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

P. Prakrithi[1,2#], Laura F. Grice[1,2,3#], Feng Zhang[1,2#], Levi Hockey[1,2], Samuel X. Tan[4], Xiao Tan[1,2], Zherui Xiong[1,2], Onkar Mulay[1,2], Andrew Causer[1,2], Andrew Newman[1,2], Duy Pham[1], Guiyan Ni[1], Kelvin Tuong[5], Xinnan Jin[1,2], Eunju Kim[3,4], Minh Tran[1], Hani Vu[1,2], Nicholas M. Muller[4], Emily E. Killingbeck[6], Mark T. Gregory[6], Siok Min Teoh[1],Tuan Vo[1], Min Zhang[7], Maria Teresa Landi[8], Kevin M. Brown[8], Mark M. Iles[9], Zachary Reitz[6], Katharina Devitt[10], Liuliu Pan[6], Arutha Kulasinghe[6], Yung-Ching Kao[4], Michael Leon[6], Sarah R. Murphy[6], Hiromi Sato[6], Jazmina Gonzalez Cruz[10], Snehlata Kumari[10], Hung N. Luu[11], Sarah E. Warren[6], Chris McMillan[12, 13], Joakim Henricson[14,15], Chris Anderson[14,15], David Muller[12,13], Arun Everest-Dass[16], Blake O'Brien[17], Mathias Seviiri[16], Matthew H. Law[18, 19, 20], H. Peter Soyer[4], Ian Frazer[10], Youngmi Kim[6,21], Mitchell S. Stark[4], Kiarash Khosrotehrani[4], Quan Nguyen[1,2]

[1] QIMR Berghofer, Immunology and Infection program, Brisbane, QLD 4006, Australia
[2] Division of Genetics and Genomics, Institute for Molecular Bioscience, The University of Queensland, QLD 4072, Australia
[3] School of Biomedical Sciences, The University of Queensland, QLD 4072, Australia
[4] Frazer Institute, Dermatology Research Centre, Experimental Dermatology Group, The University of Queensland, Brisbane, Australia
[5] Ian Frazer Centre for Children's Immunotherapy Research, Child Health Research Centre, Faculty of Medicine, The University of Queensland, Brisbane, Queensland, Australia
[6] Bruker Spatial Biology, NanoString® Technologies, Seattle WA 98109, USA
[7] School of Agriculture and Food Sciences, The University of Queensland, QLD 4072, Australia
[8] Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA
[9] Leeds Institute for Data Analytics, University of Leeds, Leeds, UK
[10] Frazer Institute, The University of Queensland, Brisbane,QLD 4072, Australia
[11] UMPC Hillman Cancer Center & School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA
[12] School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, QLD 4072, Australia
[13] Australian Infectious Diseases Research Centre, School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, QLD 4072, Australia
[14] Department of Emergency Medicine in Linköping, and Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden
[15] Department of Dermatology and venereology in Östergötland, and Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden
[16] Institute for Biomedicine and Glycomics, Griffith University, Gold Coast, QLD 4215
[17] Sullivan Nicolaides Pathology, Bowen Hills, Qld 4006, Australia
[18] QIMR Berghofer, Population Health program, Brisbane, QLD 4006, Australia
[19] School of Biomedical Sciences, Faculty of Health, Queensland University of Technology, Brisbane, QLD, 4001, Australia
[20] School of Biomedical Sciences, Faculty of Medicine, University of Queensland, Brisbane, QLD, 4072, Australia
[21] Pfizer, 21621 30th Dr SE, Bothell, WA 98021

[#]co-first authors
Correspondence: quan.nguyen@qimrberghofer.edu.au

Abstract

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

Background

Although skin cancer is the most common neoplasm, its major subtypes, basal cell carcinoma, squamous cell carcinoma, and melanoma, are rarely compared at the cellular and molecular level, limiting our understanding of their shared and distinct features. The keratinocyte cancers basal cell carcinoma (BCC; ~75-80%) and cutaneous squamous cell carcinoma (cSCC; ~20%), are the most prevalent malignancies. Though they have low mortality rates, their high incidence results in substantial public health burdens, with ~3.5 million cases treated annually in the U.S. at a cost of ~$4.8 billion (Rogers et al., 2015; Guy et al., 2015). About a third of skin cancer-related deaths in Australia are due to keratinocyte cancer (Czarnecki, 2024). In contrast, melanomas account for less than 10% of diagnosed skin cancers but exhibit a markedly higher case fatality rate, constituting the majority of all skin cancer-related deaths (Siegel et al., 2022).

All three major skin cancer types originate in the epidermis, with BCC and cSCC arising from keratinocytes (KCs) and melanoma from melanocytes (**Fig 1a**). Although foundational single-cell atlases have characterised the cellular composition of BCC (Guerrero-Juarez et al., 2022; Yerly et al., 2022, Huang et al., 2023; Ganier et al. 2024), cSCC (Ji et al., 2020; Yan et al., 2021; Zou et al., 2023) and melanoma (Tirosh et al., 2016; Jerby-Arnon et al., 2018; Karras et al., 2022; Pozniak et al., 2024) individually, a comprehensive comparison of their single-cell architecture within the tumour microenvironment has yet to be undertaken. For instance, despite common environmental risk factors such as solar UV exposure, and the many overlapping genetic and anatomical attributes, outstanding questions remain about which factors determine the initiation and progression of skin cancer and why most cSCCs and BCCs are less invasive than melanoma (Wu et al., 2014; Landi et al., 2020; Seviiri et al., 2022). Available single-cell datasets for melanoma were produced by early iterations of single cell

sequencing technologies capturing fewer cells (Tirosh et al., 2016; Zhang et al., 2022). These studies mostly described acral melanoma, uveal melanoma, or *in vitro* melanoma cell lines (as reviewed by Lim and Rigos, 2024). There are no single cell RNA sequencing (scRNAseq) datasets available for melanoma that have representative cell types spanning from benign to melanocytic lesions and invasive melanoma. Although prior scRNASeq studies have included matched healthy and tumour skin (Ji et al., 2020; Yan et al., 2021; Zou et al., 2023; Huang et al. 2023), and compared BCC and SCC signatures (Yost et al. 2019), to our knowledge, no single cell sequencing data is available for matched healthy skin, cSCC, and BCC lesions from the same individuals, an approach that enables direct within-patient comparison across disease states. Spatial data for skin cancer at single cell resolution with whole transcriptome coverage, or with multimodal measurements of RNA, proteins and metabolites is not available.

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

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

To accelerate advances in treatment, beyond current excision as the standard of care option, drug therapies that can permanently cure these skin cancers are highly desirable but are lacking and so methods to find new targets are needed. For melanoma patients, a common treatment regimen includes targeted therapies (e.g. BRAF- and MEK-inhibitors) and immunotherapies (e.g. anti-CTLA-4 and anti-PD-1). However, only 30-50% of advanced-stage melanoma patients respond to immunotherapies (Sharma et al., 2015; George et al., 2017; Herrscher et al., 2020), and these treatments can cause severe immune-related adverse events (Fink et al., 2021). Ideally, such treatments should only be offered to patients who are likely to respond, but it is not yet possible to confidently stratify patients. Similarly, only approximately 50% of advanced-stage cSCC patients

respond to immunotherapy treatment (García-Sancha et al., 2021). To better harness the potential of immunotherapies for melanomas and other skin cancers, it is necessary to identify new ligand-receptor (LR) targets as actionable (e.g. combinational therapies with PD-1/PD-L1). This, in turn, requires an understanding of cancer-immune cell interactions underlying the basis of carcinogenesis.

Here, we applied the latest single-cell and spatial technologies to map skin cancer cells, cell-cell interactions, and microenvironments. Specifically, we aimed to address keratinocyte cancers (cSCC and BCC) tumour micro-environment that vastly differs from melanomas. We integrated 12 distinct yet complementary technologies to comprehensively identify spatial, single-cell signatures and compare spatial cell-cell interactions (CCIs) to find shared and unique interactions involved in initiation and progression in the three skin cancer types. By harnessing the power of spatial multiomics, this study has provided valuable data resources to understand gene signatures, cells, spatial communities, and cell-cell interactions for the most common cancers and generated new understanding of cell types and their activities that are distinct between the three skin cancer types.

Results

Ultraplex, multimodal, multiplatform, single cell and spatial omics data resource for skin cancer

To understand the cellular microenvironment of the skin, we leveraged data from 12 molecular technologies to curate a spatial, single-cell atlas and interactome of healthy (non-sun-exposed) and cancerous skin. Biopsies were collected from 24 skin donors, consisting of patients diagnosed with cSCC (n = 7), BCC (n = 4), and melanoma (n = 7, including n = 3 for snRNAseq) and from non-cancer donors (n = 3, Visium and Xenium) (**Table S1**). Additional healthy skin samples were collected from non-sun exposed skin of cSCC patients (n = 5) (**Table S1**). The diagnosis and biopsy details are described in the Methods section. Each biopsy was measured by up to 12 technologies: Chromium single-cell RNA sequencing (scRNASeq), FLEX single nuclei sequencing (snRNAseq), whole transcriptome Visium spatial transcriptomics, single-cell resolution Xenium spatial transcriptomics, NanoString CosMx Spatial Molecular Imaging (CosMx), NanoString GeoMx Digital Spatial Profiling for proteins (Immune-oncology panel, GeoMx), GeoMx cancer transcriptome atlas (GeoMX CTA), Opal Multiplex Polaris protein assay, RNAScope RNA *in situ* hybridisation, Proximal Ligation Assay (PLA), MALDI-TOF spatial glycomics, and CODEX spatial proteomics (**Fig 1**). In the following sections, we present the results in three themes, including a single cell and spatial atlas (**Fig1** to **Fig 4**), interaction atlas (**Fig 5** to **Fig 7**) and integration with population genetics data (**Fig 8**).

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

We first sought to generate a high-quality single-cell cSCC atlas and compared transcriptomic shifts between non-cancer and cancer skin and signatures of cancer cSCC cells. We performed scRNASeq on 11 paired samples from healthy (non sun exposed) and cancer biopsies collected from five cSCC patients. A total of 45,909 skin cells passed quality control. Initial clustering identified five major clusters (“Level 1” annotations) comprising endothelial cells, fibroblasts, melanocytes, keratinocytes, and immune cell types (**Fig 2a-c, Fig S2**). To capture the complexity of the immune microenvironment, a second round of clustering and annotation was performed to find eight major immune subsets (“Level 2” annotations; **Fig 2b-c**) and further identified 19 immune cell (sub)types (“Level 3” annotations; **Fig 2b, Fig S2a,c**), representing six T cell subsets, four macrophages, two NK, two Langerhans cells (LC), two DC, B cells, plasma cells and CD14+ monocytes. Next, we characterised KCs, the dominant cell type within our cSCC scRNAseq dataset (70.6% of cells). We assigned the

following labels based upon abundant gene expression: Basal KCs (*KRT15+*, *KRT14+*), Differentiating KCs (*PKP1+, KRT10+*), Dysplastic KCs (S100A8+, S100A9+), KC interferon (*IFI27+*), KC cornified (*SBSN+, KRT2+, DSC1+*) and KC hair (combination of *KRT6B+, KRT17+, KRT16+, KRT5+*), (**Fig 2c, Fig 2Sb**). Our annotation for scRNAseq is supported by mapping the cells to the skin tissues, where the cell labels are transferred from scRNAseq to CosMX data (**Fig S2d**). The integration and comparison between scRNAseq and CosMX data provided evidence for accurate cell type annotation based on expected distribution of cell types to skin layers.

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

To date, little is known about transcriptional signatures of KC cancer cells at single cell resolution. We developed a stringent pipeline to map KC cancer cells (**Fig 2d-g**). A cell was considered a cancer KC cell if the cell had abnormal polyploidy based on CNV analysis (using consensus results from two separate CNV inference methods), and had high cancer module scores as calculated for genes that were upregulated in tumour compared to normal tissues (**Fig 2d-g**). This way, we identified a total of 745 KC cancer cells. Although we did not *a priori* restrict KC cancer cells to belong to a specific KC subtype, the majority (82.6%) of these cells were classified among dysplastic KCs, further supporting that they are most probably cancer cells from the intersection of three lines of evidence. This stringently defined 'cancer' KC population enabled the analysis of their gene signatures as discussed later.

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

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

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

Limited single-cell datasets for cutaneous melanoma were available (Tirosh et al., 2016), a mixed acral and cutaneous melanoma study (Zhang et al., 2022), and uveal melanoma and *in vitro* melanoma cell lines (as reviewed by Lim and Rigos, 2024). We opted to produce the latest and more comprehensive reference of melanoma cell types using single-nuclei sequencing of formalin fixed tissues. Three patient samples were selected and scored by 23 pathologists, with one sample defined as definitely invasive melanoma (5591 cells), another as benign naevus (3250 cells) and one severely dysplastic naevus sample (1906 cells). From 10,747 single nuclei, we identified 11 immune cell types and five KC types, as well as endothelial, fibroblast, pericyte, Schwann cell, and sweat gland clusters (**Fig 3a-c,**

Fig S4). Importantly, we used multiple lines of evidence to distinguish melanoma cells from melanocytes **(Fig 3d-g)**. Here, similar to the approach to define KC cancer cells in cSCC, we integrated CNV analysis, module scores, and spatial mapping of melanocytes to identify 118 melanoma cells with high confidence (**Fig 3d-g**). These cells were used to find signatures of melanomas as described below.

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

Overall, we found more transcriptomic variation between cancer cells from different lineages (3257 genes higher in cSCC than in melanomas, and 2713 upregulated in melanocytic lesions than in cSCC) than within a lineage, (176 genes upregulated in cancer KCs vs healthy KCs and 68 upregulated genes in melanoma cells compared to melanocytes) (**Fig 4b**). The 3257 genes higher in cSCC compared to melanomas were enriched for Myc targets, E2F pathways, G2M pathways, mTOR signaling and DNA 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 4**e) 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*). Markers of differentiation were also upregulated, such as SOX10 (melanocyte development), *FLT3* and *IRF8* (DC differentiation). Genes associated with immune evasion, including *CTLA4, CD274 (PD-L1), CXCL9, CXCR3,* and *ITGAD (CD11d)*, were also elevated. Additionally, genes linked to melanoma progression and invasion, such as *UBD (FAT10),* and *STC2*, were upregulated, as were genes associated with melanin synthesis such as *TYR* and *DCT*. In contrast, genes downregulated in melanoma but maintained in normal melanocytes included *CST6* (Cystatin E/M), a known suppressor of melanoma proliferation and migration (Xu et al., 2021).

Spatially map cells in stroma and cancer immune compartments in cSCC

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

Detailed spatial map of 21 scRNAseq/snRNAseq defined cell types to transcriptome-wide Visium data, GeoMX WTA, and panel-based, single-cell resolution CosMX data

First, using cell-type signatures for cSCC, BCC and melanoma from scRNAseq/snRNAseq data, we mapped these cell types to spatial Visium data, including 5x cSCC, 3x BCC and 4x melanoma samples across 9x patients (**Fig S10a**, **Table S1**). We spatially mapped all Level 2 cell types from scRNAseq data to our Visium data, which also matches the pathological/anatomical annotation (**Fig S10a**). The expression of marker genes in the Visium was consistent with histological annotation (**Fig S10**). For example, in melanoma Patient 48974, we observed dark-pigmented cells (i.e. melanophages) at the skin surface of the biopsy (**Fig S10a,** *top right*). The corresponding Visium spots in this region were predicted to contain a large proportion of melanocytes, based on our deconvolution results (**Fig S10a**, *middle*). These spots also expressed the highest amount of melanocyte marker *MLANA* (**Fig S10a,** *bottom*). This result demonstrates the concordance of multiple layers of

information (such as tissue morphology, expression of individual genes, and broader transcriptional signatures), establishing an accurate and comprehensive data resource for the later analysis of gene markers and all interactions.

Beyond the resolution at the regions of interest (GeoMX WTA and GeoMX protein) or spot level (Visium), we next investigated cell types by CosMx data (He et al., 2022) that provides spatial information of 131,804 cells (34.4% cSCC, 34.3% BCC, 31.3% melanoma; **Table S1**). By computational annotation, these cells were assigned to 21 cell types that can all be mapped to the tissue, demonstrating that spatial cell labelling was highly consistent with pathological annotations, but much more detailed **(Fig S10b)**. CosMX data mapped scRNAseq cell-type signatures (e.g., keratinocytes, melanocytes, fibroblast and immune cells) to distinct layers in the skin, corroborating the accuracy of our scRNAseq cell-type annotation (**Fig S3**, **Fig S10b**). The subcellular capture resolution of CosMx technology allowed the visual confirmation of cellular location of individual RNA molecules. For instance, in cSCC and BCC, 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 single cell expression of nearly 500 ligand/receptor genes, is valuable for downstream interaction and community analyses as described later.

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

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

Multimodal mapping of communities in melanoma samples with spatial RNA, protein and glycan omics

Next, we compared the annotation for the same melanoma tissue block using three single cell resolution platforms representing three modalities, the Xenium (RNA), CODEX (protein) and mass cytometry imaging (MALDI MSI for spatial glycomics), (**Fig S10, Fig S12**). The expression of molecular markers is shown as dot plots in **Fig S12** and the clustering analysis for each modality is shown in **Fig S13-S15**.

For all three modalities, the melanoma/melanocytes could be distinctly identified, using label transfer and clustering analysis. We observed the localisation of melanoma markers to the melanocyte layer, for example S100B (CODEX protein), and *S100B* and *MKI67* (CosMX), and *TYR* gene (Xenium), (**Fig S10b-f**). With 260 genes in the skin cancer panel, the single cell Xenium data could define 17 cell types where their locations matched the pathological annotation of the tumor and immune cells (**Fig S10c**). All key KC cell types were mapped with the Xenium data. In contrast, the CODEX data could not map KC cells, due to lack of protein markers for these cell types, but could clearly pinpoint additional immune cell types such as the M2 Macrophages and Neutrophils (**Fig S10d**). Although the

cell-type clustering using glycomics is less defined, it is clear that melanoma/melanocytes exhibited unique metabolomic signatures compared to other cell types in the remaining clusters (**Fig S10f, Fig S12a**).

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

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

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

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

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

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

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

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

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

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

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

For BCC, among the LR pairs enriched, we found strong signals for stromal remodelling, especially angiogenesis. Interleukins (*IL6-IL6R, IL6-IL6ST, IL1B-IL1RN-IL1R2*), chemokines (*CXCL2-CXCR1, CCL2-ACKR4*) and fibronectins (*FN1-ITGB8* and *FN1-ITGB6*) all have roles in angiogenesis and our data show that they were higher in BCC (Villani et a., 2021). Interactions in the canonical WNT signaling pathways (*WNT5A-FZD7* and *WNT5A-FZD8*; also supported by *EPCAM-EPCAM* interactions) appeared to be more active in BCC. This signaling pathway, although less well known compared to the Hedgehog (HH) signalling in BCC, is crucial in coordinating with the HH pathway to

maintain the proliferative state of BCC cells and sustain cancer stem-like cells in BCC (Yang et al., 2008).

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

For melanoma, we found strong enrichment of collagen interactions (e.g., *COL1A1*, *COL1A2*, *COL3A1*; making 15 out of all the 37 LR pairs that were specific for melanoma compared to BCC and cSCC). We found type I collagen (*COL1A*, *COL1B*) and integrin receptor (ITGA, ITGB) families to actively interact in melanoma. These included *COL1A1-ITGA2*, *COL1A1-ITGB1*, *COL1A1-ITGA5*, *COL1A2-ITGA2*, and *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 pro-tumorigenic 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**) and the top LR pairs associated with T cell and melanocytes are shown in **Fig S20f**. 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 6a**). 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 S21**). 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 Supplementary Methods). This observation is consistent with spatial single-cell gene expression data (**Fig 6c**). Both *IL34* and *CSF1R* expression were detected in our healthy and cSCC scRNASeq atlas. Low spatial resolution, transcriptome-wide Visium showed the colocalization of *IL34-CSF1R* in the dermis layer (Fig 7a) . Cells expressing the two genes are visualized on single-cell level resolution spatial data from STOmics and Curio-Seeker (Takara Bio, USA) melanoma samples and appear to be in spatial proximity (Fig 7b). Additionally, pathway comparisons for interacting Visium spots with those spots without interactions show enrichment of the antigen processing pathway and lipid metabolism (**Fig 7c, d**). We detected 531 genes upregulated in *IL34-CSF1R* positive spots in melanoma, and 758 genes for BCC. Only the melanoma gene list was found to show any enrichment for GO terms), including terms associated with immune- (**Fig 7c**) and lipid-related (**Fig 7d**) functions, suggesting the important roles of the upregulation of the *IL34-CSF1R* interactions in melanoma.

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

Beyond RNA level, we also validated protein-protein interactions based on colocalization between neighbour cells (**Fig S22**). We implemented Opal tyramide signal amplification (TSA) protocol, where primary antibody, anti-IgG polymer HRP and covalent labelling with Opal TSA fluorophores were used and the whole slide multispectral scanning was performed by Vectra Polaris. With this method, we scanned through the whole tissue section and identified image tiles containing double-positive CD8+ PD-1+ immune cells and PanCK+ PD-L1+ cancer cells (**Fig S22**).

The positive colocalization from Opal Polaris suggests evidence for interaction, but the distance can be relatively further away than direct interaction at one location. We next applied proximal ligation assay to detect two proteins within a 20 nm distance on melanoma samples (**Fig 7e**, **Fig S23**). We tested three pairs identified in our spatial omics data analysis, focusing on CD44, a dominant receptor found with distinctively more common interactions in melanoma compared to BCC and cSCC (**Fig 7e**). CD44 acts as an MMP9 docking receptor that localizes MMP9 to the cell surface, where it can degrade components of the ECM such as collagen to enhance tumour invasion (Yu and Stamenkovic, 1999). CD44 was reported to bind fibronectin (FN) to anchor cells to their surrounding ECM, potentially supporting invasion. CD44 interacts with Fibroblast growth factor (FGF1 and FGF2) in melanoma,

possibly enhancing tumour initiation and migration. Here we validated CD44-MMP9, CD44-FN1, and CD44-FGF2 interactions. The PLA signal clearly suggests the interactions occur (**Fig 7e**). Future work will validate more pairs where antibodies are available. Moreover, we showed that these LR pairs had significant prognostic values when applying to the public TCGA data (**Fig S24**), further suggesting the roles of cell-cell interactions to cancer phenotypes.

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

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

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

Next we identified ligand-receptor genes with significant correlation between spatial gene specificity scores and genetic association P value for the cell type. Such pairs suggest possible mechanisms on how the genetic association of signals in a spatial region or a cell type may be explained through dynamic interactions between cell type pairs within the spatial microenvironment. Again, we observed strong interactions between KC dysplastic in cSCC and KC-hair in BCC, and more immune interaction in melanoma, especially those involved in T cells compared to in BCC (**Fig 8d**). Visualisation of genetics association signals between T cells with Melanoma via *IL34-CSF1R* and *LTB-LTBR* suggests

tissue regions where the association with the melanoma were most strong (**Fig 8e**). These regions are at the junction of the epidermis and dermal regions for *LTB* and *LTBR* genes, and less specific for *IL34-CSF1R*. Across the whole genome, the top genetic association with spatial expression patterns is consistent with mapping key melanoma markers such as *MITF*, *TYR*, and *MX2*. The *CSF1R*, *LTB* and *LTBR* were in the top 50 genes with the highest correlation between spatial specificity for melanoma and spatial expression (**Fig 8f**). Genome-wide significant SNPs associated with genes having the highest spatial specificity for T cells or melanoma are mostly SNPs with the strongest significant values in the Manhattan plot. This suggests that the heritability of melanoma cancer risk may be exerted from effects on T cells and melanomas (**Fig 8f)**. Overall, our integrative analyses with genome-wide association studies suggest consistent and generalizable patterns.

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

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

Discussion

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

Strong evidence was drawn from integrating 12 orthogonal technologies, complementary in resolution, sensitivity, and throughput, adding spatial information and quantifying both RNA, protein and glycan modalities. Our work here reports a comprehensive guideline for other spatial multiomics studies, beyond skin cancer, to assess benefits and limitations for each tool and design a strategy to select suitable tools and combine analyses to cross-validate and gain more biological information than is attainable through any single technology. For example, the large number of genes per cell and the complete RNA extraction from the intact cells/nuclei in scRNA allows for fine-grained classification of cell types. However, the spatial neighbourhood information is lost in scRNAseq, leading to omission of the geographic cell communities within the tissue microenvironment. Further, the scRNAseq-based inference of cell-cell interaction is less accurate due to the lack of neighbourhood constraint (Pham et al., 2023). Although not at single-cell resolution, Visium data enable the inference of all ligand and

receptor genes taking into account spatial neighbourhood context. However, as single cell resolution information is not in Visium data, the inference of cell-type specific interactions is limited. This limitation may be partially addressed via cell-type deconvolution, or completely addressed by single-cell panel based platforms, which lack coverage. Regarding proteomics, either scRNAseq or spatial transcriptomics are limited in mapping traditional immune subtypes such as distinguishing CD8 T cells and CD4 T cells, which can be measured by methods such as CODEX and Vectra Polaris. Both scRNAseq, spatial transcriptomics, and CODEX, on the other hand, do not measure metabolites, an important modality that needs methods such as spatial metabolomics to profile.

Harnessing the complementary information discussed above, we devised spatial integrative analysis for 1) cell type annotation, 2) differential expression analysis, 3) ligand-receptor based interactions, 4) community/neighbourhood composition and interaction analysis, 5) multimodal validation of cell-cell interactions, and 6) mapping of genetic association signals to spatially defined regions. We discuss below how these analysis was implemented for the skin cancer context, with the intention that the approach can be related to other biological systems.

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

We found that EMT was the main pathway that was different between melanoma and cSCC-BCC, with the stronger interaction between fibroblasts and melanocytes in melanoma samples and more interaction of fibroblasts with T cells in BCC-cSCC samples. Further, the roles of activated cancer-associated fibroblasts (CAF) and fibroblasts that have undergone epithelial-mesenchymal transition in the cancer-stroma microenvironment are also crucial in BCC and cSCC progression (Sasaki et al., 2018). In addition, for colocalization-based interactions using CODEX and Xenium data, we found the complex composition of cancer cells with fibroblasts and T cells within the melanoma community. Interactions with fibroblasts have been reported as essential for melanoma progression. Given that fibroblasts are essential in skin cancer initiation and progression (Wang et al., 2012; Werner et al., 2007; Van Hove et al., 2022), the differences in fibroblast interactions between the three cancer types may help explain the differences in metastatic potential between skin cancers. Increasing evidence suggests that interactions between mutated melanocytes and fibroblasts lead to melanoma initiation and progression (Flach et al., 2011; Kim et al., 2013; Ayuso et al., 2021). Fibroblasts are most prominent in the dermis, and there is growing evidence that they, together with KCs, are key regulators of skin cancer initiation and progression (Flach et al., 2011; Kim et al., 2013; Ayuso et al., 2021). Indeed, in melanoma, it is clear that cellular interaction with the stroma is one of the key factors driving cancer initiation and progression (Kim et al., 2013; Wang et al., 2016). During embryonic development, melanocytes migrate from the neural crest to the basal layer of the epidermis (Rawles et al., 1974), where their morphology, growth and development are largely regulated by surrounding KCs (Hirobe et al., 2005). In fully developed skin under normal conditions, melanocyte-KC interactions persist, resulting in pigmentation that protects against UV radiation. However, under pathological conditions, dysplastic melanocytes shed interactions with regulatory KCs, resulting in uncontrolled proliferation and gain of stem cell self-renewal capacity (Haass et al., 2005). Therefore, a deeper understanding of

interactions between cell types in both melanoma and NMSC will be necessary to develop a more comprehensive paradigm of skin cancer treatment.

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

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

Our atlas also enables comparison of inter- and intra-tumour heterogeneity across patients. In BCC-cSCC, comparing 11 healthy and cancer samples with scRNASeq, we found an enrichment of the immune response with CD4 and CD8 T cells, M1 and M2 macrophages, NK and classical cDC cells consistently higher in cancer samples across all five patients. At the gene level, we observed a global upregulation in cancer cells of genes related to progression and invasion like *S100A7* and *KRT6B (*Chen *et al., 2019;* Chang *et al., 2011)* in our scRNASeq data. When accounting for interpatient heterogeneity, we found 39 genes that were upregulated across the entire cSCC dataset, including genes associated with immune processes such as antigen presentation, interferon-gamma

response, viral response and cell killing. Comparisons between cancer and healthy regions of each sample identified a consistent shift towards fibroblast-based signalling in cancer. While the critical roles of fibroblasts in regulating skin cancer initiation and progression are well documented (Flach et al., 2011; Ayuso et al., 2021; Kim et al., 2013), a comprehensive map of fibroblast interactions specific to partner cell types and spatial locations, as shown in this study, is still lacking.

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

Multimodal validation of ligand-receptor interactions provided strong evidence for the LR pairs and cell types involved in the interactions at RNA and protein level at local or nano-scale distance. Our first line of cross-validation for cellular interactions was via finding consistent interactions between platforms, where the broad discovery of interactions using scRNAseq and Visium can be visualised at a single cell resolution proving the colocalisation of ligand and receptor signals between two cells in the Xenium and CosMX. The validation at RNA level was then strengthened with targeted hybridization using RNAscope technology with signal amplification chemistry, allowing us to visualise at high sensitivity the co-expression of IL34-CSF1R. While transcriptomics data allows for screening many more ligand-receptor interactions than can be achieved through proteomics, the RNA-based approaches remain as an inference test, but not a direct proof of protein-protein interactions. Beyond the RNA modality, we extended the validation to protein level, with Opal TSA chemistry that allows for highly sensitive detection of protein expression. The Opal system enabled us to validate ligand-receptor with established antibodies such as PD1-PDL1. However, the colocalisation with the Opal Polaris approach lacks the resolution to find exact interactions. We, therefore, used PLA assay to detect interactions within 20 nm distance, specifically mapping the interactions to the cell membrane.

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

Together, by using a spatial multi-omics approach with 12 independent technologies, our study represents the first comprehensive comparison of spatial cellular signatures across the three skin

major cancer types, BCC, cSCC and melanoma. Although the number of samples is small (total 24 patients), the cross-validation by independent experimental platforms provides high-confidence results for the individual measured. Some of these results add consistent evidence to the existing literature, while new findings from the cohort studied here would require future validation to external cohorts. Nevertheless, our integrative analysis with population genetics studies of >300,000 individuals suggest consistent patterns and indicate highly generalizable results.

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

Methods

Patient material and ethics

All samples (**Table S1**) were collected with informed patient consent and approved for research use under ethics approval numbers 2018000165 and 2017000318 by the University of Queensland's Human Research Ethics Committees and 11QPAH477 by the Metro South Human Research Ethics Committee. All formalin-fixed, paraffin-embedded (FFPE) blocks were previously prepared following a standard fixation procedure in 10% formalin, processed in ethanol and xylene and embedded in paraffin wax. The four melanoma samples (patients 6747-085P, 21031-08TB, 48974-2B, 66487-1A) were collected during 2008-2018 and all blocks were stored at room temperature. Full patient IDs are abbreviated as 6747, 21031, 48974 and 66487 throughout this manuscript.

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

scRNASeq was performed on 14 samples representing both cancer and non-cancer cSCC and melanocyte lesions. Healthy and cSCC biopsies were paired, from patients B18, P30, P13, P04 and

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

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

scRNASeq data analysis

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

Cancerous KC cells in cSCC samples were identified based on two intersecting criteria. Copy number variation analysis was performed using both InferCNV (Tickle et al., 2019) and CopyKat (Gao et al., 2021) using default parameters. We make use of CopyKAT's ability to predict 'Aneuploid' cells and InferCNV's de-noising and QC filtering approach to retain only the cells that are likely to be 'Aneuploid'. Candidate KC cancer cells passed the first round of filtering if they were predicted to be aneuploid by both tools. Next, genes differentially expressed between KC cells from cancer and healthy biopsies were identified using either edgeR (Robinson et al., 2010) or Scanpy *(Wolf et al., 2018)* and used to calculate an "cSCC score" using a custom python script equivalent to Seurat's AddModuleScore function. Cells receiving a cSCC score in the ≥95th percentile of all scores for both methods passed the second round of filtering. Therefore, KC cancer cells were those found to be both abnormal in ploidy and enriched for genes associated with cancer biopsies. Melanoma cells were annotated in a similar way, with the module score calculated using DE genes between the malignant melanoma sample and the benign ones in the similar manner and with an ≥80th percentile threshold used for the second step of filtering based on the observed module score distribution and the number of cells.

Visium data analysis

Data generation, processing and integration, plus cell type deconvolution and CCI analysis were performed described in the **Supplementary Methods.** For the IL34-CSF1R analysis, raw LR scores from stLearn were used to classify spots as either IL34_CSF1R-positive (i.e. with a LR score > 0 for this pair) or -negative (i.e. with a LR score = 0). Integrated Visium Seurat objects for each cancer type in turn were used to perform differential gene expression analysis using Seurat's FindMarkers function (Wilcoxon test with parameters min.pct = 0.25, logfc.threshold = 0.25, adjusted p-value threshold ≤0.05) to compare gene expression between positive and negative spots. GO enrichment analysis was performed as described above. GO terms associated with upregulated genes in melanoma were

split into functionally-related groups (**Fig 8f-g**) by calculating pairwise semantic similarity values between GO terms using GOSemSim (Yu et al., 2010). K-means clustering (k = 3) was used to cluster the resulting semantic similarity values into three groups of related GO terms. Genes associated with each GO term in each group were plotted using Complex Heatmap (Gu et al., 2016).

CosMx data analysis

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

To quantify the spatial heterogeneity of each FOV, we first constructed cell-cell neighbourhood networks by applying Delaunay triangulation to cell spatial coordinates, resulting in one network per FOV. Next, we applied Rao's quadratic entropy to each cell in each network to measure the cell type heterogeneity. 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 using the *SpaMTP FindAllDEMs* function. The top 10 m/z values per cluster, per sample, were then combined and hierarchical clustering was implemented to group similar clusters together based on pseudo-bulked expression. Melanoma clusters were identified based on spatial location and confirmed by hierarchical grouping. Differential abundance analysis was again performed to identify all significantly abundant metabolites within the melanoma cluster compared to all other clusters of each sample. Common metabolites that were differentially abundant within the melanoma cluster, across all three samples, were then identified and spatially plotted across each tissue sample.

Integrative Cell Neighbourhood/Community analysis

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

To find robust communities, we combined together 30 communities, representing Visium, Xenium, and CosMX. Each community was represented by the proportion of cell types within the community. The matrix of combined communities was used to group similar communities into functional categories (considered as meta communities) such as tumour, stromal, immune or KC. This way, communities

across cancer types and platforms can be compared. For example, we compared the cancer community (e.g., CosMX_6) in cell-cell interactions across BCC, cSCC and Melanoma samples. We also compared the cancer heterogeneity at community level (e.g., CosMX_6 and Visium_2 for the cancer sample), where each community may be more or less abundant in one cancer type compared to the other cancer types.

Integrative Cell-Cell interaction analysis

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

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

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

Spatial Datasets with single-cell resolution

Some of the interacting LR pairs were visualized on high resolution spatial transcriptomics data. STOmics and Curio-Seeker (Now acquired by Takara Bio, USA) are both spatial technologies offering a single-cell resolution. Processed data for one melanoma sample from Curio was obtained from the company. The STOmics data was generated in-house, where one melanoma and two colorectal cancer samples (FFPE tissues) were profiled by Stereo-seq OMNI technology. The data was processed using the SAW pipeline and the counts data was used for visualizing cells expressing the LR genes.

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 S21, S22**). 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 7a,c-d**).

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 S23**).

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

skInteractive database

We built the skInteractive database in the form of a visualisation dashboard with a Shiny v1.7.2 application (Sievert et al., 2020). The database has two main sections, the Atlas, which shows cell types and clustering results, and the Gene Explorer, which allows the user to browse gene expression and/or LR interaction scores for the different datasets and modalities. The Atlas dashboard was

constructed using Javascript (NuxtJS framework). We converted all plots to geo map components in Apache Echarts v5.3.1 (Li et al., 2018) that provided interactive features to work with the plots. For the Gene Explorer Shiny app, we implemented multiple tabs themes and used Seurat v4.1.1 (Hao et al., 2021) to generate plots from the different data sets stored in SeuratObjects. The Shiny application included Visium (gene expression and LR scores), CosMx (gene expression) and scRNAseq data (gene expression). skInteractive Database will be ported to AWS cloud.

Data availability

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

Code availability

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

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

Author contributions statement

Q.N. conceived the project. Key contributors to data analysis include L.G. (scRNAseq and Visium data), P.P. (scRNAseq), F.Z. and G.N. and O.M (CosMx data), X.J. (scRNAseq and Visium data), M.T. (Polaris, RNAscope, and CosMx data), O.M. (GeoMx data), E.E.K (CosMx data), M.T.G. (CosMx 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., 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., 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. contributed to data analysis and interpretation. P.S., M.S.S., K.K. and I.F. provided clinical samples. All authors contributed to writing the manuscript and have reviewed and approved the manuscript.

Acknowledgements

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

Competing interests statement

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 NanoString Technologies and hold NanoString stock or stock options. The remaining authors have declared no competing interest.

References

Anborgh, P. H., J. C. Mutrie, A. B. Tuck and A. F. Chambers (2010). "Role of the metastasis-promoting protein osteopontin in the tumour microenvironment." J Cell Mol Med 14(8): 2037-2044.

Ayuso, J. M. et al. Microfluidic model with air-walls reveals fibroblasts and keratinocytes modulate melanoma cell phenotype, migration, and metabolism. Lab Chip 21, 1139–1149 (2021).

Baghdadi, M. et al. Chemotherapy-Induced IL34 enhances immunosuppression by tumor-associated macrophages and mediates survival of chemoresistant lung cancer cells. Cancer Res. 76, 6030–6042 (2016).

Baghdadi, M. et al. High co-expression of IL-34 and M-CSF correlates with tumor progression and poor survival in lung cancers. Sci. Rep. 8, 418 (2018).

Bevona, C., Goggins, W., Quinn, T., Fullerton, J. & Tsao, H. Cutaneous melanomas associated with nevi. Arch. Dermatol. 139, 1620–1624 (2003).

Bonilla, X. et al. Genomic analysis identifies new drivers and progression pathways in skin basal cell carcinoma. Nat. Genet. 48, 398–406 (2016).

Boukamp, P. Non-melanoma skin cancer: what drives tumor development and progression? Carcinogenesis 26, 1657–1667 (2005).

Cable, D. M., E. Murray, L. S. Zou, A. Goeva, E. Z. Macosko, F. Chen and R. A. Irizarry (2022). "Robust decomposition of cell type mixtures in spatial transcriptomics." Nature Biotechnology 40(4): 517-526.

Candi, E., Knight, R. A. & Melino, G. Cornification of the skin: A non-apoptotic cell death mechanism. eLS (2009) doi:10.1002/9780470015902.a0021583.

Chang, H.-H., Dreyfuss, J. M. & Ramoni, M. F. A transcriptional network signature characterizes lung cancer subtypes. Cancer 117, 353 (2011).

Chen, C. & Shan, H. Keratin 6A gene silencing suppresses cell invasion and metastasis of nasopharyngeal carcinoma via the β-catenin cascade. Mol. Med. Rep. 19, 3477–3484 (2019).

Czarnecki, D. (2024). "Mortality from Nonmelanoma Skin Cancer in Australia from 1971 to 2021." Cancers (Basel) 16(5).

Demeter, D. R. McIlwain, S. Kinoshita, N. Samusik, Y. Goltsev and G. P. Nolan (2020). "Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front." Cell 182(5): 1341-1359.e1319.

Eckhart, L., Lippens, S., Tschachler, E. & Declercq, W. Cell death by cornification. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1833, 3471–3480 (2013).Edwards, L., R. Gupta, and F. V. Filipp. 2016. 'Hypermutation of DPYD Deregulates Pyrimidine Metabolism and Promotes Malignant Progression', Mol Cancer Res, 14: 196-206.

Feller, L., Khammissa, R. A. G., Kramer, B., Altini, M. & Lemmer, J. Basal cell carcinoma, squamous cell carcinoma and melanoma of the head and face. Head & Face Medicine 12, (2016).

Fink, M. et al. Comparison of efficacy in patients with metastatic melanoma treated with ipilimumab and nivolumab who did or did not discontinue treatment due to immune-related adverse events: A real-world data study. Cancers 13, 5550 (2021).

Flach, E. H., Rebecca, V. W., Herlyn, M., Smalley, K. S. M. & Anderson, A. R. A. Fibroblasts contribute

to melanoma tumor growth and drug resistance. Mol. Pharm. 8, 2039–2049 (2011).

Foucher, E. D. et al. IL-34 induces the differentiation of human monocytes into immunosuppressive macrophages. Antagonistic effects of GM-CSF and IFNγ. PLoS One 8, e56045 (2013).

Gandarillas, A. & Freije, A. Cycling up the epidermis: Reconciling 100 years of debate. Exp. Dermatol. 23, 87–91 (2014).

Gao, R., Bai, S., Henderson, Y.C. et al. Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. Nat Biotechnol 39, 599–608 (2021).

García-Sancha, N. et al. Overcoming resistance to immunotherapy in advanced cutaneous squamous cell carcinoma. Cancers 13, (2021).

George, D. D., Armenio, V. A. & Katz, S. C. Combinatorial immunotherapy for melanoma. Cancer Gene Ther. 24, 141–147 (2017).

Giricz, O. et al. The RUNX1/IL-34/CSF-1R axis is an autocrinally regulated modulator of resistance to BRAF-V600E inhibition in melanoma. JCI Insight 3(14) (2018).

Goodson, A. G. & Grossman, D. Strategies for early melanoma detection: Approaches to the patient with nevi. Journal of the American Academy of Dermatology 60, 719–735 (2009).

Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize implements and enhances circular visualization in R. Bioinformatics 30, 2811–2812 (2014).

Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32, 2847–2849 (2016).

Guerrero-Juarez, C. F. *et al.* Single-cell analysis of human basal cell carcinoma reveals novel regulators of tumor growth and the tumor microenvironment. *Sci Adv* **8**, eabm7981 (2022).
Guy, G. P., Jr, Machlin, S. R., Ekwueme, D. U. & Yabroff, K. R. Prevalence and costs of skin cancer treatment in the U.S., 2002-2006 and 2007-2011. Am. J. Prev. Med. 48, 183–187 (2015).

Haass, N. K. & Herlyn, M. Normal human melanocyte homeostasis as a paradigm for understanding melanoma. J. Investig. Dermatol. Symp. Proc. 10, 153–163 (2005).

Hao, Y. et al. Integrated analysis of multimodal single-cell data. Cell 184, 3573–3587.e29 (2021).

Hayashido, Y., H. Kitano, T. Sakaue, T. Fujii, M. Suematsu, S. Sakurai and T. Okamoto (2014). "Overexpression of integrin αv facilitates proliferation and invasion of oral squamous cell carcinoma cells via MEK/ERK signaling pathway that is activated by interaction of integrin αvβ8 with type Ⅰ collagen." Int J Oncol 45(5): 1875-1882.

He, S. et al. High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. Nat. Biotechnol. 1–13 (2022).

Herrscher, H. & Robert, C. Immune checkpoint inhibitors in melanoma in the metastatic, neoadjuvant, and adjuvant setting. Curr. Opin. Oncol. 32, 106–113 (2020).

Hirobe, T. Role of keratinocyte-derived factors involved in regulating the proliferation and differentiation of mammalian epidermal melanocytes. Pigment Cell Res. 18, 2–12 (2005).

Hockey, L., Onkar M., Zherui X., Samuel T., Kiarash K., Christian N., and Quan N.. 2024. 'MMCCI: multimodal integrative analysis of single-cell and spatial cell-type communications to uncover overarching and condition-specific ligand-receptor interaction pathways', bioRxiv: 2024.02.28.582639.

Hou, R., Denisenko, E., Ong, H. T., Ramilowski, J. A. & Forrest, A. R. R. Predicting cell-to-cell communication networks using NATMI. Nat. Commun. 11, 5011 (2020).

Huang, L. *et al.* Single-Cell Profiling Reveals Sustained Immune Infiltration, Surveillance, and Tumor Heterogeneity in Infiltrative Basal Cell Carcinoma. *J Invest Dermatol* **143**, 2283–2294.e17 (2023).

Jerby-Arnon, L. *et al.* A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade. *Cell* **175**, 984–997.e24 (2018).

Ji, A. L. *et al.* Multimodal Analysis of Composition and Spatial Architecture in Human Squamous Cell Carcinoma. *Cell* **182**, 497–514.e22 (2020).

Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. Nature Communications 12, (2021).

Karras, P. *et al.* A cellular hierarchy in melanoma uncouples growth and metastasis. *Nature* **610**, 190–198 (2022).

Kim, E. et al. Senescent fibroblasts in melanoma initiation and progression: an integrated theoretical, experimental, and clinical approach. Cancer Res. 73, 6874–6885 (2013).

Landi, Maria Teresa, D. Timothy Bishop, Stuart MacGregor, Mitchell J. Machiela, et al. 2020. 'Genome-wide association meta-analyses combining multiple risk phenotypes provide insights into the genetic architecture of cutaneous melanoma susceptibility', Nature Genetics, 52: 494-504.

Leiter, U., Keim, U. & Garbe, C. Epidemiology of skin cancer: Update 2019. Adv. Exp. Med. Biol. 1268, 123–139 (2020).

Lewis, S. M. et al. Spatial omics and multiplexed imaging to explore cancer biology. Nat. Methods 18, 997–1012 (2021).

Li, D. et al. ECharts: A declarative framework for rapid construction of web-based visualization. Visual Informatics 2, 136–146 (2018).

Lichtenberger, B. M. & Kasper, M. Cellular heterogeneity and microenvironmental control of skin cancer. J. Intern. Med. 289, 614–628 (2021).

Lim, Su Yin, and Helen Rizos. 2024. 'Single-cell RNA sequencing in melanoma: what have we learned so far?', eBioMedicine, 100.

Longo, S. K., Guo, M. G., Ji, A. L. & Khavari, P. A. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. Nat. Rev. Genet. 22, 627–644 (2021).

Marks, R., Dorevitch, A. P. & Mason, G. Do all melanomas come from ‘moles’? A study of the histological association between melanocytic naevi and melanoma. Australas. J. Dermatol. 31, 77–80 (1990).

Martincorena, I. et al. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. Science 348, 880–886 (2015).

Martinelli, A. L. & Rapsomaniki, M. A. ATHENA: Analysis of Tumor Heterogeneity from Spatial Omics Measurements. Bioinformatics (2022) doi:10.1093/bioinformatics/btac303.

Najem, A., J. Wouters, M. Krayem, F. Rambow, M. Sabbah, F. Sales, A. Awada, S. Aerts, F. Journe,J. C. Marine and G. E. Ghanem (2021). "Tyrosine-Dependent Phenotype Switching Occurs Early inMany Primary Melanoma Cultures Limiting Their Translational Value." Front Oncol 11: 780654.

Owens, D. M. & Watt, F. M. Contribution of stem cells and differentiated cells to epidermal tumours. Nat. Rev. Cancer 3, 444–451 (2003).

Pampena, R. et al. A meta-analysis of nevus-associated melanoma: Prevalence and practical

implications. J. Am. Acad. Dermatol. 77, 938–945.e4 (2017).

Pham, Duy, Xiao Tan, Brad Balderson, Jun Xu, Laura F. Grice, Sohye Yoon, Emily F. Willis, Minh Tran, Pui Yeng Lam, Arti Raghubar, Priyakshi Kalita-de Croft, Sunil Lakhani, Jana Vukovic, Marc J. Ruitenberg, and Quan H. Nguyen. 2023. 'Robust mapping of spatiotemporal trajectories and cell–cell interactions in healthy and diseased tissues', Nature Communications, 14: 7739.

Pozniak, J. *et al.* A TCF4-dependent gene regulatory network confers resistance to immunotherapy in melanoma. *Cell* **187**, 166–183.e25 (2024).

Ratushny, V., Gober, M. D., Hick, R., Ridky, T. W. & Seykora, J. T. From keratinocyte to cancer: the pathogenesis and modeling of cutaneous squamous cell carcinoma. J. Clin. Invest. 122, 464–472 (2012).

Rawles, M. E. Origin of pigment cells from the neural crest in the mouse embryo. Physiol. Zool. 20, 248–266 (1947).

Rieckmann, J. C. et al. Social network architecture of human immune cells unveiled by quantitative proteomics. Nat. Immunol. 18, 583–593 (2017).

Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010 Jan 1;26(1):139-40. doi: 10.1093/bioinformatics/btp616.

Rogers, H. W., Weinstock, M. A., Feldman, S. R. & Coldiron, B. M. Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the U.S. population, 2012. JAMA Dermatol. 151, 1081–1086 (2015).

Sagebiel, R. W. Melanocytic nevi in histologic association with primary cutaneous melanoma of superficial spreading and nodular types: effect of tumor thickness. J. Invest. Dermatol. 100, 322S–325S (1993).

Santoriello, C., A. Sporrij, S. Yang, R. A. Flynn, T. Henriques, B. Dorjsuren, E. Custo Greig, W. McCall, M. E. Stanhope, M. Fazio, M. Superdock, A. Lichtig, I. Adatto, B. J. Abraham, M. Kalocsay, M. Jurynec, Y. Zhou, K. Adelman, E. Calo, and L. I. Zon. 2020. 'RNA helicase DDX21 mediates nucleotide stress responses in neural crest and melanoma cells', Nat Cell Biol, 22: 372-79.

Sasaki K. et al. Analysis of cancer-associated fibroblasts and the epithelial-mesenchymal transition in cutaneous basal cell carcinoma, squamous cell carcinoma, and malignant melanoma. Hum Pathol. 2018;79:1-8. (2018).

Schürch, C. M., S. S. Bhate, G. L. Barlow, D. J. Phillips, L. Noti, I. Zlobec, P. Chu, S. Black, J.

Sharma, P. & Allison, J. P. Immune checkpoint targeting in cancer therapy: toward combination strategies with curative potential. Cell 161, 205–214 (2015).

Schütz, S., L. Solé-Boldo, C. Lucena-Porcel, J. Hoffmann, A. Brobeil, A. S. Lonsdorf, M. Rodríguez Paredes, and F. Lyko. 2023. 'Functionally distinct cancer-associated fibroblast subpopulations establish a tumor promoting environment in squamous cell carcinoma', Nat Commun, 14: 5413.

Seviiri, Mathias, Matthew H. Law, Jue-Sheng Ong, Puya Gharahkhani, Pierre Fontanillas, Stella Aslibekyan, et al. 2022. 'A multi-phenotype analysis reveals 19 susceptibility loci for basal cell carcinoma and 15 for squamous cell carcinoma', Nature Communications, 13: 7650.

Shain, A. H., and B. C. Bastian. 2016. 'From melanocytes to melanomas', Nat Rev Cancer, 16: 345-58.

Siegle, J. M., A. Basin, A. Sastre-Perona, Y. Yonekubo, J. Brown, R. Sennett, M. Rendl, A. Tsirigos, J. A. Carucci and M. Schober (2014). "SOX2 is a cancer-specific regulator of tumour initiating potential in

cutaneous squamous cell carcinoma." Nat Commun 5: 4511.

Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2022. CA Cancer J. Clin. 72, 7–33 (2022).

Sievert, C. Interactive Web-based Data Visualization with R, Plotly, and Shiny. (Chapman & Hall/CRC, 2020).

Sini, M. C. et al. Genetic alterations in main candidate genes during melanoma progression. Oncotarget 9, 8531–8541 (2018).

Song, Liyang, Wenhao Chen, Junren Hou, Minmin Guo, and Jian Yang. 2024. 'Spatially resolved mapping of cells associated with human complex traits', medRxiv: 2024.10.31.24316538.

Stanley, E. R. & Chitu, V. CSF-1 receptor signaling in myeloid cells. Cold Spring Harb. Perspect. Biol. 6, (2014).

Thieu, K., Ruiz, M. E. & Owens, D. M. Cells of origin and tumor-initiating cells for nonmelanoma skin cancers. Cancer Lett. 338, 82–88 (2013).

Tickle T, Tirosh I, Georgescu C, Brown M, Haas B (2019). inferCNV of the Trinity CTAT Project. Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA.

Tirosh, I., B. Izar, S. M. Prakadan, M. H. Wadsworth, 2nd, D. Treacy, J. J. Trombetta, A. Rotem,C.Rodman, C. Lian, G. Murphy, M. Fallahi-Sichani, K. Dutton-Regester, J. R. Lin, O. Cohen, P. Shah, D. Lu, A. S. Genshaft, T. K. Hughes, C. G. Ziegler, S. W. Kazer, A. Gaillard, K. E. Kolb, A. C. Villani, C. M. Johannessen, A. Y. Andreev, E. M. Van Allen, M. Bertagnolli, P. K. Sorger, R. J. Sullivan, K. T. Flaherty, D. T. Frederick, J. Jané-Valbuena, C. H. Yoon, O. Rozenblatt-Rosen, A. K. Shalek, A. Regev, and L. A. Garraway. 2016. 'Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq', Science, 352: 189-96.

Tran, M. et al. A robust experimental and computational analysis framework at multiple resolutions, modalities and coverages. Front. Immunol. 13, 911873 (2022).

Tsao, H., C. Bevona, W. Goggins and T. Quinn (2003). "The transformation rate of moles (melanocytic nevi) into cutaneous melanoma: a population-based estimate." Arch Dermatol 139(3): 282-288.

Tuong, Z. K., A. Lewandowski, J. A. Bridge, J. L. G. Cruz, M. Yamada, D. Lambie, R. Lewandowski, R. J. Steptoe, G. R. Leggatt, F. Simpson, I. H. Frazer, H. P. Soyer and J. W. Wells (2019). "Cytokine/chemokine profiles in squamous cell carcinoma correlate with precancerous and cancerous disease stage." Scientific Reports 9(1): 17754.

Tuong, Z. K., S. W. Lukowski, Q. H. Nguyen, J. Chandra, C. Zhou, K. Gillinder, A. A. Bashaw, J. R. Ferdinand, B. J. Stewart, S. M. Teoh, S. J. Hanson, K. Devitt, M. R. Clatworthy, J. E. Powell and I. H. Frazer (2021). "A model of impaired Langerhans cell maturation associated with HPV induced epithelial hyperplasia." iScience 24(11): 103326.

Van Hove, L. & Hoste, E. Activation of fibroblasts in skin cancer. J. Invest. Dermatol. 142, 1026–1031 (2022).

Villani, R., V. Murigneux, J. Alexis, S. L. Sim, M. Wagels, N. Saunders, H. P. Soyer, L. Parmentier, S. Nikolaev, J. L. Fink, E. Roy and K. Khosrotehrani (2021). "Subtype-Specific Analyses Reveal Infiltrative Basal Cell Carcinomas Are Highly Interactive with their Environment." J Invest Dermatol 141(10): 2380-2390.

Wang, J. X., Fukunaga-Kalabis, M. & Herlyn, M. Crosstalk in skin: melanocytes, keratinocytes, stem cells, and melanoma. J. Cell Commun. Signal. 10, 191–196 (2016).

Wang, Y. et al. IL-34 is a tissue-restricted ligand of CSF1R required for the development of

Langerhans cells and microglia. Nat. Immunol. 13, 753–760 (2012).

Weatherhead, S. C., Haniffa, M. & Lawrence, C. M. Melanomas arising from naevi and de novo melanomas - Does origin matter? Br. J. Dermatol. 156, 72–76 (2007).

Weinberg, A. S., Ogle, C. A. & Shim, E. K. Metastatic cutaneous squamous cell carcinoma: An update. Dermatol. Surg. 33, 885–899 (2007).

Werner, R. N. et al. The natural history of actinic keratosis: a systematic review. Br. J. Dermatol. 169, 502–518 (2013).

Werner, S., Krieg, T. & Smola, H. Keratinocyte–Fibroblast Interactions in Wound Healing. Journal of Investigative Dermatology vol. 127 998–1008 Preprint at https://doi.org/10.1038/sj.jid.5700786 (2007).

Wolf, F., Angerer, P. & Theis, F. SCANPY: large-scale single-cell gene expression data analysis.Genome Biol 19, 15 (2018).

Wu, S., Han, J., Laden, F. & Qureshi, A. A. Long-term ultraviolet flux, other potential risk factors, and skin cancer risk: a cohort study. Cancer Epidemiol. Biomarkers Prev. 23, 1080–1089 (2014).

Xu, Z. et al. Integrin β1 is a critical effector in promoting metastasis and chemo-resistance of esophageal squamous cell carcinoma. Am. J. Cancer Res. 7, 531–542 (2017).

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 Analysis of Cystatin E/M Reveals Its Dual Functional Effects and Positive Regulation of Epithelial Cell in Human Tumors', Front Genet, 12: 733211.

Yan, G. *et al.* Single-cell transcriptomic analysis reveals the critical molecular pattern of UV-induced cutaneous squamous cell carcinoma. *Cell Death Dis* **13**, 23 (2021).
Yang, S. H., T. Andl, V. Grachtchouk, A. Wang, J. Liu, L. J. Syu, J. Ferris, T. S. Wang, A. B. Glick, S. E. Millar and A. A. Dlugosz (2008). "Pathological responses to oncogenic Hedgehog signaling in skin are dependent on canonical Wnt/beta3-catenin signaling." Nat Genet 40(9): 1130-1135.

Yerly, L. *et al.* Integrated multi-omics reveals cellular and molecular interactions governing the invasive niche of basal cell carcinoma. *Nat Commun* **13**, 4897 (2022).
Yost, K. E. *et al.* Clonal replacement of tumor-specific T cells following PD-1 blockade. *Nat Med* **25**, 1251–1259 (2019).

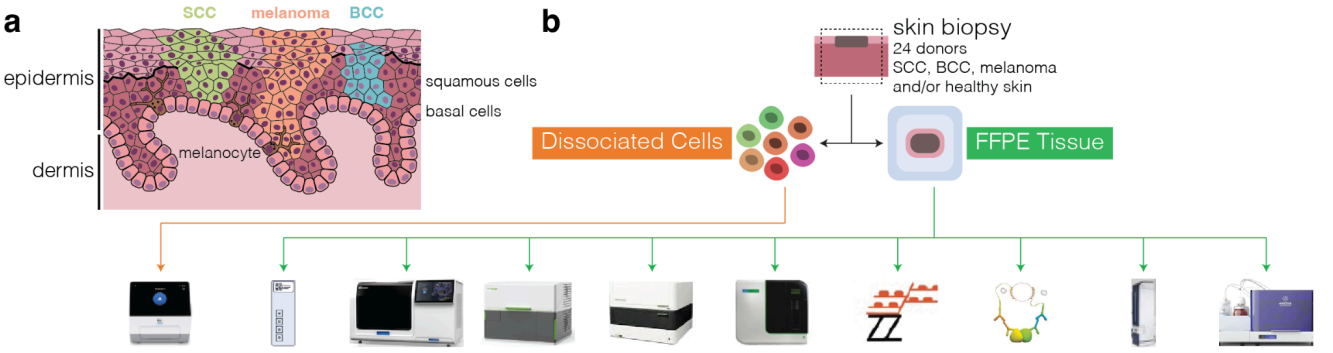
Yu, G. et al. GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. Bioinformatics 26, 976–978 (2010).

Yu, Q. and I. Stamenkovic (1999). "Localization of matrix metalloproteinase 9 to the cell surface provides a mechanism for CD44-mediated tumor invasion." Genes Dev 13(1): 35-48.

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. Yang, D. Wang, Y. Li, R. Xing, S. Teng, J. Zhao, Y. Yang, G. Zhao, K. Chen, X. Li, and J. Yang. 2022. 'A single-cell analysis reveals tumor heterogeneity and immune environment of acral melanoma', Nat Commun, 13: 7250.

Zou, D.-D. *et al.* Single-cell sequencing highlights heterogeneity and malignant progression in actinic keratosis and cutaneous squamous cell carcinoma. *Elife* **12**, (2023).

Main Figures



	scRNASeq/ snRNASeq	Visium	Xenium	CosMx	GeoMx (RNA/protein)	Polaris	RNAscope	Proximal ligation assay	Spatial glycomics	CODEX
target	RNA	RNA	RNA	RNA	RNA, protein	protein	RNA	protein	N-glycans	protein
capture method	sequencing	sequencing	probe-based imaging	probe-based imaging	probe (RNA) Ab (protein)	Ab-based imaging	probe-based imaging	Ab-based imaging	MALDI	Ab-based imaging
resolution	single cell	1-9 cells	subcellular	subcellular	ROI	subcellular	subcellular	subcellular	single cell	subcellular
capture area	N/A	whole section	whole section	FOV	FOV	whole section	whole section	whole section	whole section	whole section
# targets	~20,000	~20,000	260	960	1825 (RNA) 48 (protein)	7	5	8	96	32
image captured	✗	✓	✓	✓	✓	✓	✓	✓	✗	✓
spatial information	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓
atlas inclusion	S, B, M, HC	S, B, M, HNC	M, HNC	S, B, M	S	S	S, B	M	M	M

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.

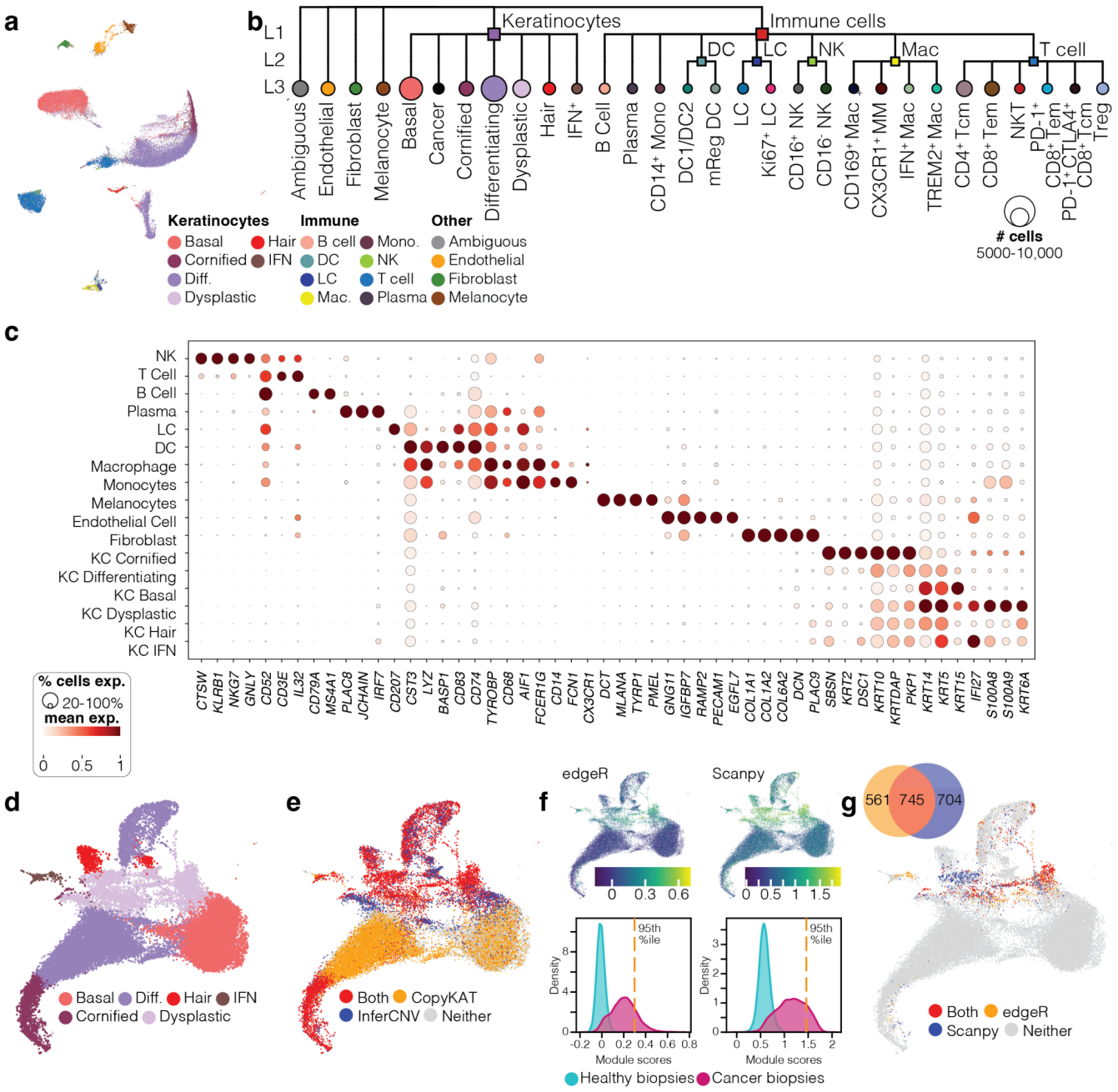


Figure 2. A single-cell atlas of cSCC, BCC and healthy skin cell types

(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 above the 95th percentile of all cell scores (f). The venn diagram indicates the number of cells passing the module score threshold by edgeR and scanpy.

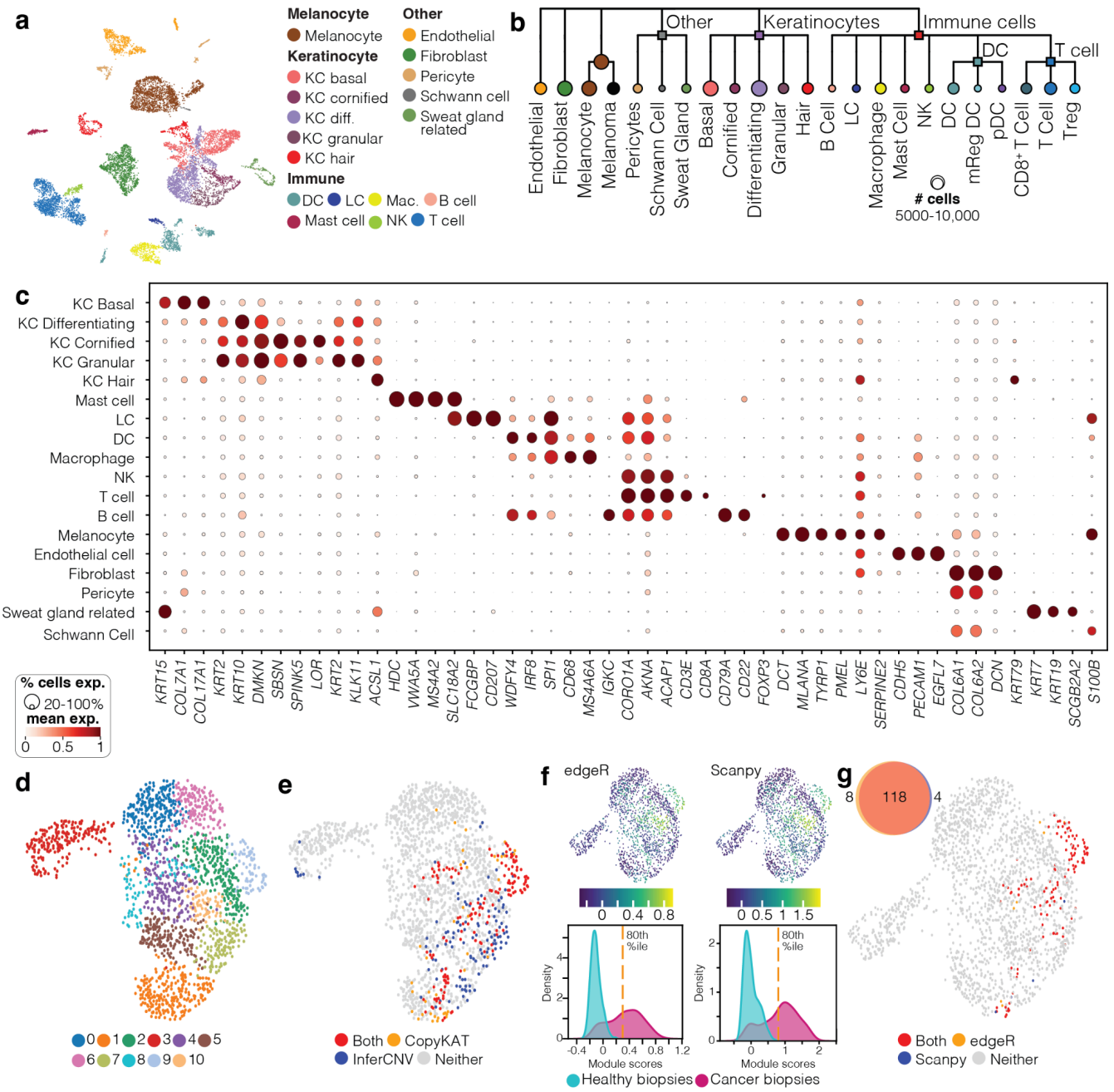


Figure 3. A single-cell atlas of melanoma cell types

(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 analysis and with a high module score by both the aforementioned methods are labelled as malignant melanocytes (red) as shown in the UMAP. The venn diagram indicates the number of cells passing the module score threshold by edgeR and scanpy.

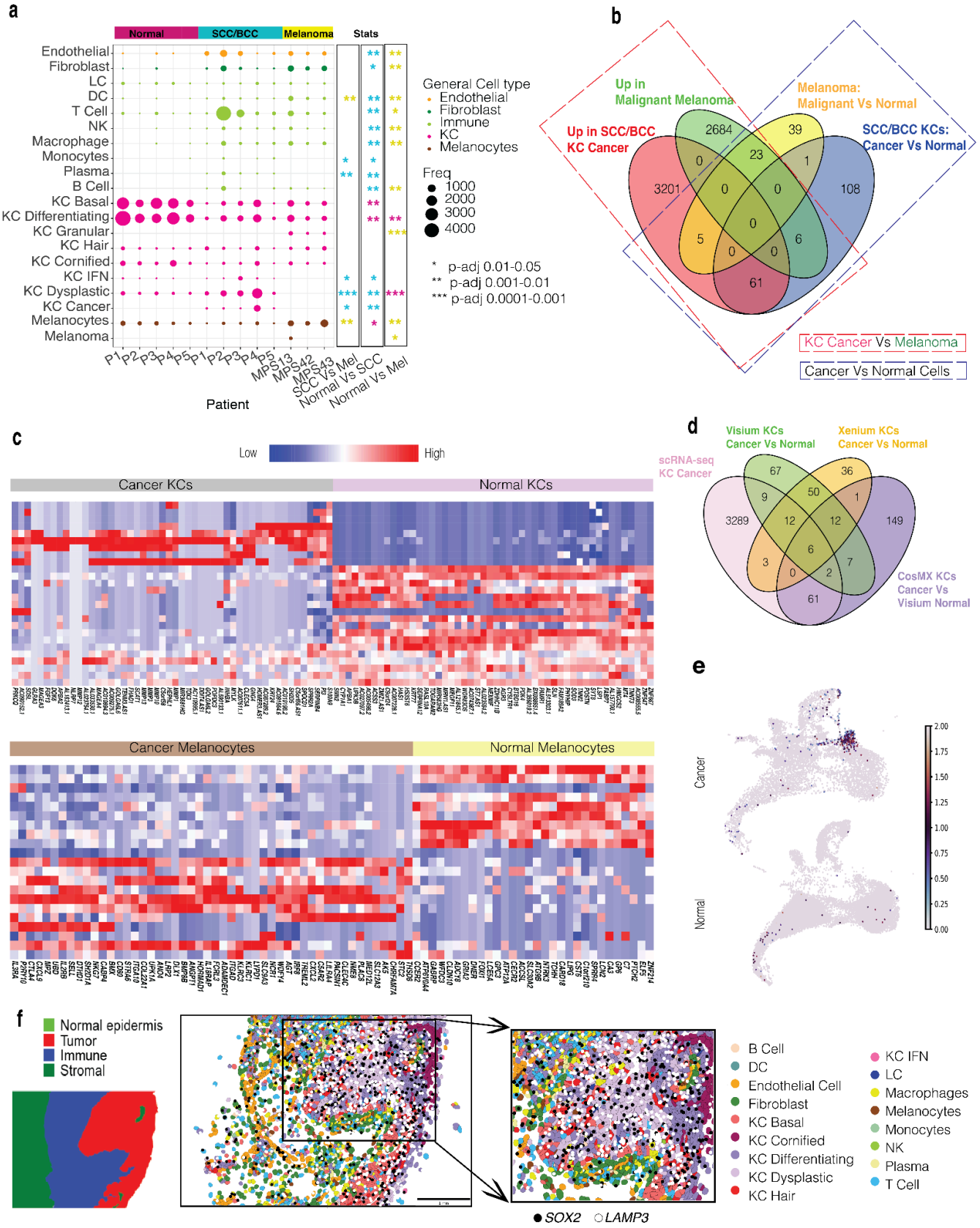


Figure 4. Biological distinctions between cSCC, BCC and melanoma in the single-cell RNASeq atlas

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

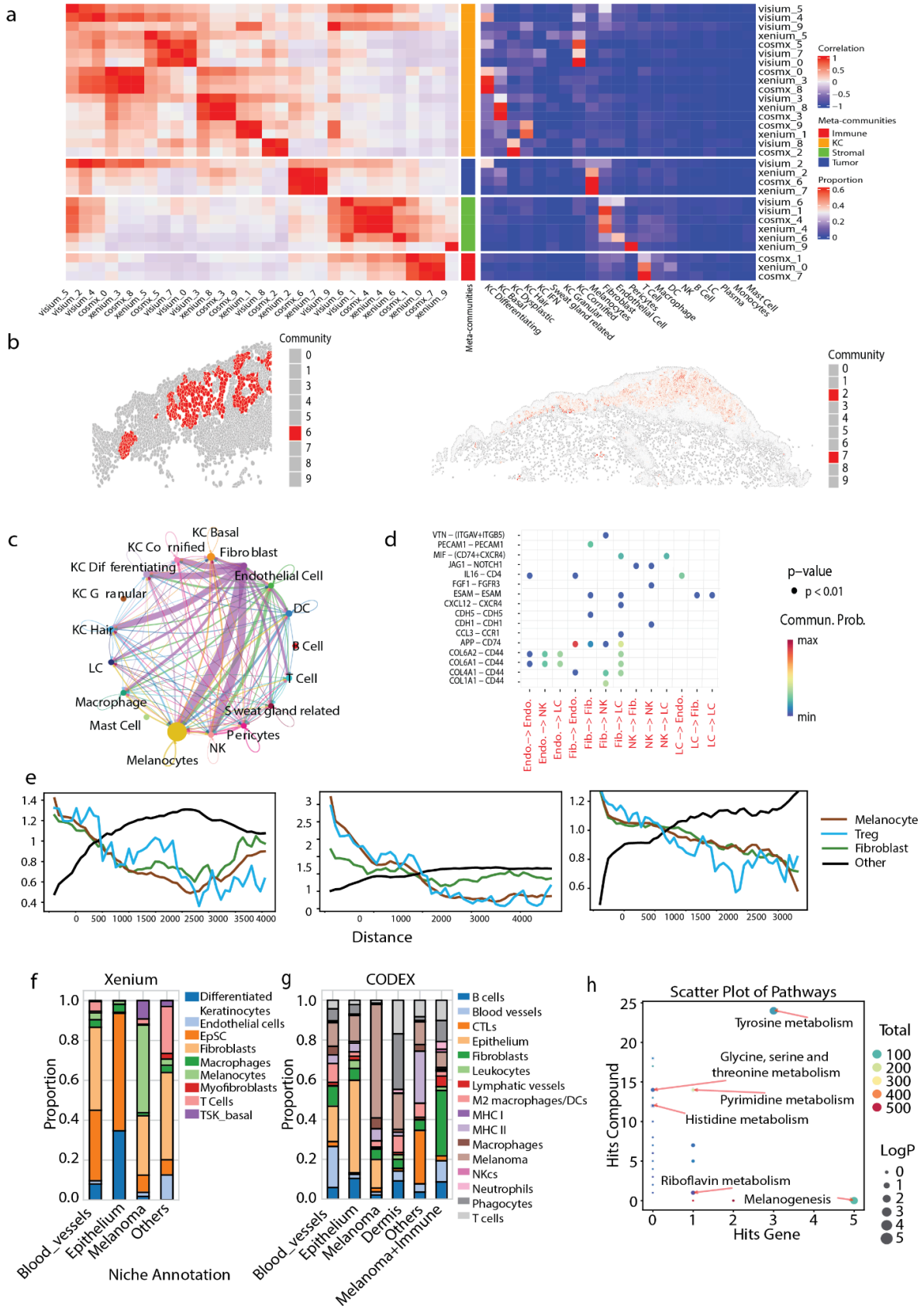


Figure 5. Integrative analysis across different individuals and technologies to build a robust spatial community atlas.

(a) Cross-modality comparison of the ten communities identified for each of Visium, CosMx and Xenium. The 4 colored bars represent super-communities (or meta-communities), which group the 10 finer communities based on their dominant cell type composition. Each row shows a community

identified from one of the three spatial platforms. The left heatmap shows similarity across communities within and between technologies, measured by pairwise Pearson correlation values between communities based on their cell type composition. This allows similar communities across technology platforms and samples to be grouped to form meta-communities. The right heatmap shows the cellular makeup of each community (i.e. proportion of each cell type per community), providing information to label the groups of communities. The central annotation shows the broad classification of communities into immune, KC, stromal or tumour-related communities, based on the cellular makeup of each.

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

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

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

(e) Cell type co-occurrence in CosMx samples between melanocytes and either other melanocytes (brown), Treg cells (blue), fibroblasts (green) or other cells (black). Each line plots the co-occurrence score (y-axis) between melanocytes and the test cell type calculated over increasing spatial distances (x-axis). The samples from left to right are melanoma 23346-105P, 30037-07BR and 6475-07FC.

(f-g) Cell type proportions of communities identified in Xenium (f) and CODEX (g) for adjacent sections from the same sample (48974-2B). The melanoma community in both datasets is enriched with melanocytes.

(h) Joint pathway analysis using upregulated genes or proteins of the melanocyte communities in Xenium and CODEX data (shown in f and g), and highly expressed glycans of the melanocyte community in MALDI data (shown in Fig S12a). The proteins, genes, metabolites are mapped to KEGG metabolic pathways. The X-axis shows the number of genes/proteins from Xenium and CODEX data found in the pathway, while the Y-axis shows glycans in the same pathway.

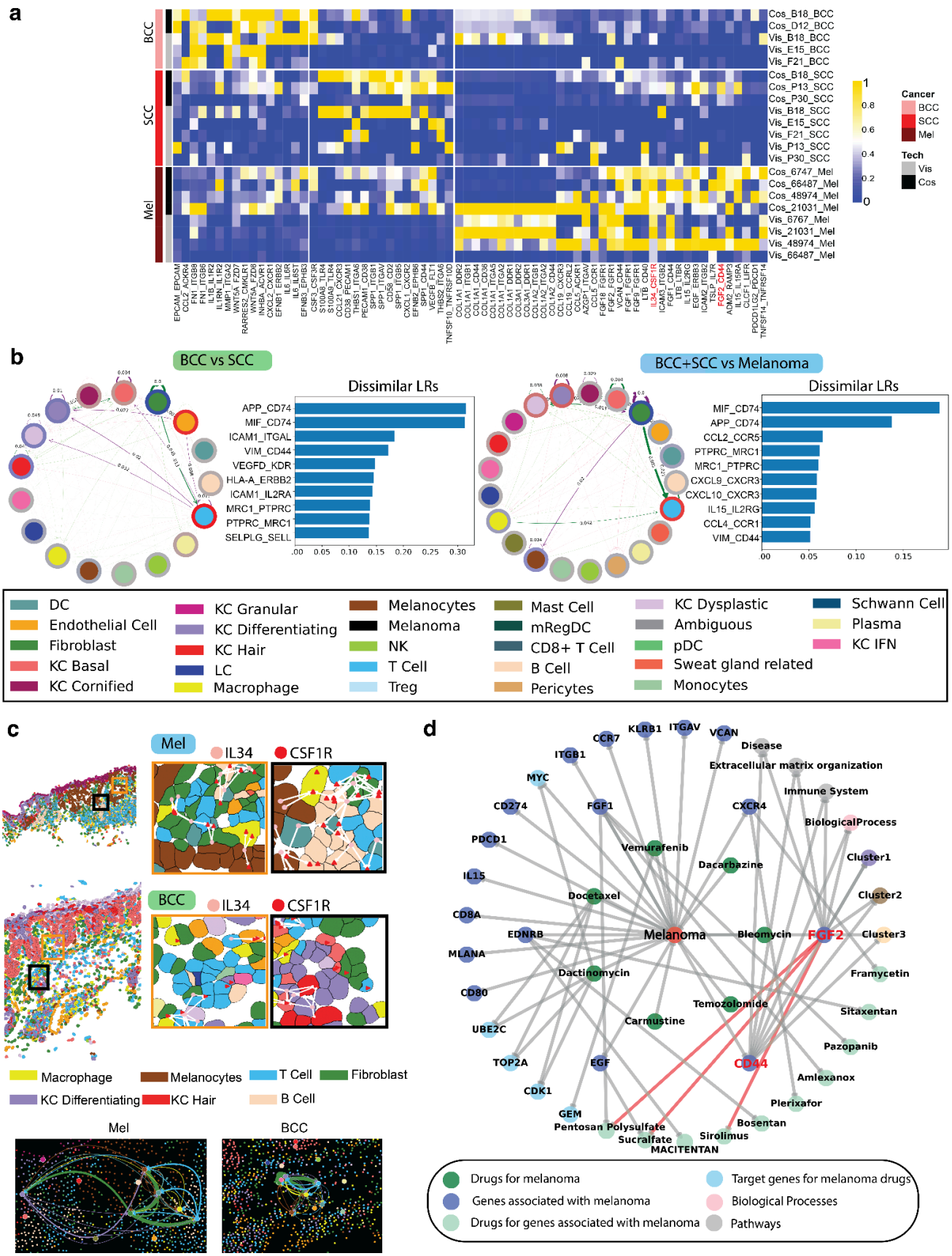


Figure 6. Differential multimodal cell-cell interaction across three cancer types using integrative analysis of Visium, CosMx, Xenium data.

(a) Heatmap of LR scores for LR pairs enriched per cancer type, with a consistent trend across samples and the two Visium and CosMx platforms. Differentially expressed LR pairs were calculated comparing each cancer type vs the others using a pseudobulked LR scores with 3 pools per sample.

Each heatmap row is a distinct CosMx or Visium sample. The two L-R pairs specific for melanoma IL34-CSF1R and FGF2-CD44 were used for experimental validations.
(b) Differential interaction analysis based on LR pairs and cell type pairs. The Venn diagram compares differential LR pair results between Melanoma and the combined BCC + cSCC datasets, calculated using edgeR with pseudobulked LR scores. The diagram highlights consistent and unique results between CosMX and Visium, where Up indicates a higher LR score in Melanoma and Down indicates a lower score in BCC + cSCC. Cell-to-cell communication between the LR pairs that are up- and downregulated in melanoma in both CosMx and Visium is shown in the two Network plots flanking the Venn diagram. In both Network plots, the purple arrows show pairs of cell types that have interactions higher in Melanoma and green arrows show interactions between cell type pairs more in the BCC + cSCC than in Melanoma. The number displayed for each arrow shows the integrated p-value across all biological replicates (the thicker arrows indicate more interactions). Interactions between the two cell types can still be significantly upregulated in melanoma even if the set of LR pairs were downregulated.
(c) Spatial mapping of cancer type-enriched LR pairs in CosMx data. One of the LR pairs that was significantly different between cancer types across technologies in Panel a, namely IL34-CSF1R (higher in melanoma) is shown. It is visualised in FOVs from melanoma sample (top) and BCC sample (bottom). For each cancer type, the cell type annotation of the FOV is shown (top left) with orange and black boxes indicating the highlighted regions (top right). Magnified boxes (top right) show the presence of the ligand (pink) and receptor (red), with white arrows showing the connections between ligands and receptors of nearby cells. An overview of interactions at tissue level is shown by large coloured arrows, representing cumulative interactions between two cell types in the tissue, with the location of the arrow root as the centroid coordinate of all cells in one cell type (bottom left).
(d) Melanoma drug target graph integrating multiple biological and pharmacological knowledge types. Nodes represent genes, drugs, and biological functions. Level 1 connections show melanoma-associated genes and drugs targeting melanoma. Level 2 links display drugs targeting the melanoma-associated genes from Level 1 and a broader gene set targeted by drugs in the network. All genes in the graph are either upregulated or have high ligand-receptor scores. Clusters 1, 2, and 3 are pathways enriched with genes shown in the graph.

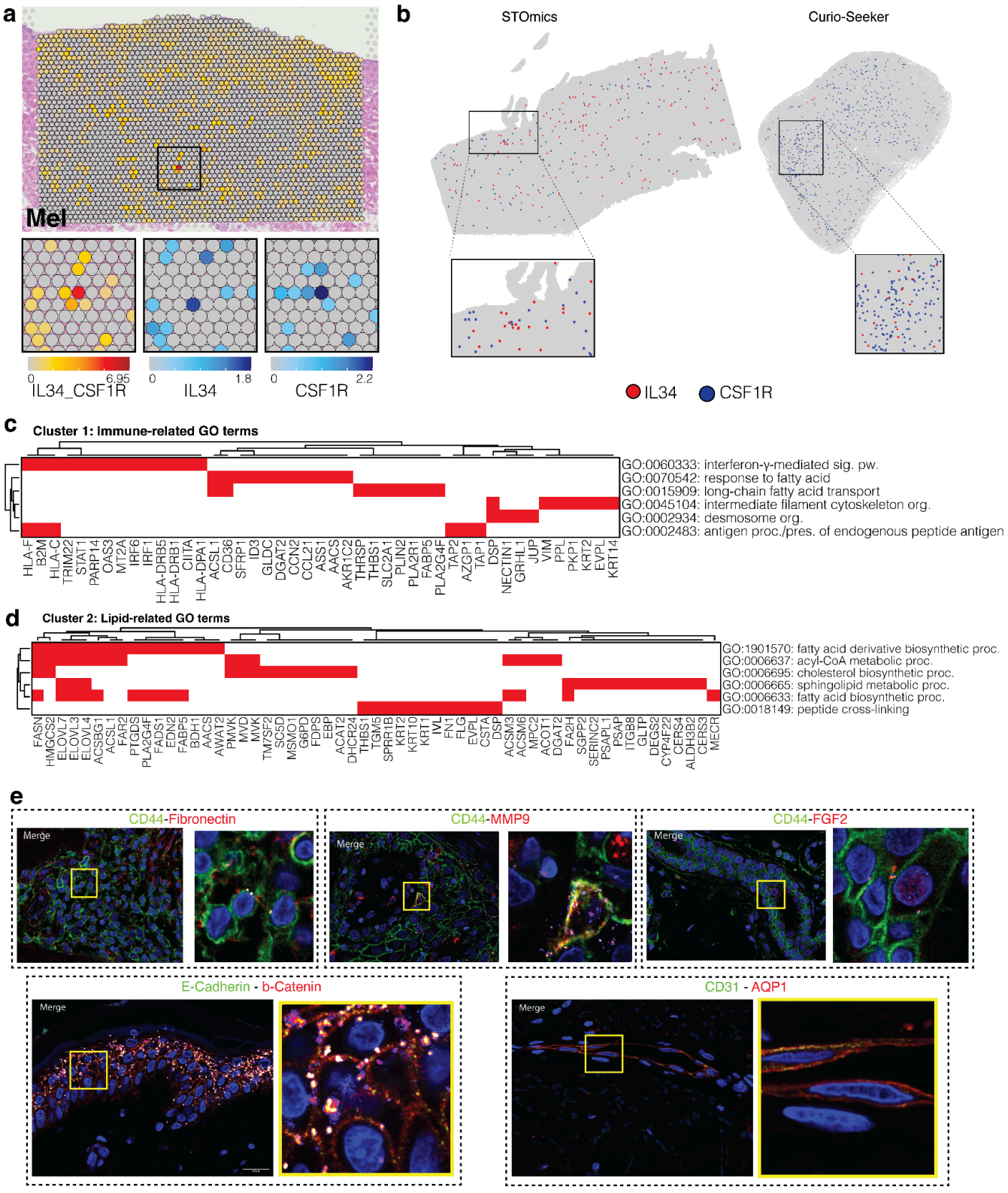


Figure 7. Multi-omics analysis of LR interaction between IL34 and CSF1R

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

(b) Melanoma high resolution spatial transcriptomics samples from STOmics and Curio-Seeker shows cells expressing IL34 and CSF1R.

(c-d) Heatmaps indicating grouped GO terms and associated genes that are enriched in IL34_CSF1R-positive spots in melanoma samples compared to IL34_CSF1R-negative spots. GO term

groups were calculated by k-means clustering (k = 3) of GO semantic similarity scores; two such groups are shown here. The full heatmap is shown in ***Fig S9b****.*
(e) Proximal ligation assay (PLA) for validating CD44 interactions in melanoma (top). A merged image of signal for the ligand and the receptor and a zoom-in window highlighting the interaction on the cell membrane. A positive PLA signal is visible if two interacting proteins are in a proximity less than 20 nm. The bottom panels show signals for positive (E-Cadherin-b-Catenin) and negative (CD31-AQP1) controls.

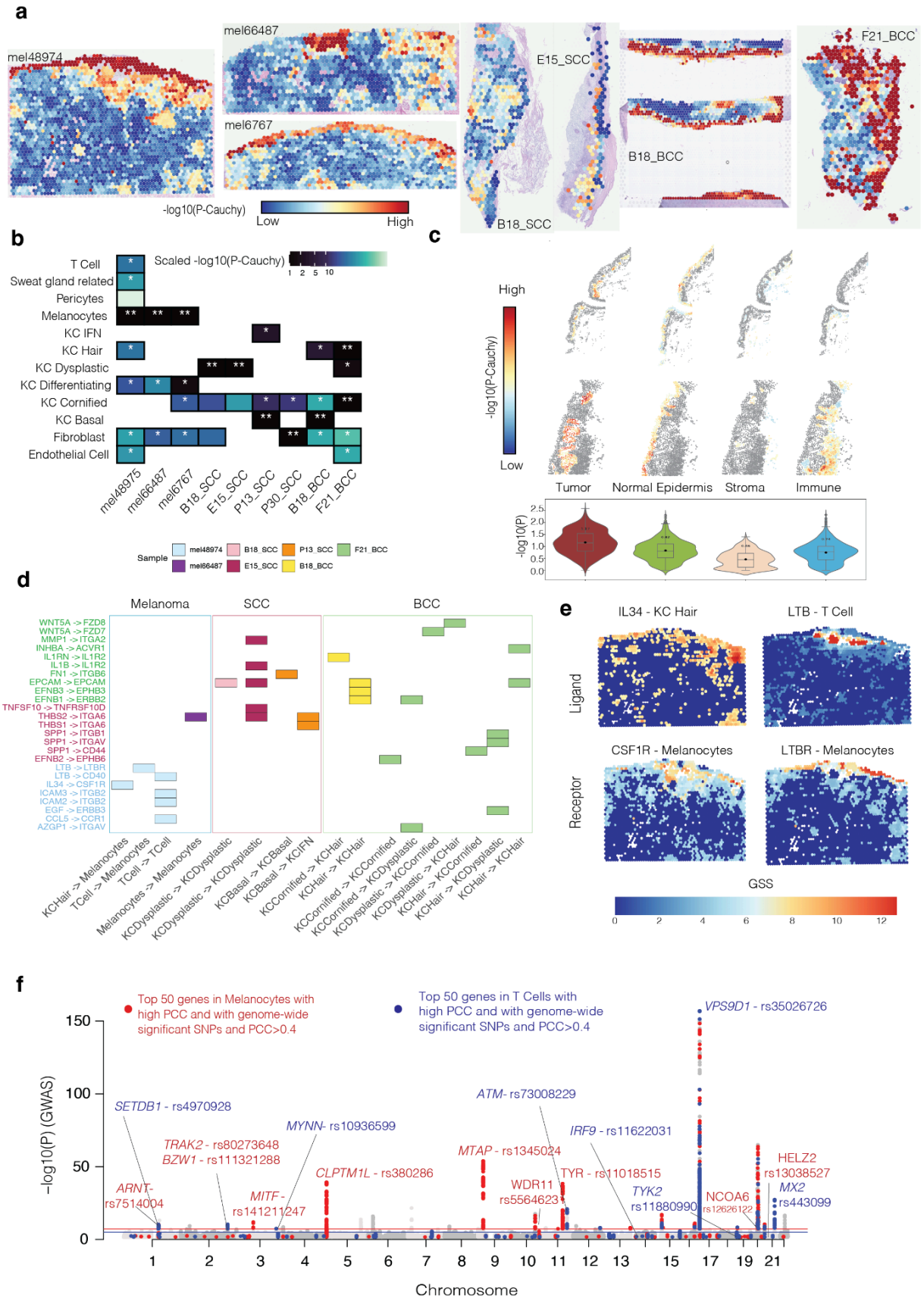


Figure 8. Mapping genetics effects from genome-wide association studies for cutaneous melanoma, cSCC, and BCC to spatial domains and cell types.

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