells, and DCs across all
380 samples (**Fig 5a**). This meta-community analysis enabled us to compare communities across platforms.
381 Both Visium and CosMX member communities (CosMX_6 and Visium_2) were significantly more
382 abundant in melanoma samples than in BCC and cSCC (**Fig S16a, b**).

383
384 A defining feature of spatial communities is the interaction between their members, which can be
385 characterized by ligand-receptor coexpression (**Fig 5c, d**) or cell-cell colocalization (**Fig 5e**). In both KC
386 cancers and melanoma, we found fibroblast interactions particularly dominant. Within the melanoma-
387 associated CosMX_6 community, interactions were highly enriched for collagen signaling, especially
388 between collagens and CD44 (including: COL6A1, COL6A2, COL4A1, and COL1A1) (**Fig 5d**). The co-
389 localization analysis in the melanoma communities in three Xenium samples shows that Treg and
390 Fibroblasts have a high co-occurrence probability with melanocytes (**Fig 5e**). This cross-platform spatial
391 transcriptomics analysis of the melanoma tumor microenvironment was followed by a multimodal
392 characterization of its regulation at the RNA, protein, and metabolite levels, as described below.
393

394 **Cross-modality community analysis characterised molecular signatures in the melanoma 395 microenvironment**

396 A benefit of applying different technologies to measure different classes of biomolecules is the ability to
397 gain a holistic insight into cell and tissue-level functions that may not be fully captured with a single
398 modality. We therefore next investigated the molecular signatures of the melanoma community across
399 multiple analytes (i.e. RNA, protein and metabolites) in spatial context, (**Fig 5f-h**). First, we performed
400 joint pathway analysis of genes/proteins with metabolites using MetaboAnalyst, providing three lines of
401 evidence that Tyrosine metabolism was enriched (supported by 3 genes and 24 compounds in KEGG
402 pathways), (**Fig 5h**). Tyrosine is a critical precursor for melanin production in melanocytes and
403 dysregulation in this pathway can contribute to melanoma development or progression (Najem et al.,
404 2021). Interestingly, we found Pyrimidine metabolism upregulated in the melanoma community. This
405 pathway is fundamental for DNA and RNA synthesis and is linked to sunlight-associated melanomas,
406 melanoma progression and treatment resistance (Edwards et al., 2016; Santoriello et al., 2020).
407 Metabolism of Glycine, Serine and Threonine was also enriched. Integrative analysis of the melanoma
408 community for adjacent sections in Xenium and CODEX suggested that more complex composition of
409 immune cells could be seen in the CODEX data, but a more detailed annotation for fibroblast and
410 keratinocyte could be achieved using the Xenium data (**Fig 5f-g**). Our approach to annotating clusters
411 in each modality before combining and comparing clusters provided flexibility in cross-modality analyses
412 (**Fig S12**).
413

414 **Single-cell predictions of CCI suggest differential interactions in healthy and cSCC cancer 415 patients**

416 Next, we compared the spatial and single-cell interactome in cSCC with other skin cancers based on
417 single-cell and spatial multi-omics to understand tumour progression and potential treatment targets in
418 the tumour microenvironment. First, the scRNASeq cSCC dataset allowed us to predict CCI occurring
419 between specific cell types, using CellChat. A total of 312 unique Ligand-Receptor (LR) pairs signalling
420 were predicted and more interactions were found in cancer samples than the matched non-cancer
421 samples from the same patients (**Fig S17**). While LR interactions varied between patients, core sets of
422 LR pairs were shared across patients. The cancer-specific core LR pairs were enriched for immune-
423 related functions including MHC class II complex assembly (**Fig S17d**), reflecting the more active
424 immune processes occurring in the cancerous biopsies. KCs were the dominant ligand contributors in
425 healthy samples (**Fig S18**). However, in cancer samples from the same patients, we observed an
426 increase in the total number of predicted LR interactions and a shift in the signalling profile for four
427 samples (P04, P30, R01 and B18), with increased signalling from endothelial, fibroblast and immune
428 cells relative to KCs (**Fig S18**).
429

430 **Spatial transcriptomics enhances the accuracy and specificity of LR interaction predictions by 431 incorporating spatial constraints**

432 Tissue dissociation during scRNASeq library preparation removes cellular spatial context, which can
433 result in false positive detection of LR as two cell types predicted to interact in scRNASeq data may
434 derive from distant tissue regions, unlikely to directly interact. We thus performed spatially-constrained
435 two-level permutation (SCTP) in our stLearn software to predict spatially-informed LR interactions. We
436 performed SCTP for CosMX data (**Supplementary Note 2**), (Pham et al., 2023). This analysis revealed
437 interactions that were predicted by scRNASeq but no colocalization was observed, suggesting possible
438 false detection (e.g., XCL1-XCR1), and others that were missed from scRNASeq data analyses, but
439 were detected by Visium (e.g. *WNT5A-ROR1*), (**Fig S19a**). Indeed, we found cases where the
440 scRNASeq missed the interactions, while all three platforms, Visium, CosMX and Xenium, strongly
441 supported the interactions both by statistical significance test and by visual co-expression of the LR
442 pairs between neighbour pairs (e.g., CXCL12-CXCR4, CCL9-CCR7) (**Fig S19b**).
443

444 **Different ligand-receptor interactions specific for a cancer type show important roles of 445 angiogenesis, integrins, and fibroblast growth factors**

446 To identify and compare highly confident interactions, we then applied our multi-platform, multi-sample
447 cell cell interaction analysis (MMCCI) approach to integrate the data to find interactions consistent
448 across biological replicates, followed by differential interaction analysis performed at cell-type network
449 levels (comparing edges connecting two interacting cell types) or at L-R levels (comparing L-R
450 interaction scores). We first performed statistical tests for differential LR scores derived from MMCCI
451 integrated results. MMCCI calculated interaction strength for each LR pair and each cell-cell pair by
452 combining stLearn SCTP scores/p-values across biological replicates. Between the integrated CosMX
453 and Visium datasets, we consistently found 16 LR pairs highly expressed in BCC, 17 in cSCC, and 37
454 in melanoma (**Fig 6a**).
455

456 For BCC, among the LR pairs enriched, we found strong signals for stromal remodelling, especially
457 angiogenesis. Interleukins (*IL6-IL6R*, *IL6-IL6ST*, *IL1B-IL1RN-IL1R2*), chemokines (CXCL2-CXCR1,
458 CCL2-ACKR4) and fibronectins (*FN1-ITGB8* and *FN1-ITGB6*) all have roles in angiogenesis and our
459 data show that they were higher in BCC (Villani et al., 2021). Interactions in the canonical WNT signaling
460 pathways (*WNT5A-FZD7* and *WNT5A-FZD8*; also supported by *EPCAM-EPCAM* interactions)
461 appeared to be more active in BCC. This signaling pathway, although less well known compared to the
462 Hedgehog (HH) signalling in BCC, is crucial in coordinating with the HH pathway to maintain the
463 proliferative state of BCC cells and sustain cancer stem-like cells in BCC (Yang et al., 2008).
464

465 For cSCC, strong enrichment of Osteopontin (*SPP1*) was found with the pairs *SPP1-ITGB1* and *SPP1-
466 ITGAV* (uniquely expressed in cSCC), while *SPP1-CD44* and *SPP1-ITGB5* were upregulated compared
467 to in BCC and Melanoma. Through interactions with integrins and *CD44*, *SPP1* may activate intracellular
468 signaling pathways like *PI3K/Akt*, *MAPK/ERK*, and *FAK* to promote cell survival, proliferation, and
469 growth (Anborgh et al., 2010). The *MAPK/ERK* pathway and *PI3K/Akt* pathways can be activated by
470 *FLT1* and *CSF3-CSF3R* interactions. The interactions associated with angiogenesis like *CD38-PECAM*,
471 *VEGFB-FLT1* (*VEGFR-1*), and *CXCL1-CXCR2* were also upregulated in cSCC, suggesting that
472 angiogenesis is upregulated in both BCC and cSCC, but through different regulation pathways.
473 Calprotectin (*S100A8* and *S100A9*) interact with *TLR4* on immune cells to promote inflammation in the
474 tumour microenvironment.
475

476 For melanoma, we found strong enrichment of collagen interactions (e.g., *COL1A1*, *COL1A2*, *COL3A1*;
477 making 15 out of all the 37 LR pairs that were specific for melanoma compared to BCC and cSCC). We
478 found type I collagen (*COL1A*, *COL1B*) and integrin receptor (*ITGA*, *ITGB*) families to actively interact
479 in melanoma. These included *COL1A1-ITGA2*, *COL1A1-ITGB1*, *COL1A1-ITGA5*, *COL1A2-ITGA2*, and
480 *COL1A2-ITGB1* (**Fig 6a**). Changes in integrin activity have been implicated in differential metastatic and

invasive risks in melanoma (Xu et al., 2017). The collagen-integrin interactions play key roles in microenvironment remodelling, creating protumorigenic and immunosuppressive niches, promoting angiogenesis, and activating the MEK/ERK signalling pathways (Hayashido et al., 2014). Among those collagen interactions, four genes were involved in DNA damage responses (with DDR1 and DDR2). The second most common ligand-receptor pairs that were uniquely increased in melanoma involve Fibroblast Growth Factor, with six out of 37 interactions, including *FGF1-CD44*, *FGF2-CD44*, *FGF2-FGFR1*, *FGF18-FGFR1*, *FGF1-FGFR1*, *FGF9-FGFR1*. FGF signaling can enhance melanoma cell proliferation by activating downstream signaling pathways such as MAPK/ERK, PI3K/AKT, and JAK/STAT, which play key roles in cell cycle progression and survival and some are therapeutic targets for melanoma (e.g., FGF2/FGFR signalling).

Different interactions between cell-cell pairs highlight the roles of fibroblast, T cells and melanocytes/keratinocytes

Using interaction network graph analysis to combine multiple replicates, followed by statistical comparisons for interacting cell types. This test was performed using MMCCI (multiplatform, multimodal CCI), where the interaction strength was defined as the cumulative p-value from all p-values of the same interaction across all samples.

At cell type level, we found that cSCC and BCC were more similar to each other than to melanoma (**Fig S20a**). The most common interactions across the three cancers involve fibroblasts with immune cells and KC cells (**Fig S20a**). We observed stronger fibroblast to T cells interaction in cSCC and BCC compared to in melanoma, whereas the fibroblast to melanocyte interaction was higher in melanoma (**Fig 6b**). cSCC and BCC also had more interactions between fibroblast and KC cells, especially the interactions with differentiating KC (**Fig 6b**).

At L-R pair level, those LR pairs specific for cancer type were enriched for key cancer-related pathways such as EMT (**Fig S20d**). MIF-CD44 was among the top two LR pairs that changed the most between cancer types (**Fig S20b**). This pair displayed strong interactions between macrophages with KC cells, fibroblast and T cells (**Fig S20c**). Among the pairs that most differentially interacting, the canonical CD80-CTLA4 was at the top most active (**Fig S20d**). The most significantly interacting cell-cell pairs for this pair in melanoma were fibroblast with melanocyte and fibroblast with T cell (**Fig S20e**). Based on L-R pair interactions, the roles of CD44, extracellular matrix, and immune system in relation to drug targets for melanoma could be investigated at a systematic level (**Fig 6d**).

Integrative confirmation of LR signalling at RNA level reveals enrichment of IL34-related antigen-presenting pathways in melanoma

The global analysis of all possible interactions between >2000 known LR pairs as described in previous sessions suggested important LR pairs for validation experiments. An example of such an LR pair is *IL34-CSF1R*, which appeared in the top interacting pairs across spatial modalities and was higher in melanoma samples than in cSCC and BCC (**Fig 7a, c, d**). IL34 is a cytokine, predominantly produced by keratinocytes, whose receptor *CSF1R* activates immune cells, in particular macrophages and Langerhans cells (Wang et al., 2012; Stanley et al., 2014). High *IL34* expression correlates with poor survival in lung cancer cell culture models and patients due to *CSF1R*-mediated activation of tumour-associated macrophages (Baghdadi et al., 2016 and 2018). High *IL34-CSF1R* interaction in melanoma was reported to be linked to drug resistance (Giricz et al., 2018).

We found co-localisation of *IL34* and *CSF1R* by RNAScope analysis (**Fig 7a-c**). Although the interaction was strongest in melanoma, using our colocalization analysis pipeline, STRISH, we also observed *IL34-CSF1R* in BCC and cSCC patients, in the dermis at the immune-rich regions (**Fig 7b**), (refer to

530 Supplementary Methods). This observation is consistent with spatial single-cell gene expression data
531 (**Fig 6c**). Both *IL34* and *CSF1R* expression were detected in our healthy and cSCC scRNASeq atlas
532 (**Fig 7d**). Low spatial resolution, transcriptome-wide Visium showed the colocalization of *IL34-CSF1R*
533 in the dermis layer and pathway comparisons for interacting Visium spots with those spots without
534 interactions show enrichment of the antigen processing pathway and lipid metabolism (**Fig 7f, g**). We
535 detected 531 genes upregulated in *IL34-CSF1R* positive spots in melanoma, and 758 spots for BCC.
536 Only the melanoma gene list was found to show any enrichment for GO terms), including terms
537 associated with immune- (**Fig 7f**) and lipid-related (**Fig 7g**) functions, suggesting the important roles of
538 the upregulation of the *IL34-CSF1R* interactions in melanoma.
539

540 **Spatial multimodality validations of ligand-receptor interactions highlight the role of CD44**

541 Beyond RNA level, we also validated protein-protein interactions based on colocalization between
542 neighbour cells (**Fig S21**). We implemented Opal tyramide signal amplification (TSA) protocol, where
543 primary antibody, anti-IgG polymer HRP and covalent labelling with Opal TSA fluorophores were used
544 and the whole slide multispectral scanning was performed by Vectra Polaris. With this method, we
545 scanned through the whole tissue section and identified image tiles containing double-positive CD8+
546 PD-1+ immune cells and PanCK+ PD-L1+ cancer cells (**Fig S21**).

547 The positive colocalization from Opal Polaris suggests evidence for interaction, but the distance can be
548 relatively further away than direct interaction at one location. We next applied proximal ligation assay to
549 detect two proteins within a 20 nm distance on melanoma samples (**Fig 7h**, **Fig S22**). We tested three
550 pairs identified in our spatial omics data analysis, focusing on CD44, a dominant receptor found with
551 distinctively more common interactions in melanoma compared to BCC and cSCC (**Fig 7h**). CD44 acts
552 as an MMP9 docking receptor that localizes MMP9 to the cell surface, where it can degrade components
553 of the ECM such as collagen to enhance tumour invasion (Yu and Stamenkovic, 1999). CD44 was
554 reported to bind fibronectin (FN) to anchor cells to their surrounding ECM, potentially supporting
555 invasion. CD44 interacts with Fibroblast growth factor (FGF1 and FGF2) in melanoma, possibly
556 enhancing tumour initiation and migration. Here we validated CD44-MMP9, CD44-FN1, and CD44-
557 FGF2 interactions. The PLA signal clearly suggests the interactions occur (**Fig 7h**). Future work will
558 validate more pairs where antibodies are available. Moreover, we showed that these LR pairs had
559 significant prognostic values when applying to the public TCGA data (**Fig S23**), further suggesting the
560 roles of cell-cell interactions to cancer phenotypes.

561 **Integrating spatial analysis with genetic association with cancer traits from population scale data**

562 We next integrated spatial omics data with summary statistics from GWAS studies to genetically map
563 skin cancer risk SNPs to spatial cell types and domains. We applied gsMAP to map genetic association
564 signals (SNP effect sizes) for cSCC, BCC, cutaneous melanoma heritability to specific cell types or
565 spatial domains. The mapping was based on 1) gene markers specifically and highly expressed in a cell
566 type or a spatial domain as defined using the spatial transcriptomics data, 2) mapping SNP to these
567 genes according to linkage disequilibrium distance, 3) testing for the significance of the cumulative
568 effects of these SNPs (Cauchy P value) compared to SNPs not associated with these markers genes
569 (Song et al., 2024). For cSCC and BCC genetics studies, we used summary statistics data from a total
570 of 10,557 controls and 537,850 controls for SCC and 36,479 cases and 540,185 controls for BCC
571 (Seviiri et al., 2022). For melanoma we used results from the analysis of with 30,143 clinically-confirmed
572 melanoma cases and 81,405 controls (Landi et al., 2020). Based on spatial gene expression in a
573 neighbourhood of a focal spot/cell, GSS (gene specificity scores) for those genes highly expressed in
574 specific spatial location was calculated. SNPs were mapped to genes in these GSS based on distance
575 to transcription start sites. Considering each spot/cell tagged with a set of SNPs next to GSS genes, the

576 proportion of trait heritability captured by these SNPs (and thus by spot/cell) relative to baseline SNPs
577 was computed based on stratified linkage disequilibrium score regression. Significance of association
578 for a spatial region or a cell type was computed by aggregating P values of spots/cells within that region.
579

580 From Visium data, we found that spatial regions enriched with genetic association for BCC, cSCC and
581 melanoma were localised to the epidermis. The association signal was, in some cases, specific to
582 locations, rather than continuous or evenly distributed in the outer layer of the skin (**Fig 8a**). Cauchy
583 aggregated significance for cell types show that top association for melanoma included melanocytes
584 and KC differentiating. Fibroblast consistently displayed a strong association signal across melanoma
585 and cSCC and BCC samples (**Fig 8b**). This was consistent with the spatial cell-cell interaction analysis,
586 which suggested the important roles of fibroblast in interactions with melanocytes, keratinocytes and
587 immune cells. Cell types most associated with cSCC and BCC cancer are KC dysplastic (more for
588 cSCC), KC hair (more enriched for BCC), and KC cornified (similar level to BCC and cSCC, and much
589 more than melanoma), (**Fig 8b**). Furthermore, we also mapped genetic association signals to single cell
590 resolution CosMX data (**Fig 8c**). The spots/cells in the tumor regions (based on pathological annotation)
591 exhibited the strongest spatial heritability explained by SNPs linked to GSS genes of the spot/cells in
592 this tissue region, followed by those in the immune regions (**Fig 8c**). The tissue region with the lower
593 spatial heritability is the stroma region.
594

595 Next we identified ligand-receptor genes with significant correlation between spatial gene specificity
596 scores and genetic association P value for the cell type. Such pairs suggest possible mechanisms on
597 how the genetic association of signals in a spatial region or a cell type may be explained through
598 dynamic interactions between cell type pairs within the spatial microenvironment. Again, we observed
599 strong interactions between KC dysplastic in cSCC and KC-hair in BCC, and more immune interaction
600 in melanoma, especially those involved in T cells compared to in BCC (**Fig 8d**). Visualisation of genetics
601 association signals between T cells with Melanoma via *IL34-CSF1R* and *LTB-LTBR* suggests tissue
602 regions where the association with the melanoma were most strong (**Fig 8e**). These regions are at the
603 junction of the epidermis and dermal regions for *LTB* and *LTBR* genes, and less specific for *IL34-CSF1R*.
604 Across the whole genome, the top genetic association with spatial expression patterns is consistent with
605 mapping key melanoma markers such as *MITF*, *TYR*, and *MX2*. The *CSF1R*, *LTB* and *LTBR* were in
606 the top 50 genes with the highest correlation between spatial specificity for melanoma and spatial
607 expression (**Fig 8f**). Genome-wide significant SNPs associated with genes having the highest spatial
608 specificity for T cells or melanoma are mostly SNPs with the strongest significant values in the
609 Manhattan plot. This suggests that the heritability of melanoma cancer risk may be exerted from effects
610 on T cells and melanomas (**Fig 8f**). Overall, our integrative analyses with genome-wide association
611 studies suggest consistent and generalizable patterns.
612

613 **The skincanceratlas database allows users to browse gene expression and LR data from three 614 omics technologies**

615 We have created a comprehensive, interactive database called sklInteractive that allows users to explore
616 our high-throughput single-cell and spatial data atlas and interactome (**Fig S1**). The Atlas section of the
617 database shows cell type clustering and annotation results from scRNASeq, Visium and CosMX data.
618 The Gene Explorer section allows the user to browse genes and LR pairs at single-cell and/or spatial
619 resolution in select samples from cSCC, BCC and melanoma patients. No coding or data downloads
620 are required, making the sklInteractive database an accessible and user-friendly way to browse this
621 resource. The sklInteractive database can be accessed at <https://skincanceratlas.com>.
622

623 Discussion

624 We present cell types, gene signatures, and differential interactome of the three major skin cancer types,
625 which collectively comprise as high as 70% of all cancers as in European ancestry populations. Despite
626 the dominant prevalence, little is known about cell type specific similarities and differences at single cell
627 resolution and in a spatial context that underly patho-etiology of these cancer types. We provide the
628 most comprehensive single cell datasets and the first spatial multiomics reference datasets of cSCC,
629 BCC and melanomas, making this resource publicly available. Through multiomic integration pipeline,
630 incorporating spatial distance and orthogonal multiplatform/multimodality validations, followed by
631 various experimental validation approaches, we built high-confidence interaction networks,
632 underpinning differential initiation and progression of each skin cancer type, which is not well understood
633 (Feller et al., 2016; Thieu et al., 2013; Kim et al., 2013; Wang et al., 2016). Cross-validated results were
634 drawn from comparing and integrating 12 orthogonal technologies, complementary in resolution,
635 sensitivity, and throughput, adding spatial information and quantifying both RNA, protein and glycan
636 modalities.

637

638 Strong results were drawn from comparing and integrating 12 orthogonal technologies, complementary
639 in resolution, sensitivity, and throughput, adding spatial information and quantifying both RNA, protein
640 and glycan modalities. Beyond applying for skin cancer, this work can be a comprehensive guideline for
641 other spatial multiomics studies to assess benefits and limitations for each tool and design a strategy to
642 select suitable tools and combine analyses to cross-validate and gain more biological information than
643 is attainable through any single technology. For example, the large number of genes per cell and the
644 complete RNA extraction from the intact cells/nuclei in scRNA allows for fine-grained classification of
645 cell types. However, the spatial neighbourhood information is lost in scRNAseq, leading to omission of
646 the geographic cell communities within the tissue microenvironment. Further, the scRNAseq-based
647 inference of cell-cell interaction is less accurate due to the lack of neighbourhood constraint (Pham et
648 al., 2023). Although not at single-cell resolution, Visium data enable the inference of all ligand and
649 receptor genes taking into account spatial neighbourhood context. However, as single cell resolution
650 information is not in Visium data, the inference of cell-type specific interactions is limited. This limitation
651 may be partially addressed via cell-type deconvolution, or completely addressed by single-cell panel
652 based platforms, which lack coverage. Regarding proteomics, either scRNAseq or spatial
653 transcriptomics are limited in mapping traditional immune subtypes such as distinguishing CD8 T cells
654 and CD4 T cells, which can be measured by methods such as CODEX and Vectra Polaris. Both
655 scRNAseq, spatial transcriptomics, and CODEX, on the other hand, do not measure metabolites, an
656 important modality that needs methods such as spatial metabolomics to profile.

657

658 Harnessing the complementary information discussed above, we devised spatial integrative analysis for
659 1) cell type annotation, 2) differential expression analysis, 3) ligand-receptor based interactions, 4)
660 community/neighbourhood composition and interaction analysis, 5) multimodal validation of cell-cell
661 interactions, and 6) mapping of genetic association signals to spatially defined regions.

662

663 For cell-cell interaction analysis, we integrated scRNAseq and Visium to map ligand-receptor based
664 interactions at transcriptome-wide level. In this combination, the scRNAseq data also informed Visium
665 data in cell type composition, allowing cell-type specific inference of LR interactions using Visium spot-
666 based, deconvolution data to infer both autocrine and paracrine interaction (Pham et al., 2023). Single-
667 cell resolution spatial transcriptomics data from Xenium and CosMX platforms was then used for
668 validating the LR pairs present in the panel. The interaction results extended beyond LR pairs, providing
669 information about cell type pairs that are interacting. In this way, we could compare pairs of cells that
670 interacted stronger or weaker in each cancer type or in cancer compared to non-cancer samples. Our
671 MMCCI method enables such statistical comparisons (Hockey et al., 2024).

673
674 We found that EMT was the main pathway that was different between melanoma and cSCC-BCC, with
675 the stronger interaction between fibroblasts and melanocytes in melanoma samples and more
676 interaction of fibroblasts with T cells in BCC-cSCC samples. Further, the roles of activated cancer-
677 associated fibroblasts (CAF) and fibroblasts that have undergone epithelial-mesenchymal transition in
678 the cancer-stroma microenvironment are also crucial in BCC and cSCC progression ²⁹. In addition, for
679 colocalization-based interactions using CODEX and Xenium data, we found the complex composition
680 of cancer cells with fibroblasts and T cells within the melanoma community. Interactions with fibroblasts
681 have been reported as essential for melanoma progression (Flach et al., 2011; Kim et al., 2013; Ayuso
682 et al., 2021). Given that fibroblasts are essential in skin cancer initiation and progression (Wang et al.,
683 2012; Werner et al., 2007; Van Hove et al., 2022), the differences in fibroblast interactions between the
684 three cancer types may help explain the differences in metastatic potential between skin cancers.
685 Increasing evidence suggests that interactions between mutated melanocytes and fibroblasts lead to
686 melanoma initiation and progression ²²⁻²⁴. Fibroblasts are most prominent in the dermis, and there is
687 growing evidence that they, together with KCs, are key regulators of skin cancer initiation and
688 progression (Flach et al., 2011; Kim et al., 2013; Ayuso et al., 2021). Indeed, in melanoma, it is clear
689 that cellular interaction with the stroma is one of the key factors driving cancer initiation and progression
690 (Kim et al., 2013; Wang et al., 2016). During embryonic development, melanocytes migrate from the
691 neural crest to the basal layer of the epidermis (Rawles et al., 1974), where their morphology, growth
692 and development are largely regulated by surrounding KCs (Hirobe et al., 2005). In fully developed skin
693 under normal conditions, melanocyte-KC interactions persist, resulting in pigmentation that protects
694 against UV radiation. However, under pathological conditions, dysplastic melanocytes shed interactions
695 with regulatory KCs, resulting in uncontrolled proliferation and gain of stem cell self-renewal capacity
696 (Haass et al., 2005). Therefore, a deeper understanding of interactions between cell types in both
697 melanoma and NMSC will be necessary to develop a more comprehensive paradigm of skin cancer
698 treatment.

699
700 For cell type annotation, our integrative analysis provided confidence for multi-level cell type
701 classification. Through integrative multi-omics analysis, we hierarchically defined 11 major cell types
702 and 30 fine-grained cell (sub-)types across the three studied skin cancers. These cell types were first
703 identified using scRNASeq data. We extensively validated the classification of less-abundant cell types
704 using spatial proteomics data from GeoMx and Polaris platforms. Moreover, annotation from scRNASeq
705 data was cross-validated by mapping them to known anatomical organisation of the skin. The presence
706 of rarer cell types at level 2 and level 3 annotation were validated with GeoMX protein data. Importantly,
707 we also validated the cell types through independent spatial mapping with Visium and CosMx. While the
708 separation of certain clusters (e.g. KC populations) from our scRNASeq atlas is not apparent in UMAP
709 space, these cell types were clearly and distinctly resolved when mapped to their spatial location in skin
710 from the same patients captured by both Visium and CosMx, for example the three layers of KC basal,
711 KC differentiating and KC cornified. This strongly suggests that the annotation of these distinct KC
712 subtypes was accurate, demonstrating the power of using complementary single-cell and spatial data
713 to validate cell subtypes. This spatial mapping also defined visible micro-structures of the skin, like hair
714 follicles and sebaceous glands, and more importantly mapped the single-cell resolution heterogeneity
715 within each of these structures. With our CosMx single-cell mapping, we were also able to perform
716 spatial community analysis, which revealed the increased cellular heterogeneity in melanoma compared
717 to cSCC and BCC. To our knowledge, this quantitative difference has not been reported previously. We
718 observed a higher proportion of Differentiating KCs in cSCC than in the two other cancer types across
719 biological replicates (**Fig 2b**) (Gandarillas et al., 2014), consistent with the defining feature of KC
720 hyperplasia in cSCC. Additionally, we found increased cornification in cSCC compared to healthy tissue,

721 again corroborating the high cell turnover and hyperkeratosis observed in cSCC. Further validation is
722 required to test these data-driven hypotheses.

723
724 For DE analysis, our multiomics approach strengthens the most common type of analysis that could be
725 integrated using most modalities/platforms. Here we combined cell-type specific DE analysis results
726 from scRNAseq data with that of CosMX, Xenium and Visium data. Important markers for KC cancer
727 and melanomas that were found consistently changed across platforms will be highly-confident
728 candidates for further experimental perturbation studies. These include SOX2, LAMP3, CXCL10, CCL5,
729 and UBE2C. Differences in LR interactions may also explain why one form of skin cancer may arise
730 over another type. For example, PTCH1 (Patched1) is defective in 70-85% of BCC, but not in cSCC
731 (Boukamp et al., 2005; Bonilla et al., 2016). The absence of ligands for the PTCH1 membrane receptor
732 in BCC leads to tumour formation under the control of the transcriptional factor GLI1 (Boukamp et al.,
733 2005). Therefore, loss of PTCH1 signalling may predispose a cell towards initiating BCC over cSCC.
734 Improved insights into the molecular crosstalk between cell types would help elucidate the molecular
735 events underlying the initiation of one cancer type over another.

736
737 Our atlas also enables comparison of inter- and intra-tumour heterogeneity across patient. In BCC-
738 cSCC, comparing 11 healthy and cancer samples with scRNASeq, we found an enrichment of the
739 immune response with CD4 and CD8 T cells, M1 and M2 macrophages, NK and classical cDC cells
740 consistently higher in cancer samples across all five patients. At the gene level, we observed a global
741 upregulation in cancer cells of genes related to progression and invasion like *S100A7* and *KRT6B* (Chen
742 et al., 2019; Chang et al., 2011) in our scRNASeq data. When accounting for interpatient heterogeneity,
743 we found 39 genes that were upregulated across the entire cSCC dataset, including genes associated
744 with immune processes such as antigen presentation, interferon-gamma response, viral response and
745 cell killing. Comparisons between cancer and healthy regions of each sample identified a consistent
746 shift towards fibroblast-based signalling in cancer. While the critical roles of fibroblasts in regulating skin
747 cancer initiation and progression are well documented (Flach et al., 2011; Ayuso et al., 2021; Kim et al.,
748 2013), a comprehensive map of fibroblast interactions specific to partner cell types and spatial locations,
749 as shown in this study, is still lacking.

750
751 Spatial cell community analyses provides new understanding of spatial patterns associated with cancer
752 biology. Here we devised an approach to find a robust community by integrating three platforms, CosMX,
753 Xenium and Visium. We were able to map a melanoma community across CODEX, Xenium and
754 Glycomics modalities and perform, for the first time, joint analysis of spatial omics data for this
755 community, showing the multimodal evidence for the upregulation of tyrosine and pyrimidine metabolism
756 pathways. The community-based colocalization analysis added more evidence for the interaction of
757 melanoma/melanocytes with regulatory T cells and fibroblast. The ligand-receptor based analysis of
758 the melanocyte-enriched community suggested much more active interaction in melanoma samples
759 compared to in BCC, specifically with strong interactions involving collagen with CD44 and Integrins.
760 Inflammatory and Mesenchymal fibroblast cells and T cells create a protumorigenic microenvironment
761 that may be associated with survival (Schütz et al., 2023; Zhang et al., 2022).

762
763 Multimodal validation of ligand-receptor interactions provided strong evidence for the LR pairs and cell
764 types involved in the interactions at RNA and protein level at local or nano-scale distance. Our first line
765 of cross-validation for cellular interactions was via finding consistent interactions between platforms,
766 where the broad discovery of interactions using scRNAseq and Visium can be visualised at a single cell
767 resolution proving the colocalisation of ligand and receptor signals between two cells in the Xenium and
768 CosMX. The validation at RNA level was then strengthened with targeted hybridization using RNAscope
769 technology with signal amplification chemistry, allowing us to visualise at high sensitivity the co-

770 expression of IL34-CSF1R. While transcriptomics data allows for screening many more ligand-receptor
771 interactions than can be achieved through proteomics, the RNA-based approaches remain as an
772 inference test, but not a direct proof of protein-protein interactions. Beyond the RNA modality, we
773 extended the validation to protein level, with Opal TSA chemistry that allows for highly sensitive
774 detection of protein expression. The Opal system enabled us to validate ligand-receptor with established
775 antibodies such as PD1-PDL1. However, the colocalisation with Opal Polaris approach lacks the
776 resolution to find exact interactions. We, therefore, used PLA assay to detect interactions within 20 nm
777 distance, specifically mapping the interactions to the cell membrane.

778

779 Our final integration type in this work involved the mapping of genetics association signals to spatial cell
780 types and tissue domains. This approach integrates population-scale information to relatively small, but
781 deeply-profiled, functional datasets using spatial multiomics platforms. With this new approach, we
782 successfully mapped significant SNPs to gene expression within tissues and computed cumulative
783 association signals for a spot, a spatial domain or a cell type. We showed the genetic association of
784 melanocytes for the melanoma trait and of KC dysplastic and KC cornified for cSCC and BCC. This
785 analysis demonstrates that while each of the skin cancers arises from a specific type of cell, their shared
786 and independent risk genes and SNPs act across a range of skin cells. The result suggests that one
787 may need to consider that at least some risk loci may be mediated by cis-regulation in keratinocytes,
788 which are involved in tightly controlling melanocyte proliferation and invasion. The analysis also provided
789 statistical evidence on the roles of genetics association in ligand-receptor interaction, such as the *IL34-*
790 *CSF1R* pair.

791 Together, by using a spatial multi-omics approach with 12 independent technologies, our study
792 represents the first comprehensive comparison of spatial cellular signatures across the three skin major
793 cancer types, BCC, cSCC and melanoma. Although the number of samples is small (total 24 patients),
794 the cross-validation by independent experimental platforms provides high-confidence results for the
795 individual measured. Some of these results add consistent evidence to the existing literature, while new
796 findings from the cohort studied here would require future validation to external cohorts. Nevertheless,
797 our integrative analysis with population genetics studies of >300,000 individuals suggest consistent
798 patterns and indicate highly generalizable results.

799 We identified both shared and distinct cellular and gene signatures for each of the three skin cancer
800 types, suggesting important cell-type specific pathways underlying differences in the initiation and
801 progression of these cancer types. The interacting cell types and LR pairs identified here represent
802 promising therapeutic targets for skin cancer treatment, including immunotherapies. The highly
803 integrated spatial multi omics dataset is available through our skin cancer website
804 (<https://skincanceratlas.com/>), which is accessible to the broader research community for visualisation
805 and analysis without requiring coding. The data would be useful in multiple scenarios, for example, to
806 provide new insights into the roles of pathways where DNA mutations have been reported, but little is
807 known about how these mutations manifest at single cell and spatial levels, including canonical markers
808 such as MAPK activation in Melanoma, Hedgehog signalling in BCC, and NOTCH/p53 signalling in
809 cSCC.

810 Methods

811 Patient material and ethics

812 All samples (**Table S1**) were collected with informed patient consent and approved for research use
813 under ethics approval numbers 2018000165 and 2017000318 by the University of Queensland's Human
814 Research Ethics Committees and 11QPAH477 by the Metro South Human Research Ethics Committee.

815 All formalin-fixed, paraffin-embedded (FFPE) blocks were previously prepared following a standard
816 fixation procedure in 10% formalin, processed in ethanol and xylene and embedded in paraffin wax. The
817 four melanoma samples (patients 6747-085P, 21031-08TB, 48974-2B, 66487-1A) were collected during
818 2008-2018 and all blocks were stored at room temperature. Full patient IDs are abbreviated as 6747,
819 21031, 48974 and 66487 throughout this manuscript.

820 For BCC and cSCC samples from eight patients (B18, E15, F21, P30, P13, P04, R01, D12), all fresh-
821 shaved biopsies were obtained in accordance with the approved ethics protocol (11QPAH477). Patients
822 presented at the Princess Alexandra Hospital Dermatology Department between October 2018 and
823 February 2020. Among these patients, three patients (B18, E15, F21) were diagnosed with both cSCC
824 and BCC and biopsies of both cancer types were collected. Five patients (B18, P30, P13, P04, R01)
825 kindly consented to participate in the collection of 4 mm punch biopsy samples of non-sun exposed non-
826 cancer skin for paired scRNA sequencing experiments. Lesion identity was confirmed by pathological
827 inspection. Portions of each sample from these patients were also preserved with 10% formalin as
828 described for FFPE samples above. To process fresh samples for scRNA sequencing, briefly, fresh-
829 shaved biopsies were collected in DMEM for immediate tissue dissociation. Tissue was incubated in 10
830 mg/mL Dispase II (cat. No. 04942078001, Roche, Darmstadt, Germany) for 45 min at 37°C, snipped
831 into small pieces with scissors, and incubated in 0.25% Trypsin for 2 min. The cells were disrupted
832 gently with a pipette and filtered through 70 µm and 40 µm cell strainers, taken up in culture medium
833 and spun down at 350 rcf. Resuspended cells were collected in PBS containing Foetal Calf Serum for
834 single-cell sequencing.

835 scRNASeq was performed on 14 samples representing both cancer and non-cancer cSCC and
836 melanocyte lesions. Healthy and cSCC biopsies were paired, from patients B18, P30, P13, P04 and
837 R01 (**Table S1**). The cancer biopsy from patient P13 was identified as being intra-epidermal carcinoma
838 (IEC), also known as Bowen's disease, a more superficial subtype of cSCC which occurs in the upper
839 epidermal layer. Patient B18 was diagnosed with both cSCC and BCC, and tissue from both cancer
840 lesions was pooled prior to library preparation. Patients E15 and F21 were also diagnosed with both
841 cSCC and BCC, but only cSCC tissue was used for scRNASeq. Two separate samples were collected
842 for P30; data were pooled after sequencing. Melanoma used for CODEX, Xenium and spatial glycomics
843 were archival samples from a retrospective patient group with thin melanomas (Stage I, Breslow depth
844 <1mm). For snRNASeq, three archived melanoma samples representing three diagnosis types (by 23
845 pathologists), including malignant, intermediate (dysplastic) and benign melanocyte lesions (MPS13,
846 MPS42, MPS43). Skin samples from healthy volunteer donors aged from 25-45 without skin cancer
847 were collected from the forearm of the donors and were preserved in FFPE format.

848 Data generation, pre-processing and cell type annotation methods are described in detail for all
849 technologies in the **Supplementary Methods**.

850 **scRNASeq data analysis**

851 scRNASeq data was generated, processed, integrated and annotated as described in the
852 **Supplementary Methods**. LR analysis for scRNASeq data was performed using CellChat (Jin et al.,
853 2021), using normalised gene expression for all patients and all genes as input. Analysis was performed
854 as per the detailed CellChat vignette. Circos plots were generated using the R package circlize (Gu et
855 al., 2014). Significant LR pairs present in ≥3 samples were visualised in a heatmap using
856 ComplexHeatmap (Gu et al., 2016). GO analysis was performed as described above for the core gene
857 suite analysis. LR pairs were split into their composite genes prior to analysis.

858 Cancerous KC cells in cSCC samples were identified based on two intersecting criteria. Copy number
859 variation analysis was performed using both InferCNV (Tickle et al., 2019) and CopyKat (Gao et al.,

860 2021) using default parameters. We make use of CopyKAT's ability to predict 'Aneuploid' cells and
861 InferCNV's de-noising and QC filtering approach to retain only the cells that are likely to be 'Aneuploid'.
862 Candidate KC cancer cells passed the first round of filtering if they were predicted to be aneuploid by
863 both tools. Next, genes differentially expressed between KC cells from cancer and healthy biopsies were
864 identified using either edgeR (Robinson et al., 2010) or Scanpy (Wolf et al., 2018) and used to calculate
865 an "cSCC score" using a custom python script equivalent to Seurat's AddModuleScore function. Cells
866 receiving an cSCC score in the ≥ 95 th percentile of all scores for both methods passed the second round
867 of filtering. Therefore, KC cancer cells were those found to be both abnormal in ploidy and enriched for
868 genes associated with cancer biopsies. Melanoma cells were annotated in a similar way, with the
869 module score calculated using DE genes between the malignant melanoma sample and the benign
870 ones in the similar manner and with an ≥ 80 th percentile threshold used for the second step of filtering
871 based on the observed module score distribution and the number of cells.

872 Visium data analysis

873 Data generation, processing and integration, plus cell type deconvolution and CCI analysis were
874 performed described in the **Supplementary Methods**. For the IL34-CSF1R analysis, raw LR scores
875 from stLearn were used to classify spots as either IL34_CSF1R-positive (i.e. with a LR score > 0 for this
876 pair) or -negative (i.e. with a LR score = 0). Integrated Visium Seurat objects for each cancer type in
877 turn were used to perform differential gene expression analysis using Seurat's FindMarkers function
878 (Wilcoxon test with parameters min.pct = 0.25, logfc.threshold = 0.25, adjusted p-value threshold ≤ 0.05)
879 to compare gene expression between positive and negative spots. GO enrichment analysis was
880 performed as described above. GO terms associated with upregulated genes in melanoma were split
881 into functionally-related groups (**Fig 8f-g**) by calculating pairwise semantic similarity values between GO
882 terms using GOSemSim (Yu et al., 2010). K-means clustering ($k = 3$) was used to cluster the resulting
883 semantic similarity values into three groups of related GO terms. Genes associated with each GO term
884 in each group were plotted using Complex Heatmap.

885 CosMx data analysis

886 CosMx data was generated, processed, integrated and annotated as described in the **Supplementary**
887 **Methods**. We first analysed CCI within individual FOVs. We used our SCTP method in stLearn (Pham
888 et al., 2023) for CCI prediction, because this tool incorporates information about LR pairs, cell types,
889 and physical distances, thus maximising data usage and providing spatially-meaningful (and therefore
890 more biologically-meaningful) results. Briefly, for a given cell, SCTP defines a neighbourhood as the set
891 of cells within a predefined spatial distance of that cell. For each LR pair and each cell in an FOV in turn,
892 LRscores are calculated as the sum of the mean ligand expression and the mean receptor expression
893 across a given cells' neighbourhood. The LRscore is further corrected by neighbourhood cell type
894 diversity, which is known to positively correlate with the likelihood of CCI (Rieckmann et al., 2017; Hou
895 et al., 2020). stLearn uses a permutation test to determine the null distribution of LR scores for
896 hypothesis testing. It defines significant cells and LR pairs. We performed cell type-specific CCI analysis
897 to examine significant LR interactions between pairs of cell types, using the outputs from the cell level
898 analysis described above. Briefly, SCTP generates a CCI_{LR} matrix by counting the number of cells with
899 significant LRscore signalling from one cell type to another for a given LR pair. Like the cell level
900 analysis, SCTP uses a permutation analysis to test whether these counts are significantly different from
901 random.

902 To quantify the spatial heterogeneity of each FOV, we first constructed cell-cell neighbourhood networks
903 by applying Delaunay triangulation to cell spatial coordinates, resulting in one network per FOV. Next,
904 we applied Rao's quadratic entropy to each cell in each network to measure the cell type heterogeneity.
905 We elected to use Rao's quadratic entropy scoring for this purpose because it can consider both the

probability of two neighbouring cells (i.e. two cells sharing an edge) being different cell types, and the spatial distance between each member of the neighbouring pair. As the natural entropy score is often used as a measure for connected graphs, we fed the customised Delaunay cell neighbourhood network through the ATHENA local quadratic scores function (Martinelli et al., 2002). For cross-cancer type comparison, we aggregated the entropy scores of cells from the same FOVs and grouped the FOVs by cancer subtype. The entropy score was calculated as a product of the spatial distance of cells and the cell type probability; these remain as constant units across FOVs, so normalisation is not required. We performed pairwise comparison of the distribution of entropy scores from each cancer subtype using a Wilcoxon rank sum test.

Xenium data analysis

Formalin-fixed paraffin-embedded (FFPE) tissue blocks were sectioned at a thickness of 5 µm and mounted onto Xenium slides, in accordance with the FFPE Tissue Preparation Guide (10x Genomics, CG000578, Rev B). In situ hybridisation was carried out overnight using 260 probes from the pre-designed Xenium Human Skin Panel (10x Genomics). DAPI staining was used to label nuclei, which were used for the estimation of cell boundaries (10x Genomics, CG000582, Rev D). Following completion of the run, H&E staining was conducted on the same tissue region. Each Xenium sample was preprocessed individually using Seurat version 5.0. During the quality control (QC) step, cells with zero expression across all genes were filtered out. Normalisation was performed using the *SCTransform* function, followed by principal component analysis (PCA) using the top 30 principal components. To annotate cell types for the Xenium cells, Seurat's label transfer workflow was employed. The melanoma single-cell RNA-seq dataset processed in the previous section was used as the reference. Anchor points between the reference and Xenium datasets were identified using the *FindTransferAnchors* function. Cell type annotations were then transferred to the Xenium data using the *TransferData* function, applying the level 2 annotation labels from the reference dataset.

CODEX data analysis

Cell segmentation for CODEX QPTIFF data was done using Cellpose as an implementation function in the Sopa package. Signal intensity for each protein channel was then mapped to the Cellpose boundaries. Outlier cells with data lower than 0.05 quantile or higher than 0.95 quantile were removed from the raw protein expression intensity matrices. The data was transformed with arcsinh and scaled to mean 0 and standard deviation 1. The cell type identification was performed based on the protein markers included in the panel using z-scores. For spatial community analysis (niche detection), we used two methods, NeighbourhoodCoordination and MonkeyBread neighbourhood clustering. Both methods clustered cells based on the cell type proportion of a neighbourhood tissue area as squared tiles (windows) or a circle of a given radius. Colocalization between cells of two cell types was computed based on distance. A network connecting cell types and communities was drawn using the network approach in the Sopa package.

Spatial Glycomics data analysis

Following general pre-processing, data from three MALDI samples were individually analysed using the R-based package *SpaMTP V1.0* (Causer et al., 2024). Mass peaks were initially binned at a resolution of 250 ppm, resulting in 5433 detectable m/z peaks. Samples were annotated against the Lipid Maps database implemented in *SpaMTP*, with the *AnnotateSM* function. Principal component analysis was run and dimensionality reduction was performed for each sample using the first 30 principal components. Louvain clustering was implemented using a resolution of 0.3, resulting in samples containing between 9 and 13 clusters. Pseudo-bulking differential metabolite abundance analysis was performed per cluster

956 using the *SpaMTP FindAllDEMs* function. The top 10 m/z values per cluster, per sample, were then
957 combined and hierarchical clustering was implemented to group similar clusters together based on
958 pseudo-bulked expression. Melanoma clusters were identified based on spatial location and confirmed
959 by hierarchical grouping. Differential abundance analysis was again performed to identify all significantly
960 abundant metabolites within the melanoma cluster compared to all other clusters of each sample.
961 Common metabolites that were differentially abundant within the melanoma cluster, across all three
962 samples, were then identified and spatially plotted across each tissue sample.

963 **Integrative Cell Neighbourhood/Community analysis**

964 For each data type, Visium, Xenium, and CosMX, we applied the NeighbourhoodCoordination method
965 to map communities of nearby cells that had similar neighbourhoods as assessed based on cell type
966 composition (Schürch et al., 2020). A neighbourhood matrix, where the proportion of cell types within
967 a neighbourhood (window) were calculated for each cell in each Visium sample, or Xenium sample or
968 CosMX FOV using the same setting. The windows for all samples and FOV within one technology (e.g.,
969 from all Visium samples) were merged into one matrix per technology. The windows were then clustered
970 using K-mean clustering, with K=10 for all samples, for consistency.

971 To find robust communities, we combined together 30 communities, representing Visium, Xenium, and
972 CosMX. Each community was represented by the proportion of cell types within the community. The
973 matrix of combined communities was used to group similar communities into functional categories
974 (considered as meta communities) such as tumour, stromal, immune or KC. This way, communities
975 across cancer types and platforms can be compared. For example, we compared the cancer community
976 (e.g., CosMX_6) in cell-cell interactions across BCC, cSCC and Melanoma samples. We also compared
977 the cancer heterogeneity at community level (e.g., CosMX_6 and Visium_2 for the cancer sample),
978 where each community may be more or less abundant in one cancer type compared to the other cancer
979 types.

980 **Integrative Cell-Cell interaction analysis**

981 We implemented our MMCCI cell cell interaction analysis pipeline to integrate data from multiple
982 samples and multiple platforms. MMCCI takes inputs as CCI results from individual samples calculated
983 by spatially-aware interaction scores and P-values from stLearn. Prior to integration, the interaction
984 scores were normalised to take into account differences in the number of cells/spots across samples.
985 The integration process resulted in two main outputs, the strength of interactions between two cell types
986 (total number of interacting cells) and the integrated P-value for the interaction (using Stouffer's method
987 to calculate the inverse cumulative distribution of all P-values).

988 **Integrative DE and pathway analysis across four platforms scRNAseq, Visium, CosMX and 989 Xenium**

990 Pseudobulking following EdgeR DE analysis pipeline with quasi likelihood ratio test were applied across
991 platforms. Shared DE genes, consistently found in all orthogonal spatial technologies, were highly
992 confident DE genes that can be considered as promising markers of a cancer type or a biological
993 pathway differentially regulated among cancer types. For each modality, significant abundant analysis
994 was performed between clusters/spatial communities. Using cell type annotation, a spatial
995 community/neighbourhood commonly found across modalities can be jointly analysed. The differential
996 markers (genes, proteins or metabolites) derived from comparing one community with the remaining
997 other communities were input into MetaboAnalyst for joint pathway analysis.

998 **Spatial multiomic validation with Proximity Ligation Assay (PLA), Opal Polaris and RNA scope**

We assessed multiple approaches to validate ligand-receptor interactions, including RNAscope, Opal Polaris and Proximal Ligation Assay for validating cell-cell interactions. In the case of IL34-CSF1R, antibodies for IL34 were not readily optimised and so we applied RNAscope to examine the colocalization at single molecule and single cell resolution (**Fig 7**). Similar to Xenium or CosMX assay in resolution, the RNAscope produces an additional advantage with the Z-probe amplifier chemistry leading to a high detection sensitivity. Our data show consistent results between RNAscope and CosMX data that prove the colocalization within neighbouring cells of IL34 and CSF1R mRNA (**Fig S22 and Fig 7c, d**). Using spatial transcriptomics data, we can also validate downstream pathways that change specifically associated with the LR pairs co-expressed in the spatial spots. For example, the functional downstream consequences of IL34_CSF1R signalling were identified based on genes that were differentially expressed between IL34_CSF1R-positive (LR score >0) and -negative (LR score of 0) spots for each cancer type (**Fig 7e-g**).

At the protein level, we detected colocalization of the protein pairs with Vectra Polaris and Proximal Ligation Assay. The PLA was performed to validate ligand-receptor interactions identified through spatial transcriptomics data. FFPE melanoma tissue sections were deparaffinized, rehydrated, and subjected to antigen retrieval. Primary antibodies specific to the ligand and receptor of interest were applied and incubated overnight at 4°C. Following the manufacturer's instructions for the NaveniFlex Tissue MR Atto647N kit (Navinci). Fluorescent signals indicative of close-proximity interactions were generated through ligation and amplification steps. After PLA single is generated, the tissue sections were stained with anti-mouse/rabbit secondary antibodies conjugated to Alexa Fluorophore for 1 hour at room temperature for visualisation of target proteins. Subsequently, tissues were counterstained with DAPI, and imaged using the STELLARIS Confocal Microscope (Leica). Our PLA results clearly showed the specific signals on the cell membrane in the positive control (E-cadherin and b-Catenin) and no signal in the negative control (CD31-AQP1), (**Fig S22**).

Integrating spatial transcriptomics with GWAS data

We used the summary statistics from a cSCC GWAS study of 10557 controls and 537850 controls and for BCC GWAS with 36479 cases and 540185 controls (Seviiri et al., 2022) and Melanoma GWAS with 30,143 clinically-confirmed melanoma cases and 81,405 controls (Landi et al., 2020). We applied gsMAP method to map genetic signals to spatial gene expression in the three skin cancer datasets (Song et al., 2024). First, tissue domains were determined by finding similar spots/cells using a graph attention autoencoder network which generated latent representations, which in turn were used to find pairwise cosine similarity between spots/cells. Next, gene specificity scores (GSS) for each spot were computed by aggregating information between similar spots/cells (domain), and rank enrichment information for top abundant genes from its homogeneous spots. The expression specificity of a gene within a focal spot was assessed by calculating the geometric mean of its expression rank across the tissue region (microdomain identified by graph attention, or cell type) of the focal spot, divided by the geometric mean of its expression rank across all spots in the ST data. Genes with a ratio higher than 1 and expressed in more spots/cells in the region/cell-type than in overall sample(s) were considered specific for the region or cell type. A high GSS score for a gene suggests that the gene was higher/enriched for the region/cell-type than most other genes in that region/cell-type.

Based on proximity to the nearest transcription start sites, GWAS SNP are assigned to GSS genes for each focal spot/cell. Given the set of assigned SNPs to each spot/cell, the SNP effects in the GWAS summary statistics for BCC, cSCC, and cutaneous melanoma and LD scores (from the 1000 Genomes Project Phase 3) were used for LD score regression analysis. Given the total set of SNPs assigned to a spot/cell, the SNP-level trait heritability (Chi-square association with the skin cancer trait of a SNP from the summary statistics) is partitioned into the SNP effect of a focal spot or cell and the effect of the SNP given the baseline SNPs that are not assigned to GSS genes. The enrichment p-value

1050 is calculated based on the partitioned regression coefficient, using one-sided Z-test for bigger than 0.
1051 To compute p-value for the association of a spatial region or a cell type across the whole sample, we
1052 aggregated P values of individual spots/cells within the spatial regions (or cell types) using Cauchy
1053 combination.

1054 **skInteractive database**

1055 We built the skInteractive database in the form of a visualisation dashboard with a Shiny v1.7.2
1056 application (Sievert et al., 2020). The database has two main sections, the Atlas, which shows cell types
1057 and clustering results, and the Gene Explorer, which allows the user to browse gene expression and/or
1058 LR interaction scores for the different datasets and modalities. The Atlas dashboard was constructed
1059 using Javascript (NuxtJS framework). We converted all plots to geo map components in Apache Echarts
1060 v5.3.1 (Li et al., 2018) that provided interactive features to work with the plots. For the Gene Explorer
1061 Shiny app, we implemented multiple tabs themes and used Seurat v4.1.1 (Hao et al., 2021) to generate
1062 plots from the different data sets stored in SeuratObjects. The Shiny application included Visium (gene
1063 expression and LR scores), CosMx (gene expression) and scRNAseq data (gene expression).
1064 skInteractive Database will be ported to AWS cloud.

1065 **Data availability**

1066 All of the sequencing data and accompanying H&E images for spatial transcriptomics both raw and
1067 processed will be deposited to ArrayExpress repository (<https://www.ebi.ac.uk/arrayexpress/>) and made
1068 publicly available according to human ethics regulations. All other experimental data (e.g. imaging data
1069 using RNAscope or by Polaris immunofluorescence) will be made available upon request. GWAS data
1070 used in this analysis is available as detailed in the relevant publications (Landi et al., 2020; Seviiri et al.,
1071 2022)

1072 **Code availability**

1073 The code to reproduce analyses and figures presented in this paper is available at
1074 <https://github.com/GenomicsMachineLearning/SkinCancerAtlas/tree/main>

1075
1076 The public Github repository of the skInteractive database can be accessed at
1077 https://github.com/BiomedicalMachineLearning/SkinCancerAtlas/tree/main/manuscript_code/skInteractiveDatabase and users can explore the interactive functionalities without the need for customised
1078 coding at <https://skincanceratlas.com/>
1079

1080 **Author contributions statement**

1081 Q.N. conceived the project. Key contributors to data analysis include L.G. (scRNAseq and Visium
1082 data), P.P. (scRNAseq), F.Z. and G.N. and O.M (CosMx data), X.J. (scRNAseq and Visium data), M.T.
1083 (Polaris, RNAscope, and CosMx data), O.M. (GeoMx data), E.E.K (CosMx data), M.T.G. (CosMx
1084 data), Z.R. (CosMx data). D.P., L.G., P.P. and A.N. built the skInteractive web tool. A.X., S.M.T, T.V.,
1085 K.D., L.P., A.K., M.L., S.R.M., S.E.W., Y.K. conducted the experiments and generated the data. M.Z.,
1086 H.S., J.G.C, S.K., H.N.L., H.P.S., M.T.L, K.M.B, M.M.I, M.S, M.H.L, I.F., Y.K., M.S.S., K.K., Q.N.
1087 contributed to data analysis and interpretation. P.S., M.S.S., K.K. and I.F. provided clinical samples.
1088 All authors contributed to writing the manuscript and have reviewed and approved the manuscript.

1089 **Acknowledgements**

1090 This work has been supported by the Australian Research Council (ARC DECRA grant DE190100116
1091 to QHN), National Health & Medical Research Council (NHMRC Project Grant 2001514 to QHN),
1092 NHMRC Investigator Grant (GNT2008928 to QHN). We thank the staff at the University of Queensland
1093 sequencing facility and the School of Biomedical Science imaging facility for their support in spatial
1094 transcriptomics and single cell RNA sequencing. We thank all members in Nguyen's Genomics and
1095 Machine Learning Lab for discussion and support in data generation and analysis. Acknowledgements
1096 for the melanoma meta-analysis data can be found in the supplemental data.

1097 **Competing interests statement**

1098 E.E.K., M.T.G., Z.R., L.P., M.L., S.R.M., H.S., S.E.W., and Y.K. are/were employees of Bruker
1099 NanoString Technologies and hold NanoString stock or stock options. The remaining authors have
1100 declared no competing interest.
1101

1102 **References**

- 1103 Anborgh, P. H., J. C. Mutrie, A. B. Tuck and A. F. Chambers (2010). "Role of the metastasis-
1104 promoting protein osteopontin in the tumour microenvironment." *J Cell Mol Med* 14(8): 2037-2044.
- 1105 Ayuso, J. M. et al. Microfluidic model with air-walls reveals fibroblasts and keratinocytes modulate
1106 melanoma cell phenotype, migration, and metabolism. *Lab Chip* 21, 1139–1149 (2021).
- 1107 Bevona, C., Goggins, W., Quinn, T., Fullerton, J. & Tsao, H. Cutaneous melanomas associated with
1108 nevi. *Arch. Dermatol.* 139, 1620–1624 (2003).
- 1109 Boukamp, P. Non-melanoma skin cancer: what drives tumor development and progression?
1110 *Carcinogenesis* 26, 1657–1667 (2005).
- 1111 Baghdadi, M. et al. Chemotherapy-Induced IL34 enhances immunosuppression by tumor-associated
1112 macrophages and mediates survival of chemoresistant lung cancer cells. *Cancer Res.* 76, 6030–6042
1113 (2016).
- 1114 Baghdadi, M. et al. High co-expression of IL-34 and M-CSF correlates with tumor progression and
1115 poor survival in lung cancers. *Sci. Rep.* 8, 418 (2018).
- 1116 Bonilla, X. et al. Genomic analysis identifies new drivers and progression pathways in skin basal cell
1117 carcinoma. *Nat. Genet.* 48, 398–406 (2016).
- 1118 Cable, D. M., E. Murray, L. S. Zou, A. Goeva, E. Z. Macosko, F. Chen and R. A. Irizarry (2022).
1119 "Robust decomposition of cell type mixtures in spatial transcriptomics." *Nature Biotechnology* 40(4):
1120 517-526.
- 1121 Chen, C. & Shan, H. Keratin 6A gene silencing suppresses cell invasion and metastasis of
1122 nasopharyngeal carcinoma via the β-catenin cascade. *Mol. Med. Rep.* 19, 3477–3484 (2019).
- 1123 Chang, H.-H., Dreyfuss, J. M. & Ramoni, M. F. A transcriptional network signature characterizes lung
1124 cancer subtypes. *Cancer* 117, 353 (2011).
- 1125 Candi, E., Knight, R. A. & Melino, G. Cornification of the skin: A non-apoptotic cell death mechanism.
1126 eLS (2009) doi:10.1002/9780470015902.a0021583.
- 1127 Chen, C. & Shan, H. Keratin 6A gene silencing suppresses cell invasion and metastasis of
1128 nasopharyngeal carcinoma via the β-catenin cascade. *Mol. Med. Rep.* 19, 3477–3484 (2019).
- 1129 Czarnecki, D. (2024). "Mortality from Nonmelanoma Skin Cancer in Australia from 1971 to 2021."
1130 *Cancers (Basel)* 16(5).
- 1131 Edwards, L., R. Gupta, and F. V. Filipp. 2016. 'Hypermutation of DPYD Dere-regulates Pyrimidine
1132 Metabolism and Promotes Malignant Progression', *Mol Cancer Res*, 14: 196-206.
- 1133 Eckhart, L., Lippens, S., Tschachler, E. & Declercq, W. Cell death by cornification. *Biochimica et
1134 Biophysica Acta (BBA) - Molecular Cell Research* 1833, 3471–3480 (2013).
- 1135 Feller, L., Khammissa, R. A. G., Kramer, B., Altini, M. & Lemmer, J. Basal cell carcinoma, squamous
1136 cell carcinoma and melanoma of the head and face. *Head & Face Medicine* 12, (2016).
- 1137
- 1138 Fink, M. et al. Comparison of efficacy in patients with metastatic melanoma treated with ipilimumab
1139 and nivolumab who did or did not discontinue treatment due to immune-related adverse events: A real-
1140 world data study. *Cancers* 13, 5550 (2021).

- 1141 Flach, E. H., Rebecca, V. W., Herlyn, M., Smalley, K. S. M. & Anderson, A. R. A. Fibroblasts
1142 contribute to melanoma tumor growth and drug resistance. *Mol. Pharm.* 8, 2039–2049 (2011).
- 1143 Foucher, E. D. et al. IL-34 induces the differentiation of human monocytes into immunosuppressive
1144 macrophages. *Antagonistic effects of GM-CSF and IFNy*. *PLoS One* 8, e56045 (2013).
- 1145 Gao, R., Bai, S., Henderson, Y.C. et al. Delineating copy number and clonal substructure in human
1146 tumors from single-cell transcriptomes. *Nat Biotechnol* 39, 599–608 (2021).
- 1147 Garcia-Sancha, N. et al. Overcoming resistance to immunotherapy in advanced cutaneous squamous
1148 cell carcinoma. *Cancers* 13, (2021).
- 1149 Gandalillas, A. & Freije, A. Cycling up the epidermis: Reconciling 100 years of debate. *Exp. Dermatol.*
1150 23, 87–91 (2014).
- 1151 George, D. D., Armenio, V. A. & Katz, S. C. Combinatorial immunotherapy for melanoma. *Cancer*
1152 *Gene Ther.* 24, 141–147 (2017).
- 1153 Giricz, O., Y. Mo, K. B. Dahlman, X. M. Cotto-Rios, C. Vardabasso, H. Nguyen, B. Matusow, M.
1154 Bartenstein, V. Polishchuk, D. B. Johnson, T. D. Bhagat, R. Shellooe, E. Burton, J. Tsai, C. Zhang, G.
1155 Habets, J. M. Greally, Y. Yu, P. A. Kenny, G. B. Fields, K. Pradhan, E. R. Stanley, E. Bernstein, G.
1156 Bollag, E. Gavathiotis, B. L. West, J. A. Sosman and A. K. Verma (2018). "The RUNX1/IL-34/CSF-1R
1157 axis is an autocrinally regulated modulator of resistance to BRAF-V600E inhibition in melanoma." *JCI*
1158 *Insight* 3(14).
- 1159 Goodson, A. G. & Grossman, D. Strategies for early melanoma detection: Approaches to the patient
1160 with nevi. *Journal of the American Academy of Dermatology* 60, 719–735 (2009).
- 1161 Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize implements and enhances circular
1162 visualization in R. *Bioinformatics* 30, 2811–2812 (2014).
- 1163 Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in
1164 multidimensional genomic data. *Bioinformatics* 32, 2847–2849 (2016).
- 1165 Guy, G. P., Jr, Machlin, S. R., Ekwueme, D. U. & Yabroff, K. R. Prevalence and costs of skin cancer
1166 treatment in the U.S., 2002-2006 and 2007-2011. *Am. J. Prev. Med.* 48, 183–187 (2015).
- 1167 Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* 184, 3573–3587.e29 (2021).
- 1168 Hayashido, Y., H. Kitano, T. Sakaue, T. Fujii, M. Suematsu, S. Sakurai and T. Okamoto (2014).
1169 "Overexpression of integrin av facilitates proliferation and invasion of oral squamous cell carcinoma
1170 cells via MEK/ERK signaling pathway that is activated by interaction of integrin avβ8 with type I
1171 collagen." *Int J Oncol* 45(5): 1875-1882.
- 1172 Hou, R., Denisenko, E., Ong, H. T., Ramilowski, J. A. & Forrest, A. R. R. Predicting cell-to-cell
1173 communication networks using NATMI. *Nat. Commun.* 11, 5011 (2020).
- 1174 Hockey, Levi, Onkar Mulay, Zherui Xiong, Samuel X. Tan, Kiarash Khosrotehrani, Christian M.
1175 Nefzger, and Quan Nguyen. 2024. 'MMCCI: multimodal integrative analysis of single-cell and spatial
1176 cell-type communications to uncover overarching and condition-specific ligand-receptor interaction
1177 pathways', *bioRxiv*: 2024.02.28.582639.
- 1178 He, S. et al. High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial
1179 molecular imaging. *Nat. Biotechnol.* 1–13 (2022).
- 1180 Hirobe, T. Role of keratinocyte-derived factors involved in regulating the proliferation and
1181 differentiation of mammalian epidermal melanocytes. *Pigment Cell Res.* 18, 2–12 (2005).
- 1182 Haass, N. K. & Herlyn, M. Normal human melanocyte homeostasis as a paradigm for understanding

- 1183 melanoma. *J. Investig. Dermatol. Symp. Proc.* 10, 153–163 (2005).
- 1184 Herrscher, H. & Robert, C. Immune checkpoint inhibitors in melanoma in the metastatic, neoadjuvant,
1185 and adjuvant setting. *Curr. Opin. Oncol.* 32, 106–113 (2020).
- 1186 Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. *Nature Communications*
1187 12, (2021).
- 1188 Kim, E. et al. Senescent fibroblasts in melanoma initiation and progression: an integrated theoretical,
1189 experimental, and clinical approach. *Cancer Res.* 73, 6874–6885 (2013).
- 1190 Lim, Su Yin, and Helen Rizos. 2024. 'Single-cell RNA sequencing in melanoma: what have we learned
1191 so far?', *eBioMedicine*, 100.
- 1192 Li, D. et al. ECharts: A declarative framework for rapid construction of web-based visualization. *Visual
1193 Informatics* 2, 136–146 (2018).
- 1194 Longo, S. K., Guo, M. G., Ji, A. L. & Khavari, P. A. Integrating single-cell and spatial transcriptomics to
1195 elucidate intercellular tissue dynamics. *Nat. Rev. Genet.* 22, 627–644 (2021).
- 1196 Lewis, S. M. et al. Spatial omics and multiplexed imaging to explore cancer biology. *Nat. Methods* 18,
1197 997–1012 (2021).
- 1198 Landi, Maria Teresa, D. Timothy Bishop, Stuart MacGregor, Mitchell J. Machiela, et al. 2020.
1199 'Genome-wide association meta-analyses combining multiple risk phenotypes provide insights into the
1200 genetic architecture of cutaneous melanoma susceptibility', *Nature Genetics*, 52: 494-504.
- 1201 Lichtenberger, B. M. & Kasper, M. Cellular heterogeneity and microenvironmental control of skin
1202 cancer. *J. Intern. Med.* 289, 614–628 (2021).
- 1203 Leiter, U., Keim, U. & Garbe, C. Epidemiology of skin cancer: Update 2019. *Adv. Exp. Med. Biol.* 1268,
1204 123–139 (2020).
- 1205 Martincorena, I. et al. Tumor evolution. High burden and pervasive positive selection of somatic
1206 mutations in normal human skin. *Science* 348, 880–886 (2015).
- 1207 Marks, R., Dorevitch, A. P. & Mason, G. Do all melanomas come from 'moles'? A study of the
1208 histological association between melanocytic naevi and melanoma. *Australas. J. Dermatol.* 31, 77–80
1209 (1990).
- 1210 Martinelli, A. L. & Rapsomaniki, M. A. ATHENA: Analysis of Tumor Heterogeneity from Spatial Omics
1211 Measurements. *Bioinformatics* (2022) doi:10.1093/bioinformatics/btac303.
- 1212 Najem, A., J. Wouters, M. Krayem, F. Rambow, M. Sabbah, F. Sales, A. Awada, S. Aerts, F. Journe,
1213 J. C. Marine and G. E. Ghanem (2021). "Tyrosine-Dependent Phenotype Switching Occurs Early in
1214 Many Primary Melanoma Cultures Limiting Their Translational Value." *Front Oncol* 11: 780654.
- 1215 Owens, D. M. & Watt, F. M. Contribution of stem cells and differentiated cells to epidermal tumours.
1216 *Nat. Rev. Cancer* 3, 444–451 (2003).
- 1217 Pampena, R. et al. A meta-analysis of nevus-associated melanoma: Prevalence and practical
1218 implications. *J. Am. Acad. Dermatol.* 77, 938–945.e4 (2017).
- 1219 Pham, Duy, Xiao Tan, Brad Balderson, Jun Xu, Laura F. Grice, Sohye Yoon, Emily F. Willis, Minh
1220 Tran, Pui Yeng Lam, Arti Raghubar, Priyakshi Kalita-de Croft, Sunil Lakhani, Jana Vukovic, Marc J.
1221 Ruitenberg, and Quan H. Nguyen. 2023. 'Robust mapping of spatiotemporal trajectories and cell-cell
1222 interactions in healthy and diseased tissues', *Nature Communications*, 14: 7739.

- 1223 Rieckmann, J. C. et al. Social network architecture of human immune cells unveiled by quantitative
1224 proteomics. *Nat. Immunol.* 18, 583–593 (2017).
- 1225 Rawles, M. E. Origin of pigment cells from the neural crest in the mouse embryo. *Physiol. Zool.* 20,
1226 248–266 (1947).
- 1227 Rogers, H. W., Weinstock, M. A., Feldman, S. R. & Coldiron, B. M. Incidence estimate of
1228 nonmelanoma skin cancer (keratinocyte carcinomas) in the U.S. population, 2012. *JAMA Dermatol.*
1229 151, 1081–1086 (2015).
- 1230 Ratushny, V., Gober, M. D., Hick, R., Ridky, T. W. & Seykora, J. T. From keratinocyte to cancer: the
1231 pathogenesis and modeling of cutaneous squamous cell carcinoma. *J. Clin. Invest.* 122, 464–472
1232 (2012).
- 1233 Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression
1234 analysis of digital gene expression data. *Bioinformatics*. 2010 Jan 1;26(1):139–40. doi:
1235 10.1093/bioinformatics/btp616.
- 1236 Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2022. *CA Cancer J. Clin.* 72, 7–
1237 33 (2022).
- 1238 Sini, M. C. et al. Genetic alterations in main candidate genes during melanoma progression.
1239 *Oncotarget* 9, 8531–8541 (2018).
- 1240 Sagebiel, R. W. Melanocytic nevi in histologic association with primary cutaneous melanoma of
1241 superficial spreading and nodular types: effect of tumor thickness. *J. Invest. Dermatol.* 100, 322S–
1242 325S (1993).
- 1243 Sharma, P. & Allison, J. P. Immune checkpoint targeting in cancer therapy: toward combination
1244 strategies with curative potential. *Cell* 161, 205–214 (2015).
- 1245 Schürch, C. M., S. S. Bhate, G. L. Barlow, D. J. Phillips, L. Noti, I. Zlobec, P. Chu, S. Black, J.
- 1246 Demeter, D. R. McIlwain, S. Kinoshita, N. Samusik, Y. Goltsev and G. P. Nolan (2020). "Coordinated
1247 Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front."
1248 *Cell* 182(5): 1341-1359.e1319.
- 1249 Stanley, E. R. & Chitu, V. CSF-1 receptor signaling in myeloid cells. *Cold Spring Harb. Perspect. Biol.*
1250 6, (2014).
- 1251 Song, Liyang, Wenhao Chen, Junren Hou, Minmin Guo, and Jian Yang. 2024. 'Spatially resolved
1252 mapping of cells associated with human complex traits', medRxiv: 2024.10.31.24316538.
- 1253 Seviiri, Mathias, Matthew H. Law, Jue-Sheng Ong, Puya Gharakhani, Pierre Fontanillas, Stella
1254 Aslibekyan, et al. 2022. 'A multi-phenotype analysis reveals 19 susceptibility loci for basal cell
1255 carcinoma and 15 for squamous cell carcinoma', *Nature Communications*, 13: 7650.
- 1256 Sievert, C. Interactive Web-based Data Visualization with R, Plotly, and Shiny. (Chapman & Hall/CRC,
1257 2020).
- 1258 Shain, A. H., and B. C. Bastian. 2016. 'From melanocytes to melanomas', *Nat Rev Cancer*, 16: 345–
1259 58.
- 1260 Santoriello, C., A. Sporrij, S. Yang, R. A. Flynn, T. Henriques, B. Dorjsuren, E. Custo Greig, W.
1261 McCall, M. E. Stanhope, M. Fazio, M. Superdock, A. Lichtig, I. Adatto, B. J. Abraham, M. Kalocsay, M.
1262 Juryneć, Y. Zhou, K. Adelman, E. Calo, and L. I. Zon. 2020. 'RNA helicase DDX21 mediates
1263 nucleotide stress responses in neural crest and melanoma cells', *Nat Cell Biol*, 22: 372-79.
- 1264 Schütz, S., L. Solé-Boldo, C. Lucena-Porcel, J. Hoffmann, A. Brobeil, A. S. Lonsdorf, M. Rodríguez

- 1265 Paredes, and F. Lyko. 2023. 'Functionally distinct cancer-associated fibroblast subpopulations
1266 establish a tumor promoting environment in squamous cell carcinoma', Nat Commun, 14: 5413.
- 1267 Siegle, J. M., A. Basin, A. Sastre-Perona, Y. Yonekubo, J. Brown, R. Sennett, M. Rendl, A. Tsirigos, J.
1268 A. Carucci and M. Schober (2014). "SOX2 is a cancer-specific regulator of tumour initiating potential in
1269 cutaneous squamous cell carcinoma." Nat Commun 5: 4511.
- 1270 Tickle T, Tirosh I, Georgescu C, Brown M, Haas B (2019). inferCNV of the Trinity CTAT Project.
1271 Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA.
- 1272 Tirosh, I., B. Izar, S. M. Prakadan, M. H. Wadsworth, 2nd, D. Treacy, J. J. Trombetta, A.
1273 Rotem,C.Rodman, C. Lian, G. Murphy, M. Fallahi-Sichani, K. Dutton-Regester, J. R. Lin, O. Cohen, P.
1274 Shah, D. Lu, A. S. Genshaft, T. K. Hughes, C. G. Ziegler, S. W. Kazer, A. Gaillard, K. E. Kolb, A. C.
1275 Villani, C. M. Johannessen, A. Y. Andreev, E. M. Van Allen, M. Bertagnolli, P. K. Sorger, R. J. Sullivan,
1276 K. T. Flaherty, D. T. Frederick, J. Jané-Valbuena, C. H. Yoon, O. Rozenblatt-Rosen, A. K. Shalek, A.
1277 Regev, and L. A. Garraway. 2016. 'Dissecting the multicellular ecosystem of metastatic melanoma by
1278 single-cell RNA-seq', Science, 352: 189-96.
- 1279 Tran, M. et al. A robust experimental and computational analysis framework at multiple resolutions,
1280 modalities and coverages. Front. Immunol. 13, 911873 (2022).
- 1281 Thieu, K., Ruiz, M. E. & Owens, D. M. Cells of origin and tumor-initiating cells for nonmelanoma skin
1282 cancers. Cancer Lett. 338, 82–88 (2013).
- 1283 Tuong, Z. K., A. Lewandowski, J. A. Bridge, J. L. G. Cruz, M. Yamada, D. Lambie, R. Lewandowski, R.
1284 J. Steptoe, G. R. Leggatt, F. Simpson, I. H. Frazer, H. P. Soyer and J. W. Wells (2019).
1285 "Cytokine/chemokine profiles in squamous cell carcinoma correlate with precancerous and cancerous
1286 disease stage." Scientific Reports 9(1): 17754.
- 1287 Tuong, Z. K., S. W. Lukowski, Q. H. Nguyen, J. Chandra, C. Zhou, K. Gillinder, A. A. Bashaw, J. R.
1288 Ferdinand, B. J. Stewart, S. M. Teoh, S. J. Hanson, K. Devitt, M. R. Clatworthy, J. E. Powell and I. H.
1289 Frazer (2021). "A model of impaired Langerhans cell maturation associated with HPV induced
1290 epithelial hyperplasia." iScience 24(11): 103326.
- 1291 Tsao, H., C. Bevona, W. Goggins and T. Quinn (2003). "The transformation rate of moles (melanocytic
1292 nevi) into cutaneous melanoma: a population-based estimate." Arch Dermatol 139(3): 282-288.
- 1293 Van Hove, L. & Hoste, E. Activation of fibroblasts in skin cancer. J. Invest. Dermatol. 142, 1026–1031
1294 (2022).
- 1295 Villani, R., V. Murigneux, J. Alexis, S. L. Sim, M. Wagels, N. Saunders, H. P. Soyer, L. Parmentier, S.
1296 Nikolaev, J. L. Fink, E. Roy and K. Khosrotehrani (2021). "Subtype-Specific Analyses Reveal
1297 Infiltrative Basal Cell Carcinomas Are Highly Interactive with their Environment." J Invest Dermatol
1298 141(10): 2380-2390.
- 1299 Wu, S., Han, J., Laden, F. & Qureshi, A. A. Long-term ultraviolet flux, other potential risk factors, and
1300 skin cancer risk: a cohort study. Cancer Epidemiol. Biomarkers Prev. 23, 1080–1089 (2014).
- 1301 Werner, R. N. et al. The natural history of actinic keratosis: a systematic review. Br. J. Dermatol. 169,
1302 502–518 (2013).
- 1303 Weatherhead, S. C., Haniffa, M. & Lawrence, C. M. Melanomas arising from naevi and de novo
1304 melanomas - Does origin matter? Br. J. Dermatol. 156, 72–76 (2007).
- 1305 Weinberg, A. S., Ogle, C. A. & Shim, E. K. Metastatic cutaneous squamous cell carcinoma: An update.
1306 Dermatol. Surg. 33, 885–899 (2007).
- 1307 Wang, Y. et al. IL-34 is a tissue-restricted ligand of CSF1R required for the development of

- 1308 Langerhans cells and microglia. *Nat. Immunol.* 13, 753–760 (2012).
- 1309 Werner, S., Krieg, T. & Smola, H. Keratinocyte–Fibroblast Interactions in Wound Healing. *Journal of*
1310 *Investigative Dermatology* vol. 127 998–1008 Preprint at <https://doi.org/10.1038/sj.jid.5700786> (2007).
- 1311 Wang, J. X., Fukunaga-Kalabis, M. & Herlyn, M. Crosstalk in skin: melanocytes, keratinocytes, stem
1312 cells, and melanoma. *J. Cell Commun. Signal.* 10, 191–196 (2016).
- 1313 Xu, Z. et al. Integrin β 1 is a critical effector in promoting metastasis and chemo-resistance of
1314 esophageal squamous cell carcinoma. *Am. J. Cancer Res.* 7, 531–542 (2017).
- 1315 Wolf, F., Angerer, P. & Theis, F. SCANPY: large-scale single-cell gene expression data
1316 analysis. *Genome Biol.* 19, 15 (2018).
- 1317 Xu, D., S. Ding, M. Cao, X. Yu, H. Wang, D. Qiu, Z. Xu, X. Bi, Z. Mu, and K. Li. 2021. 'A Pan-Cancer
1318 Analysis of Cystatin E/M Reveals Its Dual Functional Effects and Positive Regulation of Epithelial Cell
1319 in Human Tumors', *Front Genet*, 12: 733211.
- 1320 Yu, G. et al. GOSemSim: an R package for measuring semantic similarity among GO terms and gene
1321 products. *Bioinformatics* 26, 976–978 (2010).
- 1322 Yu, Q. and I. Stamenkovic (1999). "Localization of matrix metalloproteinase 9 to the cell surface
1323 provides a mechanism for CD44-mediated tumor invasion." *Genes Dev* 13(1): 35-48.
- 1324 Yang, S. H., T. Andl, V. Grachtchouk, A. Wang, J. Liu, L. J. Syu, J. Ferris, T. S. Wang, A. B. Glick, S.
1325 E. Millar and A. A. Dlugosz (2008). "Pathological responses to oncogenic Hedgehog signaling in skin
1326 are dependent on canonical Wnt/beta3-catenin signaling." *Nat Genet* 40(9): 1130-1135.
- 1327 Zhang, C., H. Shen, T. Yang, T. Li, X. Liu, J. Wang, Z. Liao, J. Wei, J. Lu, H. Liu, L. Xiang, Y. Yang, M.
1328 Yang, D. Wang, Y. Li, R. Xing, S. Teng, J. Zhao, Y. Yang, G. Zhao, K. Chen, X. Li, and J. Yang. 2022.
1329 'A single-cell analysis reveals tumor heterogeneity and immune environment of acral melanoma', *Nat
1330 Commun.*, 13: 7250.

1331 **Main Figure Legends**

1332 **Figure 1. Integrating 12 single cell and spatial technologies to create cell type, community and**
1333 **cell-cell interaction reference atlases for skin cancer.**

1334 (a) Simplified cross-section of the human epidermis, highlighting squamous cells, melanocytes and
1335 basal cells. Coloured regions represent cSCC (green), which originates from squamous cells, melanoma
1336 (orange), which originates from melanocytes, and BCC (blue), which originates from basal cells. Two
1337 orange melanocytes are shown in the dermal region as occurs in invasive melanoma; other cells in the
1338 lower dermis layer are not depicted.

1339 (b) Overview of sample design and technologies used to generate data for this project. ROI - region of
1340 interest; FOV - field of view; S - cSCC; B - BCC; M - melanoma; HC - healthy (cancer patient); HNC -
1341 healthy (non-cancer patient donor). Technologies included are single cell RNA sequencing for fresh
1342 samples, single nuclei sequencing for formalin-fixed samples, Visium, Xenium, CosMX, GeoMX DSP
1343 for whole transcriptome, GeoMX DSP for proteins, Polaris, RNAscope, the proximal ligation assay, ,
1344 spatial glycomics and CODEX.

1345 **Figure 2. A single-cell atlas of cSCC, BCC and healthy skin cell types**

1346 (a) UMAP plot showing the integration of 45,909 healthy and cSCC cells from 11 samples of five
1347 patients, indicating results of Level 2 cell type annotation. Seventeen cell types were identified - eight
1348 immune cell clusters, six KC clusters, endothelial cells, fibroblasts, and melanocytes - plus an additional
1349 cluster of ambiguous cells.

1350 (b) Dendrogram showing the cell classification hierarchy, including Level 3 annotation of immune cells.

1351 (c) Distinguishing markers of 17 Level 2 cell types. Markers are a combination of predicted markers for
1352 each cluster, plus known canonical markers for each cell type.

1353 (d) Subclustering of keratinocytes showing the six Level 2 subtypes.

1354 (e-g) Classification of cancerous KC cells. Candidate cells were first classified as being aneuploid (red)
1355 if both InferCNV and CopyKat predicted them to be as such (e). Cells were then assigned an "cSCC
1356 score" (Module score calculated based on the cumulative expression of genes differentially expressed
1357 in KCs in the cancer samples as compared to those from the normal samples) using differentially
1358 expressed genes identified using two different methods, edgeR and scanpy (f). Finally, cells were
1359 classified as KC Cancer (g) if they were classified as aneuploid (e) and also received an cSCC score
1360 above the 95th percentile of all cell scores (f). The venn diagram indicates the number of cells passing
1361 the module score threshold by edgeR and scanpy.

1362 **Figure 3. A single-cell atlas of melanoma cell types**

1363 (a) UMAP plot showing the integration of 10,747 melanoma cells from three patient samples, indicating
1364 results of Level 2 cell type annotation. Eighteen cell types were identified - melanocytes, seven immune
1365 cell clusters, five KC clusters, and five other cell types.

1366 (b) Dendrogram showing the Level 2 cell classification hierarchy.

1367 (c) Distinguishing markers of 18 Level 2 cell types. Markers are a combination of predicted markers for
1368 each cluster, plus known canonical markers for each cell type.

1369 (d) Result of Level 2 reclustering and cell type annotation for melanocytes.

1370 (e-g) Results for classification of cancerous melanoma cells. Melanocytes from the patient with the
1371 malignant tumor were classified as likely melanomas if they were both predicted to have aneuploid
1372 genomes (red) by both InferCNV and CopyKat (e). Cells were then assigned a "melanoma score" (f).
1373 Specifically, a module score was computed using genes upregulated in the melanoma sample compared
1374 to the benign sample using both edgeR pseudobulking and scanpy non-parametric test. For the sample
1375 from melanoma patient, a majority of the cells with a score >80th percentile cut-off were from the
1376 Melanoma sample cluster (Clusters 9 and 10), (g) and finally the cells inferred 'Aneuploid' by the CNV
1377 analysis and with a high module score by both the aforementioned methods are labelled as malignant

1380 melanocytes (red) as shown in the UMAP. The venn diagram indicates the number of cells passing the
1381 module score threshold by edgeR and scanpy.
1382

1383 **Figure 4: Biological distinctions between cSCC, BCC and melanoma in the single-cell RNASeq**
1384 **atlas**

1385 (a) Dot plot showing the percentage of each Level 2 cell type within patient samples. Dots are coloured
1386 by cell type category and dot size indicates their percentage within each sample; all columns sum to
1387 100. Results of differential abundance statistical tests are shown to the right, comparing abundance in
1388 cSCC vs melanoma, cSCC vs healthy skin, and melanoma vs healthy skin. Asterisks indicate the sample
1389 in which the cell type was found to be more abundant, either healthy skin (pink), cSCC-BCC (blue) or
1390 melanoma (yellow).

1391 (b) A venn diagram of the top significant upregulated genes across cancerous and non-cancerous KCs
1392 and melanocytes. (red) Upregulated in cSCC/BCC KC Cancer cells compared to Malignant Melanocytes
1393 from melanoma samples, (green) Upregulated in Malignant Melanocytes from melanoma samples
1394 compared to cSCC/BCC KC Cancer cells, (yellow) Upregulated in Malignant melanocytes compared to
1395 other melanocytes in melanoma samples, (blue) Upregulated in cancer KCs compared to other KCs in
1396 cSCC/BCC sample.

1397 (c) Heatmaps showing top 50 differentially expressed genes across Cancer vs Normal KCs (top left),
1398 Melanocytes vs Melanoma (bottom). Each column of the heatmap indicates a pseudo-bulked pool.

1399 (d) Integrative, multiple platform analysis of differentially expressed genes. From left to right, the Venn
1400 diagram shows the overlap between DE genes between cSCC cancer KCs vs normal KCs across
1401 scRNAseq and for KCs in cancerous tissues compared to those from the normal tissues from non-
1402 cancer donors with spatial datasets of Visium, Xenium and CosMX.

1403 e) UMAP plot for scRNAseq data showing the expression of SOX2 in cancer vs non-cancer samples,
1404 which matches the location of KC cancer cells in UMAP shown in Fig 2.

1405 f) Tissue gene expression plot of CosMX data showing two of the five shared markers SOX2 and
1406 LAMP3. Pathological annotation of the region is shown on the left.

1407
1408 **Figure 5. Integrative analysis across different individuals and technologies to build a robust**
1409 **spatial community atlas.**

1410 (a) Cross-modality comparison of the ten communities identified for each of Visium, CosMx and Xenium.
1411 Each row shows a community identified from one of the three spatial platforms. The left heatmap shows
1412 similarity across communities within and between technologies, measured by pairwise Pearson
1413 correlation values between communities based on their cell type composition. This allows similar
1414 communities across technology platforms and samples to be grouped to form meta-communities. The
1415 right heatmap shows the cellular makeup of each community (i.e. proportion of each cell type per
1416 community), providing information to label the groups of communities. The central annotation shows the
1417 broad classification of communities into immune, KC, stromal or tumour-related communities, based on
1418 the cellular makeup of each.

1419 (b) Spatial localisation of cells belonging to communities CosMx_6 (left) and Xenium_2 and Xenium_7
1420 (right). Together with Visium_2, these communities form a meta-community that is enriched for
1421 melanocytes.

1422 (c) Inter-community communication within melanoma CosMx_6. The chord plot visualises cell-cell
1423 communication mediated by Collagen signaling pathways, using the CellChat pathway database. Lines
1424 connect communicating cell types; line thickness represents greater communication between cell pairs.

1425 (d) Ligand-receptor interactions between pairs of cell types within the melanoma community CosMx_6.
1426 Top significant L-R pairs and corresponding cell type pairs are shown.

1428 (e) Cell type co-occurrence in CosMx samples between melanocytes and either other melanocytes
1429 (brown), Treg cells (blue), fibroblasts (green) or other cells (black). Each line plots the co-occurrence
1430 score (y-axis) between melanocytes and the test cell type calculated over increasing spaital distances
1431 (x-axis). The samples from left to right are melanoma 23346-105P, 30037-07BR and 6475-07FC.
1432 (f-g) Cell type proportions of communities identified in Xenium (f) and CODEX (g) for adjacent sections
1433 from the same sample (48974-2B). The melanoma community in both datasets is enriched with
1434 melanocytes.
1435 (h) Joint pathway analysis using upregulated genes or proteins of the melanocyte communities in
1436 Xenium and CODEX data (shown in f and g), and highly expressed glycans of the melanocyte
1437 community in MALDI data (shown in Fig S12a). The proteins, genes, metabolites are mapped to KEGG
1438 metabolic pathways. The X-axis shows the number of genes/proteins from Xenium and CODEX data
1439 found in the pathway, while the Y-axis shows glycans in the same pathway.

1440
1441
Figure 6: Differential multimodal cell-cell interaction across three cancer types using
integrative analysis of Visium, CosMx, Xenium data.

1442 (a) Heatmap of LR scores for LR pairs enriched per cancer type, with a consistent trend across samples
1443 and the two Visium and CosMx platforms. Differentially expressed LR pairs were calculated comparing
1444 each cancer type vs the others using a pseudobulked LR scores with 3 pools per sample. Each heatmap
1445 row is a distinct CosMx or Visium sample. The two L-R pairs specific for melanoma IL34-CSF1R and
1446 FGF2-CD44 were used for experimental validations.

1447 (b) Differential interaction analysis based on LR pairs and cell type pairs. The Venn diagram compares
1448 differential LR pair results between Melanoma and the combined BCC + cSCC datasets, calculated
1449 using edgeR with pseudobulked LR scores. The diagram highlights consistent and unique results
1450 between CosMX and Visium, where Up indicates a higher LR score in Melanoma and Down indicates a
1451 lower score in BCC + cSCC. Cell-to-cell communication between the LR pairs that are up- and
1452 downregulated in melanoma in both CosMx and Visium is shown in the two Network plots flanking the
1453 Venn diagram. In both Network plots, the purple arrows show pairs of cell types that have interactions
1454 higher in Melanoma and green arrows show interactions between cell type pairs more in the BCC +
1455 cSCC than in Melanoma. The number displayed for each arrow shows the integrated p-value across all
1456 biological replicates (the thicker arrows indicate more interactions). Interactions between the two cell
1457 types can still be significantly upregulated in melanoma even if the set of LR pairs were downregulated.

1458 (c) Spatial mapping of cancer type-enriched LR pairs in CosMx data. Two LR pairs that were significantly
1459 different between cancer types across technologies in Panel A are shown, namely IL34-CSF1R (higher
1460 in melanoma) and S100A8-TLR4 (higher in BCC). Both are visualised in FOVs from melanoma sample
1461 (top) and BCC sample (bottom). For each cancer type, the cell type annotation of the FOV is shown (top
1462 left) with orange and black boxes indicating the highlighted regions (right). An overview of interactions
1463 at tissue level is shown by large coloured arrows, representing cumulative interactions between two cell
1464 types in the tissue, with the location of the arrow root as the centroid coordinate of all cells in one cell
1465 type (bottom left). Magnified boxes (right) show the presence of ligands (pink) and receptors (red), with
1466 white arrows showing the connections between ligands and receptors of nearby cells.

1467 (d) Melanoma drug target graph integrating multiple biological and pharmacological knowledge types.
1468 Nodes represent genes, drugs, and biological functions. Level 1 connections show melanoma-
1469 associated genes and drugs targeting melanoma. Level 2 links display drugs targeting the melanoma-
1470 associated genes from Level 1 and a broader gene set targeted by drugs in the network. All genes in
1471 the graph are either upregulated or have high ligand-receptor scores. Clusters 1, 2, and 3 are pathways
1472 enriched with genes shown in the graph.

1473
1474
Figure 7. Multi-omics analysis of LR interaction between IL34 and CSF1R

1477 (a) H&E staining of a tissue section from cSCC patient B18, with a pathologist's annotation highlighting
1478 a cancer nest region.
1479 (b) Mapping of IL34_CSF1R co-localisation in RNAscope data from an adjacent section of the tissue
1480 block to that shown in **Panel a**. Image tiles containing both CSF1R and IL34 RNA were identified using
1481 a window scanning strategy in STRISH. The colour scale represents the -log(p-value) of each window's
1482 LR co-expression score, while the red box indicates the tissue border. The white box highlights the
1483 tissue zone shown in **Panel c**.
1484 (c) Magnified region corresponding to the white box in **Panel b**, showing DAPI-stained cell nuclei (blue)
1485 with CSF1R (yellow) and IL34 (red) RNA visible as coloured dots.
1486 (d) Expression of IL34 and CSF1R in healthy and cSCC samples from scRNASeq data. This interaction
1487 was statistically predicted to occur in both healthy and cancer samples in our CellChat analysis.
1488 (e) Exemplar spatial plots showing the LR score for IL34_CSF1R from patient 48974. The black box
1489 indicates a region highlighted below the main image. Here, zoomed-in boxes show the IL34_CSF1R LR
1490 score (left) and IL34 (middle) and CSF1R (right) gene expression for the same tissue region.
1491 (f-g) Heatmaps indicating grouped GO terms and associated genes that are enriched in IL34_CSF1R-
1492 positive spots in melanoma samples compared to IL34_CSF1R-negative spots. GO term groups were
1493 calculated by k-means clustering ($k = 3$) of GO semantic similarity scores; two such groups are shown
1494 here. The full heatmap is shown in **Fig S9b**.
1495 (h) Proximal ligation assay (PLA) for validating CD44 interactions in melanoma. A merged image of
1496 signal for the ligand and the receptor and a zoom-in window highlighting the interaction on the cell
1497 membrane. A positive PLA signal is visible if two interacting proteins are in a proximity less than 20 nm.
1498

1499 **Figure 8. Mapping genetics effects from genome-wide association studies for cutaneous
1500 melanoma, cSCC, and BCC to spatial domains and cell types.**

1501 (a) Gene specificity score (GSS) and association of spatial spots with skin cancer heritability. GSS score
1502 for each gene in a spot/cell represents the enrichment of the gene as a top rank most abundant gene in
1503 the spot/cell and its neighbour spots/cells in an anatomical region, a spatial domain, or a cell type. The
1504 p-value shows the spatial heritability enrichment significance of a spot with a trait based on SNPs
1505 mapped to the genes with high GSS scores (one-sided Z-test for stratified coefficient different to 0). The
1506 p-value is more significant if the SNPs that are mapped to the high GSS genes explain a higher
1507 proportion of heritability for the trait.
1508 (b) Cell types with the highest enrichment of heritability explained by SNPs tagged to GSS genes of
1509 cells in a cell type. The white asterisks indicate the most enriched cell-type for heritability of cutaneous
1510 melanoma, cSCC and BCC traits.
1511 (c) gsMAP significance spatial heritability enrichment is shown at single-cell resolution across the tissue
1512 (upper tissue plots) or per annotated skin regions (lower violin plots) from the cosMx data of the sample
1513 mel48974.
1514 (d) LR pairs with significant association with SNP heritability explained by the corresponding cell types.
1515 The rectangles show cases where both L and R genes had PCC >0.3 between GSS of the gene and
1516 the gsMAP P-values (the significance level for the LD stratified coefficients for the spot bigger than 0).
1517 The results suggest which LR pairs are related with the heritability of a cell type pairs.
1518 (e) GSS of two LR pairs showing specificity of the L and R genes to tissue regions at the immune-rich
1519 dermal layers and the epidermis of the skin.
1520 (f) Manhattan plot showing top significant GWAS SNPs co-localizing with genes in melanocytes (red)
1521 and T cells (blue) that had the highest Pearson correlation between GSS and the gsMAP trait association
1522 P-value or associated with SNPs with genome-wide significance. The Y-axis shows the -log(P-value)
1523 from GWAS analysis.
1524

1525 **Supplemental Figure Legends**

1526
1527 **Figure S1. The skincanceratlas database allows interactive analysis of genes, cells, interactions**
1528 **across the multi-omics skin cancer data resource.**

1529 *The skincanceratlas function allows users to visualise clustering, cell type and spatial data for cSCC,*
1530 *BCC and melanoma. The Gene Explorer function (bottom) allows visualisation of gene expression and*
1531 *LR interactions at single cell and/or spatial resolution from data generated with Visium, CosMX and*
1532 *scRNAseq for all three skin cancer types.*

1533
1534 **Fig S2 - Annotation of KC and Immune cells in BCC-cSCC.**

1535 (a) Marker expression for the level 3 annotation of immune cells.

1536 (b) Marker expression for the level 3 annotation of KC cells.

1537 (c) UMAP displaying the overall level 3 annotation of all 30 cell types.

1538 (d) Visualisation of the single cell (sub)clusters that are mapped to the tissue using the CosMX data.
1539 scRNAseq data was used as the reference to transfer cell annotation labels to CosMX cells using the
1540 RCTD deconvolution method. The KC basal, KC hair, KC differentiating, KC cornified clearly form
1541 distinct layers in the cSCC skin sample B18. The cell types defined in scRNAseq data clearly mapped
1542 to distinct anatomical layers of the skin, providing cross-platform evidence for correct cell type annotation
1543 of the scRNAseq reference data, which did not have a spatial context.

1544
1545 **Figure S3. A single-cell atlas of cSCC cell types**

1546 (a-b) UMAP plots showing the integration of ~46,000 cells from 11 samples of five cSCC patients,
1547 showing cell distribution by patient sample (a) and cancer status (b).

1548 (c) Expression of selected genes that are amongst the top 10 most differentially expressed genes
1549 between cancer and healthy cells when compared across the entire dataset. Each column shows only
1550 the healthy (left) or cancer (right) cells split from the full UMAP plot shown in (a-b).

1551 (d) Heatmap showing expression in healthy and cancer samples of the 57 core cancer genes shared
1552 between all five patients (as in Panel F).

1553 (e) Bar plot showing the distribution of each cell type across eleven samples.

1554 (f) Genes overexpressed in healthy (left) or cancer (right) cells across individual patients. Core suites of
1555 genes that are detected across four or more of the five patients are highlighted in the centre of each
1556 diagram; suites contain a total of 217 healthy and 136 cancer genes respectively.

1557 (g) The top six non-redundant gene ontology terms enriched in the two gene suites, ranked by -log10
1558 adjusted P-value.

1559
1560 **Fig S4 - Annotation of melanoma samples**

1561 (a) UMAP shows the annotation of all the three annotation levels from common to specific cell types.
1562 UMAP showing Level 3 annotation also displays melanoma cells, which were not defined by Level 1
1563 and Level 2 annotation.

1564 (b) Heatmap showing marker expression for the level 3 annotation.

1565
1566 **Fig S5 - Pathway analysis for gene markers of cSCC-BCC and melanoma**

1567 (a) Heatmap showing enriched pathways for 3267 genes KC Cancer Vs melanoma, tested against the
1568 MSigDB. All DE genes upregulated in Cancer cells in cSCC-BCC compared to normal KC cells (left)
1569 and in melanomas compared to normal melanocytes were used. Pathway names are shown in columns
1570 and genes in these pathways are shown in rows. The red bars indicate -log(P-value) and the terms are
1571 sequentially ordered with the most significant pathway on the left.

1572 (b) MSigDB pathways for 2713 genes up in Melanoma Vs KC Cancer.

1573 (c) Integrated pathway analysis, using KEGG and Reactome databases for genes higher in cancer KCs

1574 compared to the normal KCs in cSCC-BCC samples observed in scRNA-seq (left; 176 KC Cancer vs
1575 normal KCs) and for genes higher in malignant melanocytes compared to the normal melanocytes in
1576 snRNAseq Melanoma samples (68 genes up in Melanoma Vs normal melanocytes).
1577 (d) DE genes from comparing KC cells and Melanocytes.

1578

1579 **Figure S6. Cell type annotation for the Visium and Xenium data of samples from three normal**
1580 **control donors.**

1581 The donors did not have skin cancer. The top and middle panels show Visium and Xenium data of
1582 adjacent tissue sections, respectively, and a zoom-in view of a Xenium FOV is shown at the bottom
1583 panel. The legend colors indicate cell types.

1584

1585 **Figure S7. Captured regions of interest for GeoMx protein quantification**

1586 Morphology imaging of cSCC biopsies for GeoMx protein data, from patients R01 (a, top 3 tissues), P04
1587 (a, bottom tissue) and B18 (b). Boxes show zoomed-in regions of interest (ROIs) selected for marker
1588 quantification, connected to their location in the full tissue image by red lines. Morphology markers show
1589 DNA (blue, DAPI), CD3 (red), CD45 (yellow) and PanCK (green). ROIs were segmented by either CD45
1590 or PanCK expression.

1591

1592 **Figure S8. Analysis of GeoMx gene markers and cell type compositions using GeoMX RNA and**
1593 **protein assays.**

1594 (a) Heatmap of 33 genes which were differentially expressed between PanCK+ and CD45+ segments
1595 within the GeoMx RNA dataset, and were also identified as members of the core cancer- or healthy-
1596 associated gene suites. Column annotations separate samples by GeoMx segment (PanCK+ and
1597 CD45+ segments) and patient, while row annotation indicates whether genes were present in the
1598 cancer- or healthy-associated gene suites from **Fig S3f**. Gene label colour indicates whether the gene
1599 was upregulated in the CD45+ (blue) or PanCK+ (green) segment in the GeoMx analysis.

1600 (b) Validation of the presence of key cell types from our scRNASeq atlas, using RNA markers captured
1601 by GeoMx; this plot is complementary to **Fig S9n** which shows the same comparison for equivalent
1602 protein markers. Selected markers are CD163 and CD68 (M2 markers), MS4A1 (B cell marker), CD8A
1603 and CD8B (CD8+ T cell markers), ITGAX (dendritic cell marker), FOXP3 and FOXP3 (Treg markers),
1604 and FAP and FN1 (fibroblast markers). Corresponding protein markers for each gene are given in
1605 brackets. Top annotation bars separate samples by segment (PanCK+ and CD45+ segments) and
1606 patient. Marker association with particular cell types was taken from NanoString product information.
1607 Asterisks indicate markers that were differentially expressed between CD45+ and PanCK+ segments;
1608 markers that were not found to be differentially expressed are included as their expression provides
1609 evidence for the presence of these cell types.

1610 (c) Top 30 differentially expressed genes between PanCK+ and CD45+ segments within the GeoMx
1611 RNA dataset. A total of 267 differentially expressed genes were identified between segments.
1612 Differentially expressed genes were sorted by FDR and the top 30 genes were selected regardless of
1613 whether they were enriched in the PanCK+ or CD45+ segments.

1614 (d) Expression of all the 48 proteins (Immuno-oncology panel) in GeoMX proteomics data. The heatmap
1615 shows normalised data across FOVs.

1616 (e) Top 5 GO enrichment results per segment from an analysis of the top 100 differentially expressed
1617 genes associated with the CD45+ and PanCK+ segments of the GeoMx RNA data. Differentially
1618 expressed genes were sorted by absolute fold-change values and the top 100 genes were selected from
1619 each of the CD45+ and PanCK+ segments.

1620 (f) Inferred cell type proportions of ROIs from patients B18, R01 and P04 in the CD45+ segment of the
1621 RNA GeoMx data. RNA expression information was used to perform cell type deconvolution for the
1622 segmented ROIs. Bar width is proportional to the number of nuclei per ROI. Top annotations indicate

1623 the sample of origin for each ROI.

1624

1625 **Figure S9. Spatial mapping of cancer cell types within cSCC skin cancer as measured by GeoMx**
1626 **spatial transcriptomics**

1627 (a) Left, H&E stained image of cSCC tissue from patient B18, with pathologist's annotation indicating
1628 cancer (green) and immune cell (blue) tissue regions. This annotation was used in combination with
1629 morphology marker staining to define regions of interest (ROIs) for GeoMx spatial proteomics. Right,
1630 selected ROIs corresponding to the highlighted regions (ROI 002, patient B18), segmented based on
1631 CD45 (top) and PanCK (bottom) markers.

1632 (b) Validation of the presence of key cell types from our scRNASeq atlas, using protein markers captured
1633 by GeoMx. Selected markers are CD163 and CD68 (M2 markers), CD20 (B cell marker), CD8 (CD8+ T
1634 cell marker), CD11c (dendritic cell marker), FOXP3 and CD25 (Treg markers), FAP alpha and
1635 Fibronectin (fibroblast markers). Top annotation bars separate samples by segment (PanCK+ and
1636 CD45+ segments) and patient. Marker association with particular cell types was taken from Nanostring
1637 product information. FAP-a. = FAP-alpha, Fibro. = Fibronectin. Asterisks indicate markers that were
1638 differentially expressed between CD45+ and PanCK+ segments; markers that were not found to be
1639 differentially expressed are included as their expression provides evidence for the presence of these
1640 cell types.

1641

1642 **Fig S10 Integrative analysis of spatial multiomics atlas for melanoma - Cross-modality spatial mapping**
1643 across three cancer types

1644 (a) Visium mapping for BCC (left), cSCC (middle), and Melanoma (right), with Hematoxylin and Eosin
1645 (H&E) staining (top), cell type annotation (middle), and marker expression levels (bottom).

1646 (b) Spatial localisation of marker gene expression in CosMx data from BCC (left), cSCC (middle), and
1647 melanoma (right). Cell type annotations and marker expression are shown together (bottom), with
1648 selected regions highlighted in enlarged panels (middle). Pathological annotations are provided in the
1649 top thumbnails.

1650

1651 (c) Pathological annotation of the Xenium melanoma sample (48974-2B) on H&E image (top), and
1652 corresponding cell type annotation and TYR marker gene expression shown for the same region
1653 (bottom).

1654 (d) Cell type annotation for CODEX data on melanoma tissue

1655 (e) Immunofluorescent staining of S100A8 protein marker corresponds to the tissue region containing
1656 melanoma cells in Panel e.

1657 (f) Cell type annotation for spatial Glycomics data on an adjacent tissue section from the same block
1658 as shown in Panels d-e.

1659

1660 **Figure S11: Spatial cellular heterogeneity across cancer types using CosMx**

1661 (a) Spatial cell type annotations of cSCC patient B18 FOV 15, the FOV with the highest overall
1662 heterogeneity score for this sample.

1663 (b) Spataial cell heterogeneity plot for the FOV shown in Panel a. Cells are coloured by heterogeneity
1664 score (measured by cell neighbourhood analysis), highlighting regions of the FOV containing diverse
1665 cell types in close proximity.

1666 (c) Boxplot of heterogeneity scores across three cancer types, indicating the lowest overall
1667 heterogeneity score for cSCC. Asterisks indicate significant p-values from paired Wilcoxon test.

1668 (d) Boxplot of heterogeneity scores across tissue regions of each cancer type. Each FOV was
1669 designated as being either healthy, immune or cancer by a pathologist. Wilcoxon ranked sum test was
1670 applied.

1672 **S12. Integrative analysis of spatial transcriptomics, proteomics, and glycomics data for**
1673 **melanoma community.**

1674 (a) Joint pathway analysis using DE genes/proteins of melanocyte communities in Xenium and CODEX
1675 data, and glycans of melanocytes in MALDI data, to map to KEGG metabolic pathways. The X-axis
1676 shows enriched genes/proteins in Xenium and CODEX data, while the Y-axis shows enriched glycans.

1677 (b-c) Cell type proportions of the community identified in CODEX data (b) and Xenium data (c). The
1678 melanoma community in both datasets is enriched with melanocytes.

1679 (d) Clustering of the melanoma spatial glycomics sample, with the identification of the melanoma-
1680 enriched cluster 2. A heatmap of top markers for each cluster is on the left, and a tissue plot with cluster
1681 label is shown on the right. The expression for four representative metabolites is shown at the bottom
1682 on the panel.

1683 (e) Annotation for CODEX data from the adjacent tissue section, showing the melanoma cluster, with a
1684 heatmap of protein markers for each cell type, the tissue plot showing all key cell types, and the the
1685 expression of the three proteins at the bottom.

1686 (f) Similar to d) and e) the annotation of the melanoma adjacent tissue section for the Xenium data is
1687 shown.

1688 **Fig S13: A spatial community mapping of three melanoma samples for Xenium data**

1689 (a) Community mapping for one sample based on cell type composition. The heatmap shows cell type
1690 composition per community.

1691 (b) Expression of marker genes at single cell resolution, showing cell type (left), H&E image, and
1692 expression.

1693 (c) Communities and cell type composition for two other melanoma samples.

1694 (d) Cell-cell interaction analysis of the melanoma community. The line shows co-occurrence score with
1695 melanocytes, where a high score indicates the cells are closely distributed with melanocytes across
1696 the whole tissue. The scores change when the distance to search for neighbouring cells increases
1697 (i.e., less specific to a spatial domain).

1698 **Fig S14: A spatial community mapping of three melanoma samples for CODEX data**

1699 (a) A detailed cell type annotation and an example of colocalization between two cell types (Melanoma
1700 and Fibroblasts) are shown.

1701 (b) Community mapping with a zoom-in view of communities, highlighting two melanoma communities
1702 (shown. as melanoma-1 and melanoma-2).

1703 (c) Community network analysis, with most similar communities connected in a nearest neighbourhood

1704 (d) Examples of cell-type composition for each community of all three melanoma samples

1705 (e) Spatial distribution of communities, defined from merged analysis of all three samples.

1706 **Figure S15. Analysis of spatial glycomics data.**

1707 (a) Melanoma community analysis using spatial glycomics data. The colors show clustering results.
1708 Identification of a melanoma cluster (cluster 6 in samples 1 and 3, cluster 8 in sample 2; indicated with
1709 an asterisk) is based on histology and relative location compared to clusters from CODEX data.

1710 (b) Heatmap showing the expression of top metabolite markers (named by the mass m/z values) for
1711 each cluster identified in A. Top annotation bar shows sample of origin. The green box and associated
1712 green labels highlight the metabolites enriched across the three melanoma clusters.

1713 (c) Venn diagram highlighting the similarities and differences between markers identified for the
1714 melanoma cluster in each sample.

1715 (d) Expression of the top four metabolite markers shared between samples. The top metabolites are the
1716 most differentially expressed compared to all other clusters.

1721 **Figure S16. Integrative analysis across different individuals and technologies to build a robust**
1722 **spatial community atlas.**

1723 (a-b) Boxplots showing the proportion of each community deriving from BCC, cSCC and melanoma in
1724 Visium (a) and CosMX (b) data. Wilcoxon ranked sum test was applied.

1725 (c) Spatial distribution of ten communities in CosMX data in representative FOVs from BCC (patient B18,
1726 FOV 13, left), cSCC (patient B18, FOV 14, second left) and melanoma (patient 48974-2B, FOV 12,
1727 second right). Pathological annotations for these FOVs are given in **Fig S10b**. Spatial location of cells
1728 from CosMX_6, a member of the melanoma meta-community (Fig 5), is shown next to the original
1729 distribution of all communities (right).

1730 (d) Spatial distribution of ten communities in Xenium melanoma sample (right) relative to the location of
1731 cells from Community Xenium_2 from the melanoma meta-community (left).

1732
1733 **Figure S17. Cell-cell interaction analysis of scRNASeq data for cSCC**

1734 (a) Heatmap indicating significant LR pairs across samples. 312 significant pairs were detected from the
1735 entire suite of expressed genes in each sample. For visualisation purposes, only those LR pairs
1736 predicted in at least 4 of 10 samples are shown, with their corresponding -log10(pval) values. Non-
1737 significant LR pairs in each sample are white. LR communication was calculated between pairs of Level
1738 2 cell types; duplicate LR pairs in each sample (i.e. predicted for multiple cell types within one sample)
1739 were collapsed, with only the highest -log10(pval) score shown.

1740 (b) Patient-specific LR predictions in healthy (left) or cancer (right) cells across patients. The core suites
1741 of LR pairs that are shared across four or more of the five patients are highlighted in the centre of each
1742 Venn diagram.

1743 (c) Overlap between core LR suites in cancer and healthy patients.

1744 (d) The top six non-redundant GO terms that are enriched in the cancer-specific suite of LR genes,
1745 ranked by -log10 adjusted P-value. GO analysis was performed against a background gene universe of
1746 all LR pairs that are present in the CellChat database and are also expressed in our dataset.

1747
1748 **Figure S18. CCI analysis in cSCC samples using scRNASeq**

1749 Changes in cell type participation in cellular crosstalk between ligands and receptors in healthy and
1750 cancer samples from different patients. Circle diagrams show cells expressing ligands on the left, and
1751 receptors on the right. Connecting line colours indicate sender cell identity, and line widths indicate the
1752 number of unique LR connections predicted between pairs of cell types. Colours around the edge of the
1753 ring indicate sender and receiver cells' identities. Numbers below each plot indicate the number of cells
1754 (C) and significant interactions (I) for each sample. Interactions were predicted for pairs of Level 2 cell
1755 types but summarised to broader cell type categories for visualisation.

1756
1757 **Figure S19. Comparing LR detection between scRNASeq data without and different spatial**
1758 **transcriptomics platforms with spatial context.**

1759 (a) Examples of cases mis-detected by scRNASeq (WNT5A-ROR1) or detected by scRNASeq but not
1760 colocalize in spatial data, suggesting potential false positive.

1761 (b) Consistent detection of LR pairs by Visium, CosMX and Xenium data

1762
1763 **Figure S20. Comparing interactions across the three cancer types**

1764 (a) MMCCI integrated interactions for each cancer type and across multiple samples and spatial
1765 transcriptomics platforms. From left to right, circos interaction plots for BCC, cSCC and Melanoma. The
1766 last plot on the right is the integrated interactions for all cancer types for all samples.

1767 (b) Comparing differential interactions at cell type level for BCC vs cSCC and for the integrated BCC-
1768 cSCC vs Melanoma using all L-R pairs. For cSCC vs BCC comparison, purple arrows show more
1769 interactions in BCC and green arrows indicating more in cSCC. For cSCC-BCC vs Melanoma, the purple
1770 arrows show more in melanoma and green arrows show more in cSCC-BCC.

1771 (c) Differential interactions at cell type level using MIF-CD74 pair as the top different L-R pair shown in
1772 panel b. The difference indicates which cell type pairs are more actively interacting by considering the
1773 MIF-CD74 pair.

1774 (d) Differential CCI network plots for LR pairs in the EMT pathway found from MSigDb pathway analysis.
1775 The pathway-centric interaction analysis shows specific cell type pairs where the EMT pathway was
1776 most significantly different. Pathway enrichment analysis from MSigDb on upregulated LR pairs common
1777 across CosMx and Visium 1) Melanoma and 2) cSCC+BCC.

1778 (e) Interactions between cell-types in Melanoma compared to cSCC only.

1779 (f) Top LR pairs interacting between T cells and melanocytes.

1780

1781 **Figure S21. Spatial cell type and community mapping at the protein level by targeted
1782 multispectral imaging (Polaris)**

1783 (a) Polaris protein imaging data from cSCC patient B18 with pathological annotation of cancer and
1784 immune regions, based on tissue morphology. A panel of six antibodies was used to profile protein
1785 markers and major cell types of interest including CD8 (T cell marker), FoxP3 (T-regulatory cell marker),
1786 CD68 (monocyte/macrophage marker), PanCK (epithelial cell marker), PD-1 and PD-L1 (immune
1787 checkpoint inhibitor, with the former expressed on immune cells and the latter on tumour cells); DAPI
1788 was also included as a nuclear stain. White and green dashed lines outline the cancer and immune
1789 regions of the tissue, as annotated by a pathologist. The red box indicates the tissue border as
1790 determined later during STRISH analysis, shown here to aid comparison with subsequent panels.

1791 (b) Cell type classification based on clustering and gating of single-cell resolution protein expression
1792 signals.

1793 (c) Mapping of cancer-immune cell co-localisation using STRISH, identifying image tiles containing
1794 double-positive CD8+ PD-1+ immune cells and PanCK+ PD-L1+ cancer cells. The identified
1795 colocalisation zone correlates with the cancer infiltrate region identified in Panel a.

1796 (d) Voronoi diagram mapping the spatial distribution of cell communities, defined by further clustering of
1797 cell type classifications.

1798 (e) Cell type composition of the five cell communities defined in Panel d, revealing higher-level groups
1799 of communities representing cancer only (red box), immune cells only (orange box), and cancer-immune
1800 infiltrate (pink box). Heatmap colours represent the neighbourhood enrichment score for each
1801 community-cell type combination.

1802 (f-g) Cell type co-occurrence measurements of CD8+ PD-1+ FoxP3+ immune cells (f) and PanCK+
1803 cancer cells (g) with either the same or other cell types in the dataset. Each line plots the co-occurrence
1804 score between the reference cell type and the test cell type calculated over increasing spatial distances

1805

1806 **Figure S22. Experimental validation of ligand-receptor interaction using proximal ligation assay
1807 for melanoma samples.** In each row, fluorescence signal for a single antibody and PLA is shown,
1808 followed by a merge image and a Zoom-in window highlighting the interaction on the cell membrane. A
1809 positive PLA signal is visible within two proteins in a proximity less than 20 nm. From top row to
1810 bottom row, a positive control E-Cadherin and b-Catenin, a negative control for CD31 and AQP1, and
1811 the interaction results for the three pairs CD44-Fibronectin, CD44-MMP9, and CD44-FGF2 ligand-
1812 receptor pairs.

1813

1814 **Figure S23. Prognostic value of melanoma specific genes identified with scRNA-seq and spatial
1815 transcriptomics.** The plots indicate survival curves of patients from the TCGA SKCM dataset for groups
1816 stratified as high and low expressing patients for

1817 (a) genes upregulated in malignant melanocytes compared to normal melanocytes identified by DE
1818 analysis with edgeR

1819 (b) LR genes upregulated in Melanoma compared to SCC and BCC

(c) a random set of genes excluding the genes used in a and b. The melanoma specific genes identified in our study are associated with poor prognosis where the high expressing patients indicate significantly poor survival while the random genes show no significant stratification.

Main Figures

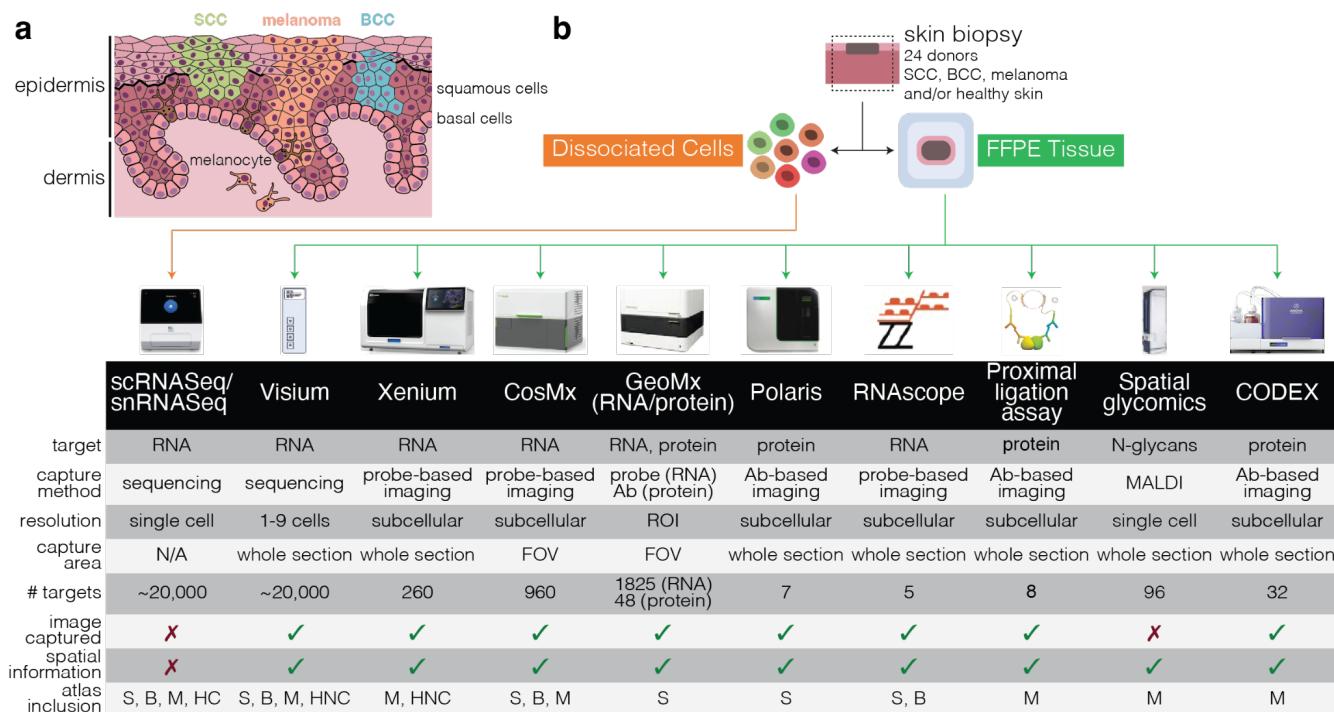


Figure 1. Integrating 12 single cell and spatial technologies to create cell type, community and cell-cell interaction reference atlases for skin cancer.

(a) Simplified cross-section of the human epidermis, highlighting squamous cells, melanocytes and basal cells. Coloured regions represent cSCC (green), which originates from squamous cells, melanoma (orange), which originates from melanocytes, and BCC (blue), which originates from basal cells. Two orange melanocytes are shown in the dermal region as occurs in invasive melanoma; other cells in the lower dermis layer are not depicted.

(b) Overview of sample design and technologies used to generate data for this project. ROI - region of interest; FOV - field of view; S - cSCC; B - BCC; M - melanoma; HC - healthy (cancer patient); HNC - healthy (non-cancer patient donor). Technologies included are single cell RNA sequencing for fresh samples, single nuclei sequencing for formalin-fixed samples, Visium, Xenium, CosMX, GeoMX DSP for whole transcriptome, GeoMX DSP for proteins, Polaris, RNAscope, the proximal ligation assay, , spatial glycomics and CODEX.

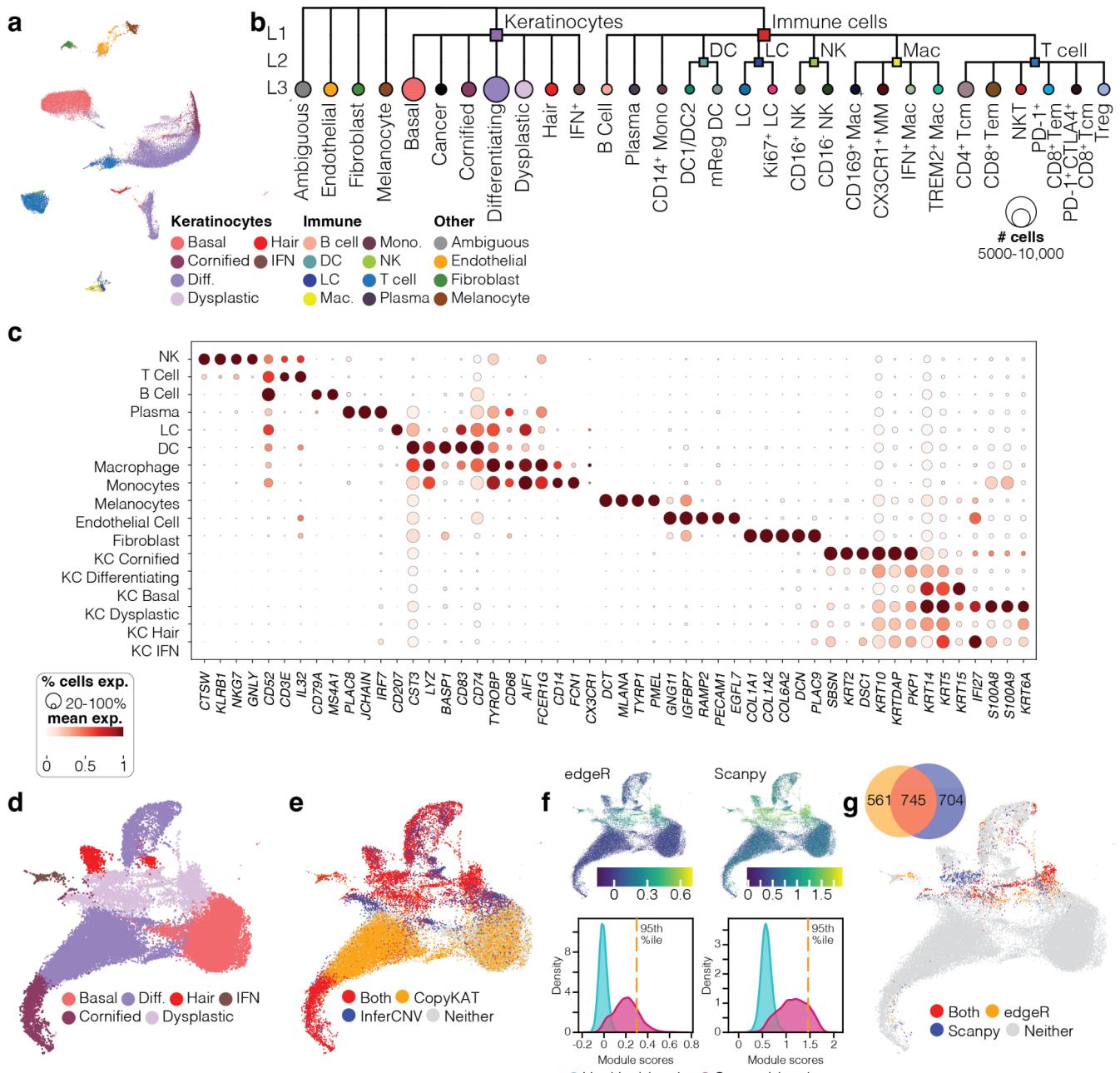


Figure 2. A single-cell atlas of cSCC, BCC and healthy skin cell types

(a) UMAP plot showing the integration of 45,909 healthy and cSCC cells from 11 samples of five patients, indicating results of Level 2 cell type annotation. Seventeen cell types were identified - eight immune cell clusters, six KC clusters, endothelial cells, fibroblasts, and melanocytes - plus an additional cluster of ambiguous cells.

(b) Dendrogram showing the cell classification hierarchy, including Level 3 annotation of immune cells.

(c) Distinguishing markers of 17 Level 2 cell types. Markers are a combination of predicted markers for each cluster, plus known canonical markers for each cell type.

(d) Subclustering of keratinocytes showing the six Level 2 subtypes.

(e-g) Classification of cancerous KC cells. Candidate cells were first classified as being aneuploid (red) if both InferCNV and CopyKat predicted them to be as such (e). Cells were then assigned an “cSCC score” (Module score calculated based on the cumulative expression of genes differentially expressed in KCs in the cancer samples as compared to those from the normal samples) using differentially expressed genes identified using two different methods, edgeR and scanpy (f). Finally, cells were classified as KC Cancer (g) if they were classified as aneuploid (e) and also received an cSCC score

1868 above the 95th percentile of all cell scores (f). The venn diagram indicates the number of cells passing
1869 the module score threshold by edgeR and scanpy.
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883

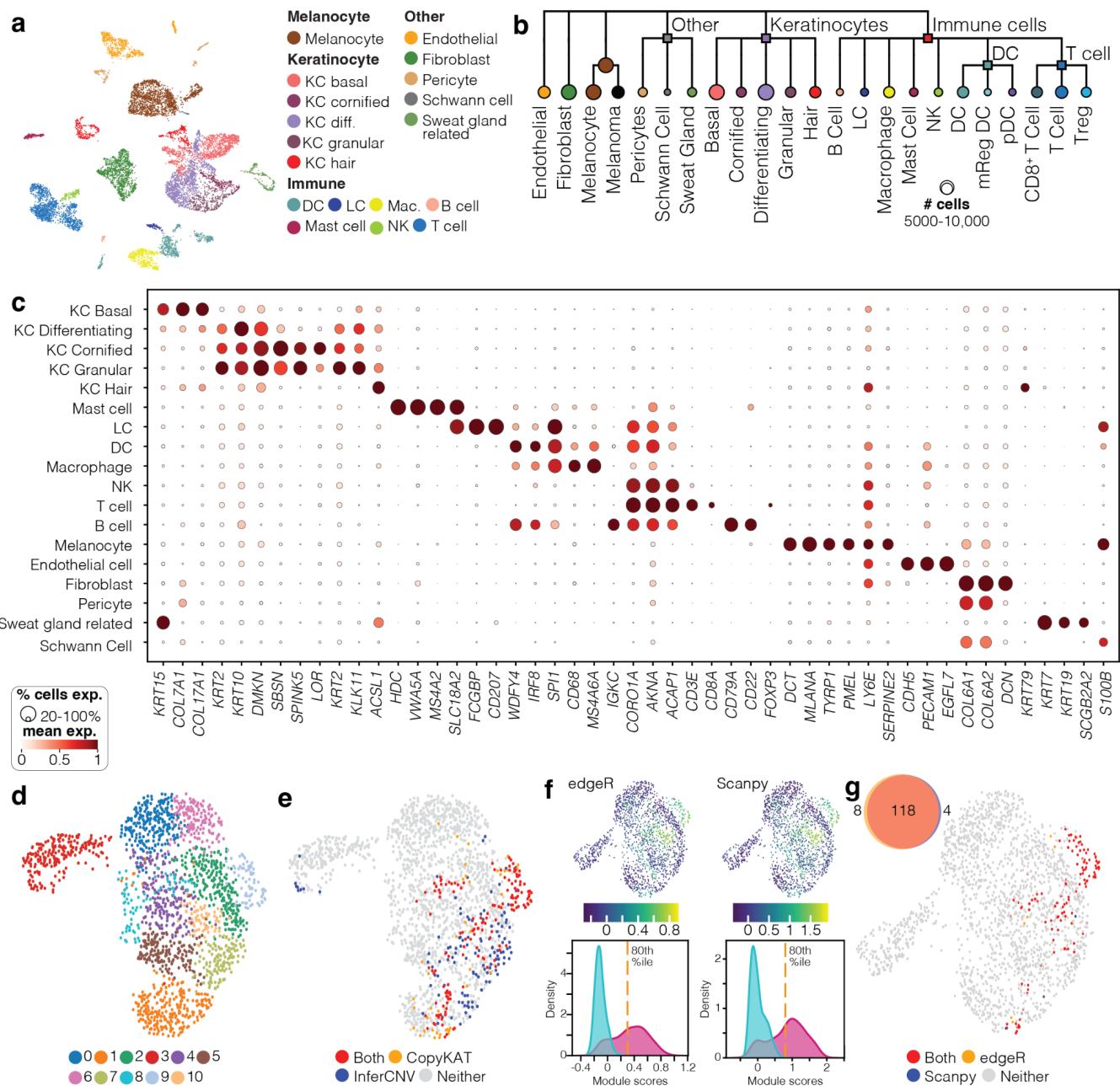


Figure 3. A single-cell atlas of melanoma cell types

(a) UMAP plot showing the integration of 10,747 melanoma cells from three patient samples, indicating results of Level 2 cell type annotation. Eighteen cell types were identified - melanocytes, seven immune cell clusters, five KC clusters, and five other cell types.

(b) Dendrogram showing the Level 2 cell classification hierarchy.

(c) Distinguishing markers of 18 Level 2 cell types. Markers are a combination of predicted markers for each cluster, plus known canonical markers for each cell type.

(d) Result of Level 2 reclustering and cell type annotation for melanocytes.

(e-g) Results for classification of cancerous melanoma cells. Melanocytes from the patient with the malignant tumor were classified as likely melanomas if they were both predicted to have aneuploid genomes (red) by both InferCNV and CopyKat (e). Cells were then assigned a "melanoma score" (f). Specifically, a module score was computed using genes upregulated in the melanoma sample compared to the benign sample using both edgeR pseudobulking and scanpy non-parametric test. For the sample from melanoma patient, a majority of the cells with a score >80th percentile cut-off were from the Melanoma sample cluster (Clusters 9 and 10), (g) and finally the cells inferred 'Aneuploid' by the CNV

1900 analysis and with a high module score by both the aforementioned methods are labelled as malignant
1901 melanocytes (red) as shown in the UMAP. The venn diagram indicates the number of cells passing the
1902 module score threshold by edgeR and scanpy.
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916

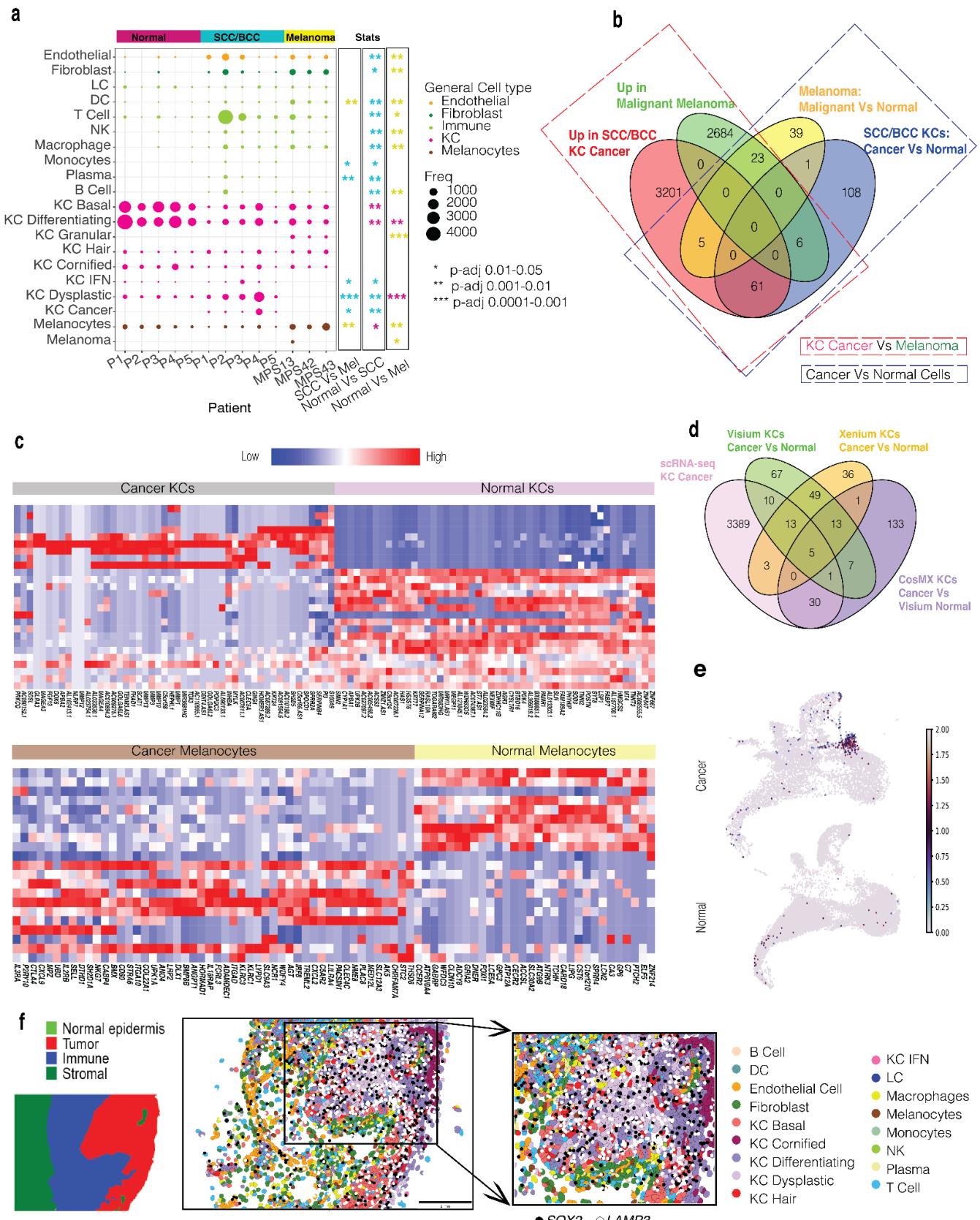
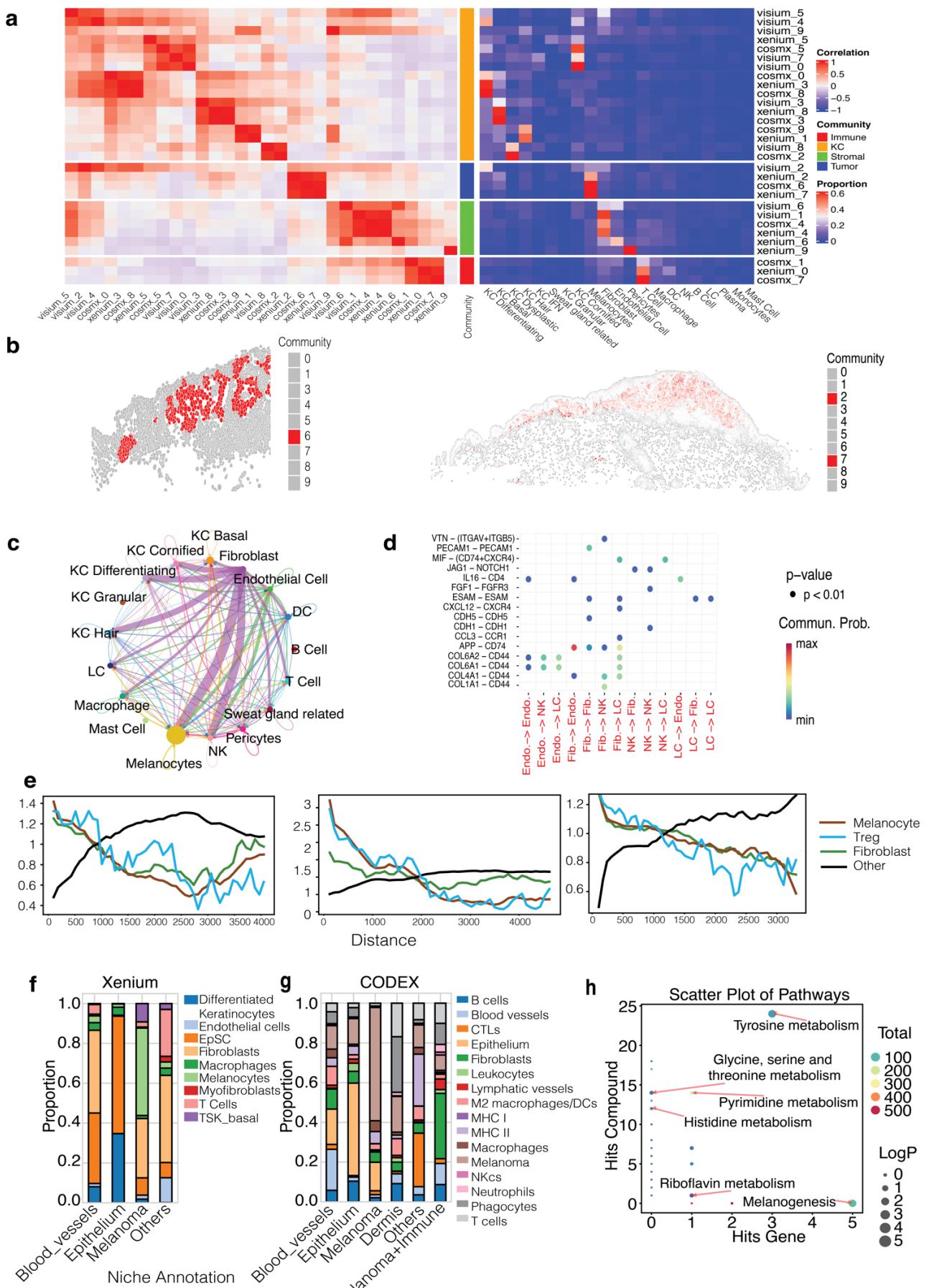


Figure 4: Biological distinctions between cSCC, BCC and melanoma in the single-cell RNASeq atlas

(a) Dot plot showing the percentage of each Level 2 cell type within patient samples. Dots are coloured by cell type category and dot size indicates their percentage within each sample; all columns sum to 100. Results of differential abundance statistical tests are shown to the right, comparing abundance in cSCC vs melanoma, cSCC vs healthy skin, and melanoma vs healthy skin. Asterisks indicate the sample

1924 in which the cell type was found to be more abundant, either healthy skin (pink), cSCC-BCC (blue) or
1925 melanoma (yellow).
1926 (b) A venn diagram of the top significant upregulated genes across cancerous and non-cancerous KCs
1927 and melanocytes. (red) Upregulated in cSCC/BCC KC Cancer cells compared to Malignant Melanocytes
1928 from melanoma samples, (green) Upregulated in Malignant Melanocytes from melanoma samples
1929 compared to cSCC/BCC KC Cancer cells, (yellow) Upregulated in Malignant melanocytes compared to
1930 other melanocytes in melanoma samples, (blue) Upregulated in cancer KCs compared to other KCs in
1931 cSCC/BCC sample.
1932 (c) Heatmaps showing top 50 differentially expressed genes across Cancer vs Normal KCs (top left),
1933 Melanocytes vs Melanoma (bottom). Each column of the heatmap indicates a pseudo-bulked pool.
1934 (d) Integrative, multiple platform analysis of differentially expressed genes. From left to right, the Venn
1935 diagram shows the overlap between DE genes between cSCC cancer KCs vs normal KCs across
1936 scRNAseq and for KCs in cancerous tissues compared to those from the normal tissues from non-
1937 cancer donors with spatial datasets of Visium, Xenium and CosMX.
1938 e) UMAP plot for scRNAseq data showing the expression of SOX2 in cancer vs non-cancer samples,
1939 which matches the location of KC cancer cells in UMAP shown in **Fig 2**.
1940 f) Tissue gene expression plot of CosMX data showing two of the five shared markers SOX2 and
1941 LAMP3. Pathological annotation of the region is shown on the left.
1942



1943
1944
1945

Figure 5. Integrative analysis across different individuals and technologies to build a robust spatial community atlas.

1946 (a) Cross-modality comparison of the ten communities identified for each of Visium, CosMx and Xenium.
1947 Each row shows a community identified from one of the three spatial platforms. The left heatmap shows
1948 similarity across communities within and between technologies, measured by pairwise Pearson
1949 correlation values between communities based on their cell type composition. This allows similar
1950 communities across technology platforms and samples to be grouped to form meta-communities. The
1951 right heatmap shows the cellular makeup of each community (i.e. proportion of each cell type per
1952 community), providing information to label the groups of communities. The central annotation shows the
1953 broad classification of communities into immune, KC, stromal or tumour-related communities, based on
1954 the cellular makeup of each.
1955 (b) Spatial localisation of cells belonging to communities CosMx_6 (left) and Xenium_2 and Xenium_7
1956 (right). Together with Visium_2, these communities form a meta-community that is enriched for
1957 melanocytes.
1958 (c) Inter-community communication within melanoma CosMx_6. The chord plot visualises cell-cell
1959 communication mediated by Collagen signaling pathways, using the CellChat pathway database. Lines
1960 connect communicating cell types; line thickness represents greater communication between cell pairs.
1961 (d) Ligand-receptor interactions between pairs of cell types within the melanoma community CosMx_6.
1962 Top significant L-R pairs and corresponding cell type pairs are shown.
1963 (e) Cell type co-occurrence in CosMx samples between melanocytes and either other melanocytes
1964 (brown), Treg cells (blue), fibroblasts (green) or other cells (black). Each line plots the co-occurrence
1965 score (y-axis) between melanocytes and the test cell type calculated over increasing spaital distances
1966 (x-axis). The samples from left to right are melanoma 23346-105P, 30037-07BR and 6475-07FC.
1967 (f-g) Cell type proportions of communities identified in Xenium (f) and CODEX (g) for adjacent sections
1968 from the same sample (48974-2B). The melanoma community in both datasets is enriched with
1969 melanocytes.
1970 (h) Joint pathway analysis using upregulated genes or proteins of the melanocyte communities in
1971 Xenium and CODEX data (shown in f and g), and highly expressed glycans of the melanocyte
1972 community in MALDI data (shown in Fig S12a). The proteins, genes, metabolites are mapped to KEGG
1973 metabolic pathways. The X-axis shows the number of genes/proteins from Xenium and CODEX data
1974 found in the pathway, while the Y-axis shows glycans in the same pathway.

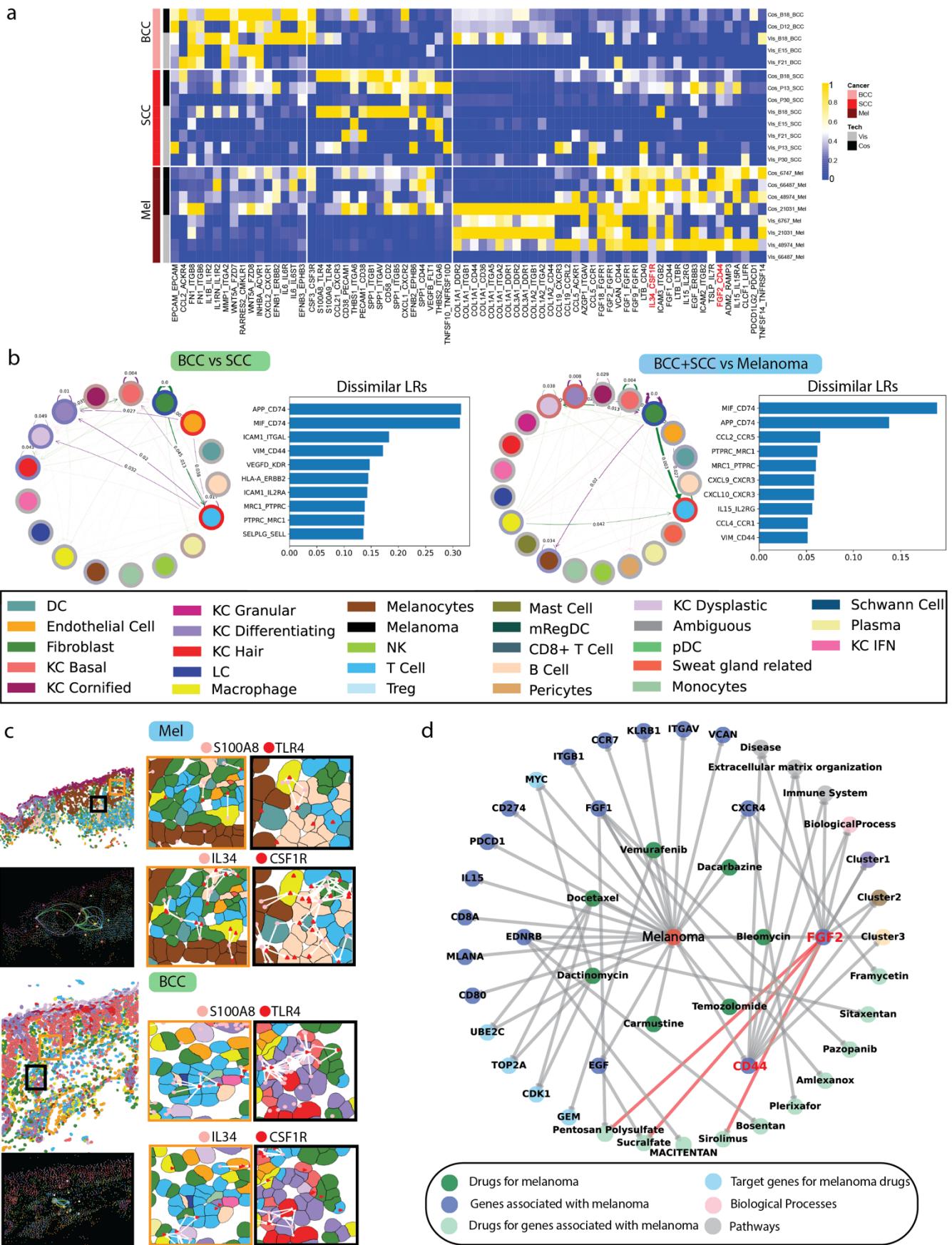


Figure 6: Differential multimodal cell-cell interaction across three cancer types using integrative analysis of Visium, CosMx, Xenium data.

1982 (a) Heatmap of LR scores for LR pairs enriched per cancer type, with a consistent trend across samples
1983 and the two Visium and CosMx platforms. Differentially expressed LR pairs were calculated comparing
1984 each cancer type vs the others using a pseudobulked LR scores with 3 pools per sample. Each heatmap
1985 row is a distinct CosMx or Visium sample. The two L-R pairs specific for melanoma IL34-CSF1R and
1986 FGF2-CD44 were used for experimental validations.
1987 (b) Differential interaction analysis based on LR pairs and cell type pairs. The Venn diagram compares
1988 differential LR pair results between Melanoma and the combined BCC + cSCC datasets, calculated
1989 using edgeR with pseudobulked LR scores. The diagram highlights consistent and unique results
1990 between CosMX and Visium, where Up indicates a higher LR score in Melanoma and Down indicates a
1991 lower score in BCC + cSCC. Cell-to-cell communication between the LR pairs that are up- and
1992 downregulated in melanoma in both CosMx and Visium is shown in the two Network plots flanking the
1993 Venn diagram. In both Network plots, the purple arrows show pairs of cell types that have interactions
1994 higher in Melanoma and green arrows show interactions between cell type pairs more in the BCC +
1995 cSCC than in Melanoma. The number displayed for each arrow shows the integrated p-value across all
1996 biological replicates (the thicker arrows indicate more interactions). Interactions between the two cell
1997 types can still be significantly upregulated in melanoma even if the set of LR pairs were downregulated.
1998 (c) Spatial mapping of cancer type-enriched LR pairs in CosMx data. Two LR pairs that were significantly
1999 different between cancer types across technologies in Panel A are shown, namely IL34-CSF1R (higher
2000 in melanoma) and S100A8-TLR4 (higher in BCC). Both are visualised in FOVs from melanoma sample
2001 (top) and BCC sample (bottom). For each cancer type, the cell type annotation of the FOV is shown (top
2002 left) with orange and black boxes indicating the highlighted regions (right). An overview of interactions
2003 at tissue level is shown by large coloured arrows, representing cumulative interactions between two cell
2004 types in the tissue, with the location of the arrow root as the centroid coordinate of all cells in one cell
2005 type (bottom left). Magnified boxes (right) show the presence of ligands (pink) and receptors (red), with
2006 white arrows showing the connections between ligands and receptors of nearby cells.
2007 (d) Melanoma drug target graph integrating multiple biological and pharmacological knowledge types.
2008 Nodes represent genes, drugs, and biological functions. Level 1 connections show melanoma-
2009 associated genes and drugs targeting melanoma. Level 2 links display drugs targeting the melanoma-
2010 associated genes from Level 1 and a broader gene set targeted by drugs in the network. All genes in
2011 the graph are either upregulated or have high ligand-receptor scores. Clusters 1, 2, and 3 are pathways
2012 enriched with genes shown in the graph.
2013
2014

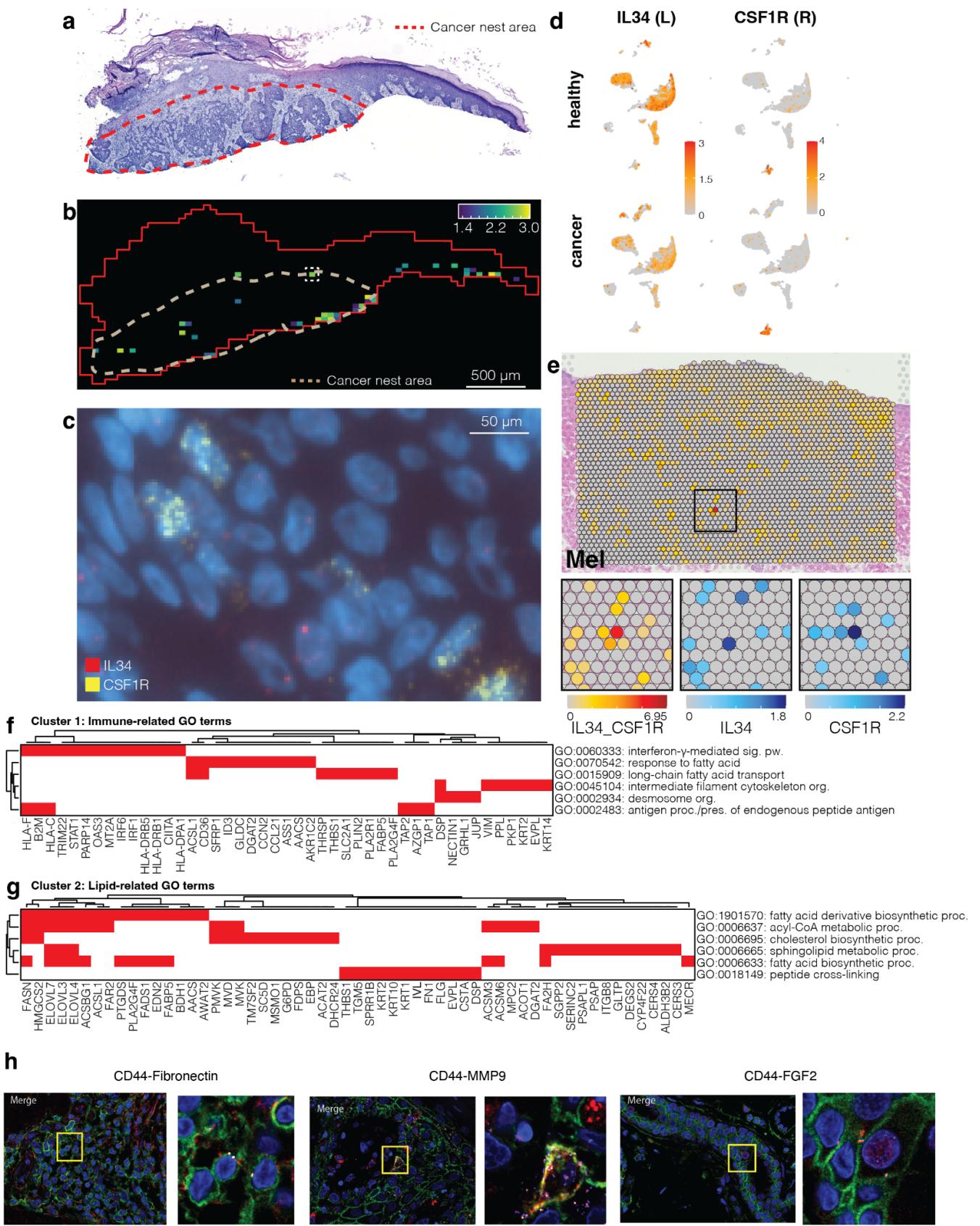


Figure 7. Multi-omics analysis of LR interaction between IL34 and CSF1R

(a) H&E staining of a tissue section from cSCC patient B18, with a pathologist's annotation highlighting a cancer nest region.

(b) Mapping of IL34_CSF1R co-localisation in RNAscope data from an adjacent section of the tissue block to that shown in **Panel a**. Image tiles containing both CSF1R and IL34 RNA were identified using a window scanning strategy in STRISH. The colour scale represents the $-\log(p\text{-value})$ of each window's

2022 *LR co-expression score, while the red box indicates the tissue border. The white box highlights the*
2023 *tissue zone shown in **Panel c**.*

2024 *(c) Magnified region corresponding to the white box in **Panel b**, showing DAPI-stained cell nuclei (blue)*
2025 *with CSF1R (yellow) and IL34 (red) RNA visible as coloured dots.*

2026 *(d) Expression of IL34 and CSF1R in healthy and cSCC samples from scRNASeq data. This interaction*
2027 *was statistically predicted to occur in both healthy and cancer samples in our CellChat analysis.*

2028 *(e) Exemplar spatial plots showing the LR score for IL34_CSF1R from patient 48974. The black box*
2029 *indicates a region highlighted below the main image. Here, zoomed-in boxes show the IL34_CSF1R LR*
2030 *score (left) and IL34 (middle) and CSF1R (right) gene expression for the same tissue region.*

2031 *(f-g) Heatmaps indicating grouped GO terms and associated genes that are enriched in IL34_CSF1R-*
2032 *positive spots in melanoma samples compared to IL34_CSF1R-negative spots. GO term groups were*
2033 *calculated by k-means clustering ($k = 3$) of GO semantic similarity scores; two such groups are shown*
2034 *here. The full heatmap is shown in **Fig S9b**.*

2035 *(h) Proximal ligation assay (PLA) for validating CD44 interactions in melanoma. A merged image of*
2036 *signal for the ligand and the receptor and a zoom-in window highlighting the interaction on the cell*
2037 *membrane. A positive PLA signal is visible if two interacting proteins are in a proximity less than 20 nm.*

2038

2039

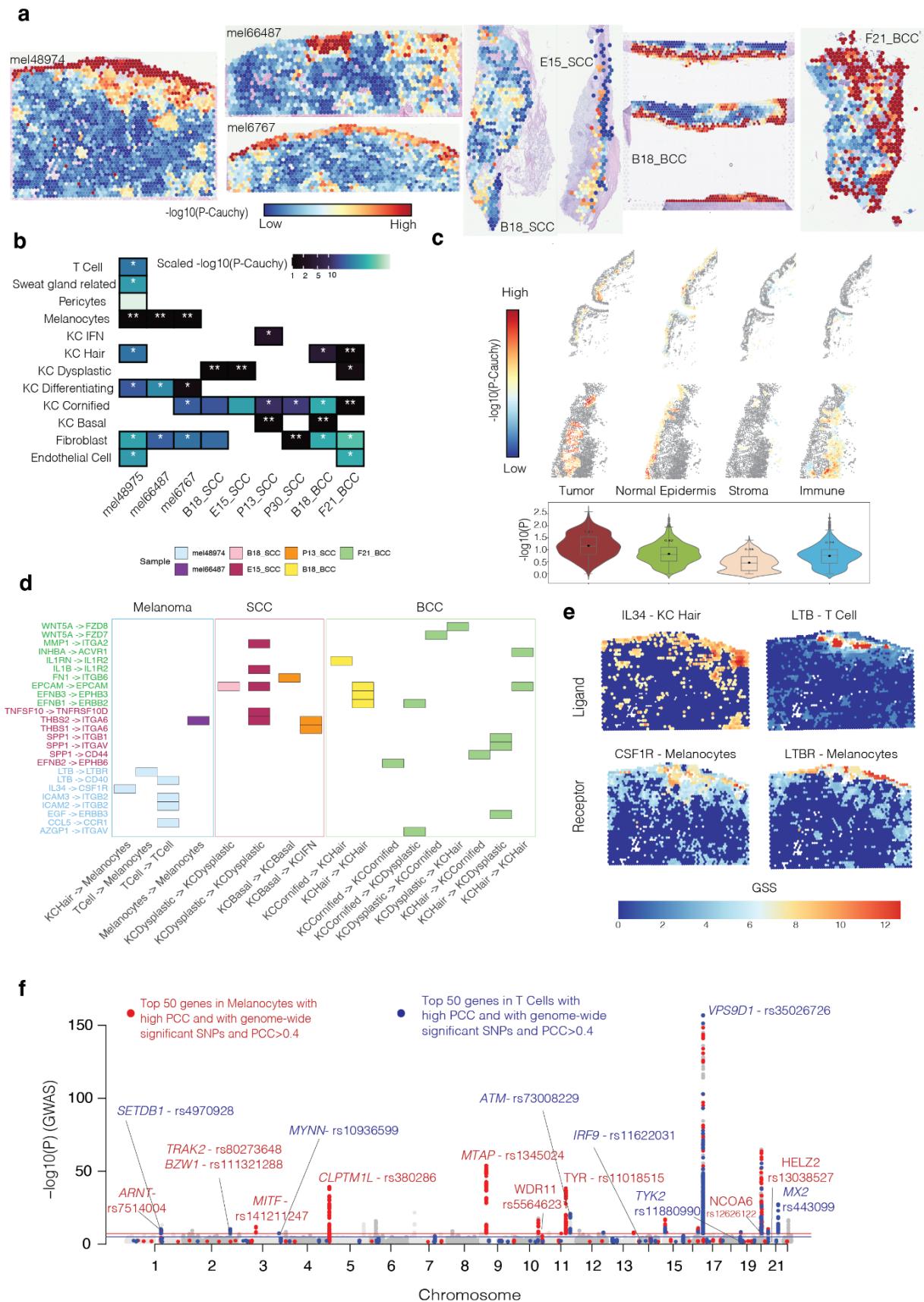


Figure 8. Mapping genetics effects from genome-wide association studies for cutaneous melanoma, cSCC, and BCC to spatial domains and cell types.

(a) Gene specificity score (GSS) and association of spatial spots with skin cancer heritability. GSS score for each gene in a spot/cell represents the enrichment of the gene as a top rank most abundant gene in

2045 the spot/cell and its neighbour spots/cells in an anatomical region, a spatial domain, or a cell type. The
2046 p-value shows the spatial heritability enrichment significance of a spot with a trait based on SNPs
2047 mapped to the genes with high GSS scores (one-sided Z-test for stratified coefficient different to 0). The
2048 p-value is more significant if the SNPs that are mapped to the high GSS genes explain a higher
2049 proportion of heritability for the trait.

2050 (b) Cell types with the highest enrichment of heritability explained by SNPs tagged to GSS genes of
2051 cells in a cell type. The white asterisks indicate the most enriched cell-type for heritability of cutaneous
2052 melanoma, cSCC and BCC traits.

2053 (c) gsMAP significance spatial heritability enrichment is shown at single-cell resolution across the tissue
2054 (upper tissue plots) or per annotated skin regions (lower violin plots) from the cosMx data of the sample
2055 mel48974.

2056 (d) LR pairs with significant association with SNP heritability explained by the corresponding cell types.
2057 The rectangles show cases where both L and R genes had PCC >0.3 between GSS of the gene and
2058 the gsMAP P-values (the significance level for the LD stratified coefficients for the spot bigger than 0).
2059 The results suggest which LR pairs are related with the heritability of a cell type pairs.

2060 (e) GSS of two LR pairs showing specificity of the L and R genes to tissue regions at the immune-rich
2061 dermal layers and the epidermis of the skin.

2062 (f) Manhattan plot showing top significant GWAS SNPs co-localizing with genes in melanocytes (red)
2063 and T cells (blue) that had the highest Pearson correlation between GSS and the gsMAP trait
2064 association P-value or associated with SNPs with genome-wide significance. The Y-axis shows the -
2065 log(P-value) from GWAS analysis.