nature genetics

Article

https://doi.org/10.1038/s41588-026-02552-8

Single-cell spatial transcriptomic analysis of human skin anatomy

Received: 6 September 2025

Accepted: 23 February 2026

Published online: 23 March 2026

Check for updates

Paula Restrepo[1,2,3], Alexis Wilder[2,3], Aubrey Houser[1,2,3], Harkirat Singh Sandhu[2,3], Angie Ramirez[4], M. Grace Hren[2,3], Raman Gill[1,2,3], Abiha Kazmi[2,3], Larry Chen[2,3], Alexandra Nigro[2,3], Ichiro Imanishi[2,3], Deniz Demircioglu[4,5,6], Dan Hasson[2,4,5,6], Alan Soto[7], Stephanie McQuillan[7], Edgar Gonzalez-Kozlova[5,8], Rachel Brody[7], Benjamin Ungar[3], Maria Kasper[9], Catherine P. Lu[10], Philip Torina[11], Jesse M. Lewin[3], Sacha Gnjatic[5,8], Sai Ma[12,13] & Andrew L. Ji[2,3,5,6,8,14]

The skin is the largest human organ and a site of substantial disease burden, yet its cellular and molecular organization across the body is largely undefined. Here we construct an organ-wide single-cell spatial atlas of ~1.2 million cells from normal adult human skin, resolving the location of 45 cell types across 114 samples encompassing 15 anatomic sites. We uncover site-specific stereotypic cell-type composition and their organization into ten multicellular neighborhoods, most notably a perivascular neighborhood reminiscent of skin-associated lymphoid tissue. Within this neighborhood, ligand–receptor (L–R) analyses identify a central role for tumor necrosis factor in maintaining *CCL19*$^{+}$ perivascular fibroblasts, highlighting homeostatic immune–stromal crosstalk. Finally, comparing neighborhood dynamics in spatial transcriptomics of skin disease, we find pan-disease immune alterations in this perivascular neighborhood, suggesting spatial compartmentalization of pathogenic activity. Thus, multicellular neighborhoods underlie the skin's multiscale molecular to macroanatomic organization, orchestrate cell–cell interactions and anatomic site specialization and exhibit architectural disruption in disease.

Human skin has critical roles such as thermoregulation and defense against external insults. Across the body, skin shows marked functional and anatomical specialization in varied epidermal thickness, hair follicle density, moisture, pH, lipid composition and microbial communities[1–3]. These differences reflect distinct anatomic susceptibilities to dermatologic conditions, which in turn represent a substantial global burden on quality of life[4,5].

Despite these observations, how the skin's unique spatial microenvironments are maintained is poorly understood. Such knowledge could inform efforts to improve skin health, exemplified by a recent human trial that successfully modified skin thickness through intradermal injection of volar fibroblasts[6]. Prior studies on skin anatomic site specialization are limited by several factors. Although in vitro studies have identified positional memory programs in fibroblasts, their interpretation is complicated by recent single-cell RNA-sequencing (scRNA-seq)-defined subpopulations[7–12]. Bulk transcriptomic studies cannot distinguish whether site differences reflect cell abundance or cell state[13]. Single-cell and spatial transcriptomics have thus far compared only a few sites[14,15]. Critically, most studies lack accounting for donor-level variation.

Here we construct an organ-wide single-cell MERFISH[16] spatial atlas of normal human skin encompassing 1.2 million cells from 15 anatomic sites and 22 donors, including 7 donors sampled at 12 sites, enabling site-specific comparisons that account for interindividual heterogeneity. Our analyses reveal diverse cellular composition and multicellular neighborhood architecture across sites, mediated by distinct cell–cell communication networks. Among these, we identify tumor necrosis factor (TNF) as an essential regulator of immune–fibroblast crosstalk in

A full list of affiliations appears at the end of the paper. e-mail: andrew.ji@mssm.edu

a perivascular neighborhood reminiscent of skin-associated lymphoid tissue (SALT) and link neighborhood disruptions to skin disease. Finally, we provide an interactive webtool for exploring these data (https://rstudio-connect.hpc.mssm.edu/humanskin-spatialcensus/) to aid the advancement of human skin biology and disease studies.

Results

A single-cell spatial MERFISH atlas of normal human skin

To map the cellular and spatial diversity of human skin within and across anatomic sites, we profiled clinically and histologically normal skin (NS) from 22 donors (10 males, 12 females; age = 25–83 years) across 15 anatomic sites (Supplementary Table 1). Up to 12 matched sites per donor were collected during autopsies with a postmortem interval (PMI) of <24 h (n = 99 samples, n = 9 donors), along with live tissue from surgically discarded panniculectomy (n = 6 samples, n = 6 donors) and tumor-free Mohs reconstruction tissue (n = 9 samples, n = 7 donors) to include facial sites. To resolve the >40 skin cell types defined by scRNA-seq studies, MERFISH balances high transcript sensitivity that can be negatively impacted in capture-based spatial sequencing methods[17] with the multiplexing capacity necessary to distinguish subpopulations and interrogate cell–cell communication. Therefore, we designed two custom 500-gene MERFISH panels (434 overlapping genes) comprising approximately equal parts of cell-type markers and ligand–receptor (L–R) pairs (Methods; Extended Data Fig. 1a and Supplementary Table 2).

Using an optimized protocol, we generated MERFISH data from 114 samples, detecting 105,046,537 transcripts (Fig. 1a; Methods). Attesting to high sample preparation consistency, adjacent tissue sections were highly concordant in transcript detection (Spearman ρ = 0.98; Extended Data Fig. 1b), as was overlapping gene expression across panels (ρ = 0.86; Extended Data Fig. 1c). MERFISH RNA counts correlated with GTEx bulk RNA-seq (ρ = 0.77–0.81; Extended Data Fig. 1d), confirming accuracy. Unexpectedly, DV200 was not a good predictor of MERFISH data quality (Extended Data Fig. 1e), whereas autopsy PMI modestly correlated with reduced transcript detection (ρ = −0.29; Extended Data Fig. 1f), likely from RNA degradation. After segmentation and filtering, we recovered 1,201,886 cells with an average of 63 ± 78 transcripts and 27 ± 20 genes per cell (mean ± s.d.), comparable to previous MERFISH studies (Extended Data Fig. 1g,h; Methods)[18,19]. Most transcript density variation was attributable to gene panel and donor (Extended Data Fig. 1i).

Integration of the full MERFISH dataset revealed 18 major cell populations with distinct spatial localizations and accurate canonical marker expression (Fig. 1b,c and Extended Data Fig. 2a,b; Methods). Batch effects from collection source, donor or gene panel were minimal (Extended Data Fig. 2c). To validate MERFISH cell-type annotations, we integrated NS scRNA-seq from 84 donors across 14 previously published studies (n = 285,887 cells; Extended Data Fig. 2d–f and Supplementary Table 3; Methods)[10,12,14,15,20–30]. MERFISH recapitulated scRNA-seq cell types, while additionally capturing dissociation-sensitive adipocytes and sebocytes[31] (Extended Data Fig. 2g). Despite lower transcript and gene counts in autopsy samples (Extended Data Fig. 1h), all cell types remained detectable, suggesting minimal impact on clustering and annotation (Extended Data Fig. 2h). Together, these data represent an organ-wide single-cell spatial transcriptomic atlas of adult human skin.

MERFISH localizes 45 cellular subpopulations in human skin

We then resolved major epithelial, stromal and immune subpopulations (Fig. 1d). Within epithelium, MERFISH distinguished five interfollicular epidermis (IFE) keratinocyte (KC) states—basal, spinous I, spinous II, cycling and granular—expressing expected markers and localized appropriately (Extended Data Fig. 3a–c). Additional epithelial states included infundibular basal/differentiated KCs; bulge, outer root sheath/inner root sheath populations; basal/differentiated sebocytes; eccrine duct, gland and myoepithelial cells; and melanocytes (Fig. 1d). Their canonical markers matched analogous scRNA-seq clusters and spatially localized to expected tissue areas (Extended Data Fig. 3a–c)[17].

Notably, we observed two spinous KC clusters, Spn KC I and Spn KC II. Spn KC II was most prominent in the scalp and sole and displayed an elevated infundibular-like (*S100A8*, *SOX9*, *GJB2* and *GJB6*) gene expression profile similar to previously described *S100A8*+ scalp KCs[26] and *SOX9*+ palmoplantar KCs[15] (Extended Data Fig. 3b–g and Supplementary Tables 4 and 5). To directly compare these, we assessed whether scalp and sole KCs were transcriptionally distinct in our integrated scRNA-seq reference. Basal and spinous KCs from both sites clustered together and segregated clearly from infundibulum clusters, which were enriched in scalp (Extended Data Fig. 3d–g). Sole KCs showed higher *GJB2*, *GJB6* and *SOX9* expression, whereas scalp KCs expressed more *S100A8* (Extended Data Fig. 3h), but the lack of donor-matched datasets complicates distinguishing site-specific expression from donor-specific expression. Neither scRNA-seq nor MERFISH resolved discrete scalp-specific or sole-specific IFE KC subsets; thus, we hypothesize that Spn KC II represents a shared, specialized spinous population across these sites.

We then characterized 15 stromal subpopulations, including fibroblasts, pericytes, endothelial cells (ECs), adipocytes and Schwann cells (Fig. 1d,f and Extended Data Fig. 4a). We identified eight fibroblast subpopulations—papillary (Papil Fib), two reticular (Retic Fib I and Retic Fib II), two perivascular (Perivasc Fib I and Perivasc Fib II), perineural, dermal papilla (DP)-like and dermal sheath (DS) subsets (Fig. 1f and Extended Data Fig. 4a–c). MERFISH localized all fibroblast groups, including previously described *CCL19*$^+$ Perivasc Fib I fibroblasts near vasculature[19]. Notably, *ANGPTL7* marked two populations—perineural fibroblasts (*ANGPTL7*$^+$/*ITGA6*$^+$/*KLF5*$^+$) that colocalized with Schwann cells; and Retic Fib II (*ANGPTL7*$^+$/*PRG4*$^+$/*COMP*$^+$) previously observed in palmoplantar skin[22], which localized diffusely throughout the sole dermis (Extended Data Fig. 4c–e). Our perineural fibroblasts also confirmed previously identified fibroblast subpopulation coexpressing *ANGPTL7*, *ITGA6*, *KLF5* and *EGFR* (Extended Data Fig. 4e and Supplementary Tables 4 and 5)[10,15]. The two *APOE*$^+$/*CXCL12*$^+$/*C3*$^+$ perivascular populations were distinguished by *CCL19* expression in Perivasc Fib I, which were enriched near superficial blood vessels and resembled recently named 'F3' fibroblastic reticular cell-like *CCL19*$^+$/*HLA-DRA*$^+$/*CD74*$^+$ fibroblasts[32]. In contrast, Perivasc Fib II cells were more enriched around deeper vessels. DP-like cells occupied the epithelial-adjacent regions of the upper hair follicle, while DS cells lined the hair follicles (Fig. 1d,f and Extended Data Fig. 4c). ECs clustered into lymphatic (LECs, *TFPI*$^+$/*CCL21*$^+$) and two vascular EC (VEC) subsets distinguished by *ACKR1* expression resembling a high endothelial venule-like EC (HEC; Fig. 1d and Extended Data Fig. 4a)[33].

Finally, we identified 13 immune cell types, including DC subtypes (CD1C$^+$, CLEC9A$^+$ and CCR7$^+$), Langerhans cells, monocytes, macrophages, mast cells, plasma cells, cycling immune cells and T cell subsets (naive, CD4$^+$ T_H, CD4$^+$ T_{reg} and CD8$^+$ T cells; Fig. 1d,e,g), consistent with scRNA-seq profiles (Extended Data Fig. 4f and Supplementary Tables 4 and 5). While Langerhans cells localized in the IFE and hair follicle epithelia (Fig. 1g and Extended Data Fig. 3c), most other immune cells localized to perivascular regions, except for macrophages, which were additionally scattered throughout the dermis (Fig. 1g). Expression profiles of all 41 overlapping subpopulations identified between MERFISH and scRNA-seq correlated well, as scRNA-seq did not capture sebocytes, adipocytes or eccrine myoepithelial cells (Fig. 1e).

Collectively, MERFISH identified and localized 45 unique cell populations in adult skin at a comparable granularity to scRNA-seq, yielding a high-resolution single-cell spatial atlas of human skin (Supplementary Video 1).

Stereotypic cellular distribution patterns across anatomic sites

To explore functional specialization across body sites, we quantified cellular diversity (Shannon index, a proxy for specialization) and density

https://doi.org/10.1038/s41588-026-02552-8

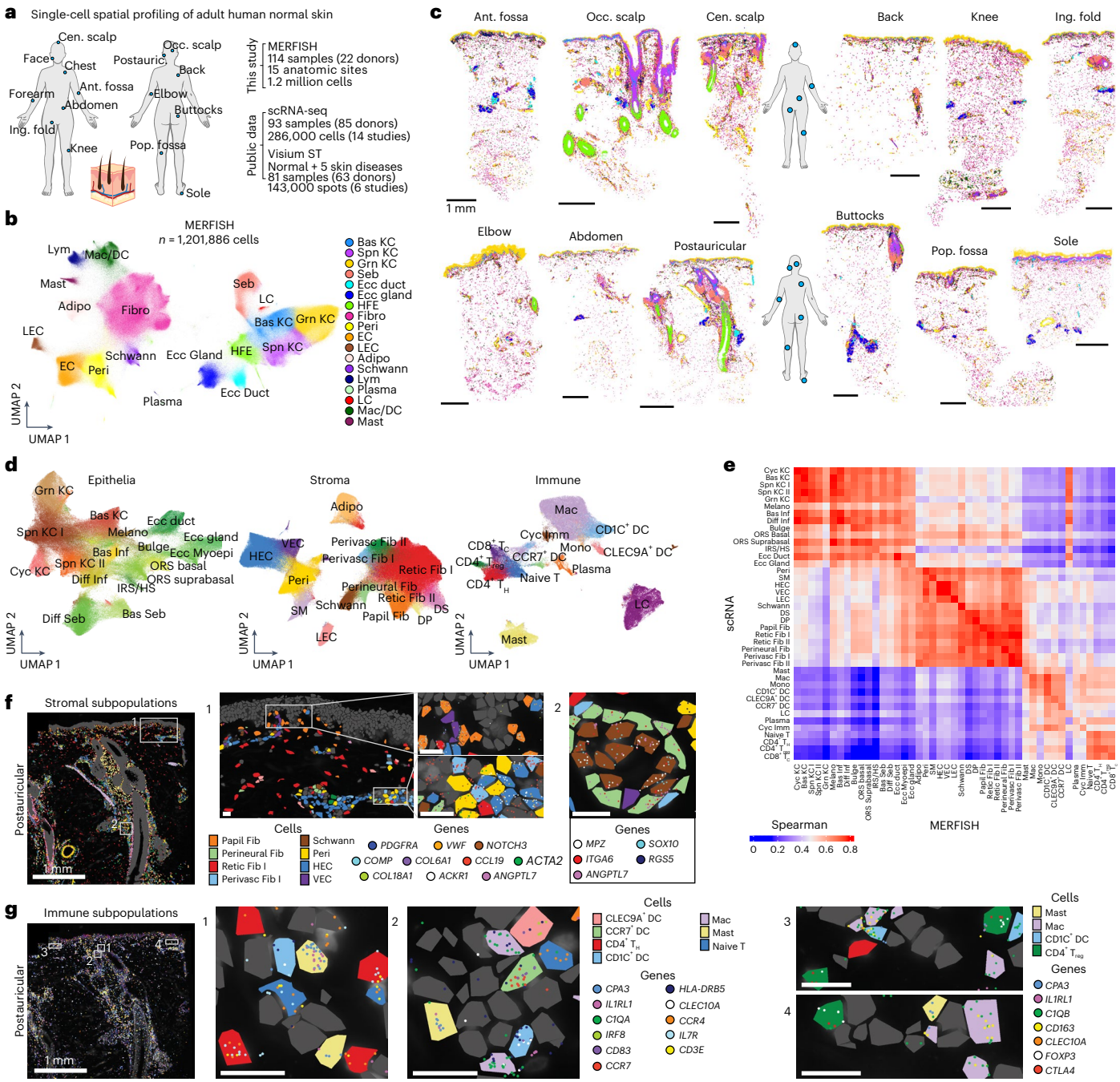

Fig. 1 | Single-cell spatial transcriptomic profiling of adult human skin using MERFISH. a, Overview of the study. **b**, UMAP of integrated MERFISH data from 114 human skin samples labeled by broad cell types. **c**, Single-cell spatial localization of cell types labeled in **b** from 12 anatomic sites from a single donor (D165) profiled with MERFISH. Scale bars, 1 mm. **d**, UMAP of epithelial, stromal and immune subpopulations identified by MERFISH and labeled by cell type. **e**, Pairwise Spearman correlations of overlapping gene expression in MERFISH and scRNA-seq cell types. **f**,**g**, Spatial localization of stromal (**f**) and immune (**g**) subpopulations in D165 postauricular sample, where each dot represents a transcript. Scale bars for all zoom-in panels, 25 µm. Adipo, adipocyte; Bas, basal; Cyc, cycling; DC, dendritic cell; Diff, differentiated; DP, dermal papilla; DS, dermal sheath; Ecc, eccrine; Fib, fibroblast; HFE, hair follicle epithelia; HS, hair shaft; Imm, immune; Inf, infindibulum; IRS, inner root sheath; LC, Langerhans cell; LEC, lymphatic endothelial cell; Lym, lymphocyte; Mac, macrophage; Melano, melanocyte; ORS, outer root sheath; Papil, papillary; Peri, pericyte; Perivasc, perivascular; Retic, reticular; Seb, sebaceous; SM, smooth muscle; Spn, spinous; T_C, cytotoxic T cell; T_H, T helper cell; T_{reg}, regulatory T cell; Ant. fossa, antecubital fossa; Pop. fossa, popliteal fossa; Cen. scalp, central scalp; Occ. scalp, occipital scalp; Ing. fold, inguinal fold; postauric, postauricular; UMAP, Uniform Manifold Approximation and Projection.

(cells per 100 µm^2). Hair-dense regions, such as the scalp, face and postauricular sites, showed the highest diversity, while extensor sites, such as the elbow and knee, featured the lowest (Fig. 2a,b). While higher diversity at face and scalp may reflect an increase in hair follicle-associated populations, the sole also exhibited high diversity despite being hairless. Furthermore, flexural sites, such as the antecubital and popliteal fossae, were more diverse than neighboring extensor regions despite similar regional coordinates along the extremities. Finally, centrally located body sites (buttocks, abdomen and back) were low in diversity, suggesting diversity increases centrifugally from central to peripheral sites. Similarly, density broadly correlated with diversity, suggesting a centrifugal density axis across the body plan (Fig. 2a,b).

https://doi.org/10.1038/s41588-026-02552-8

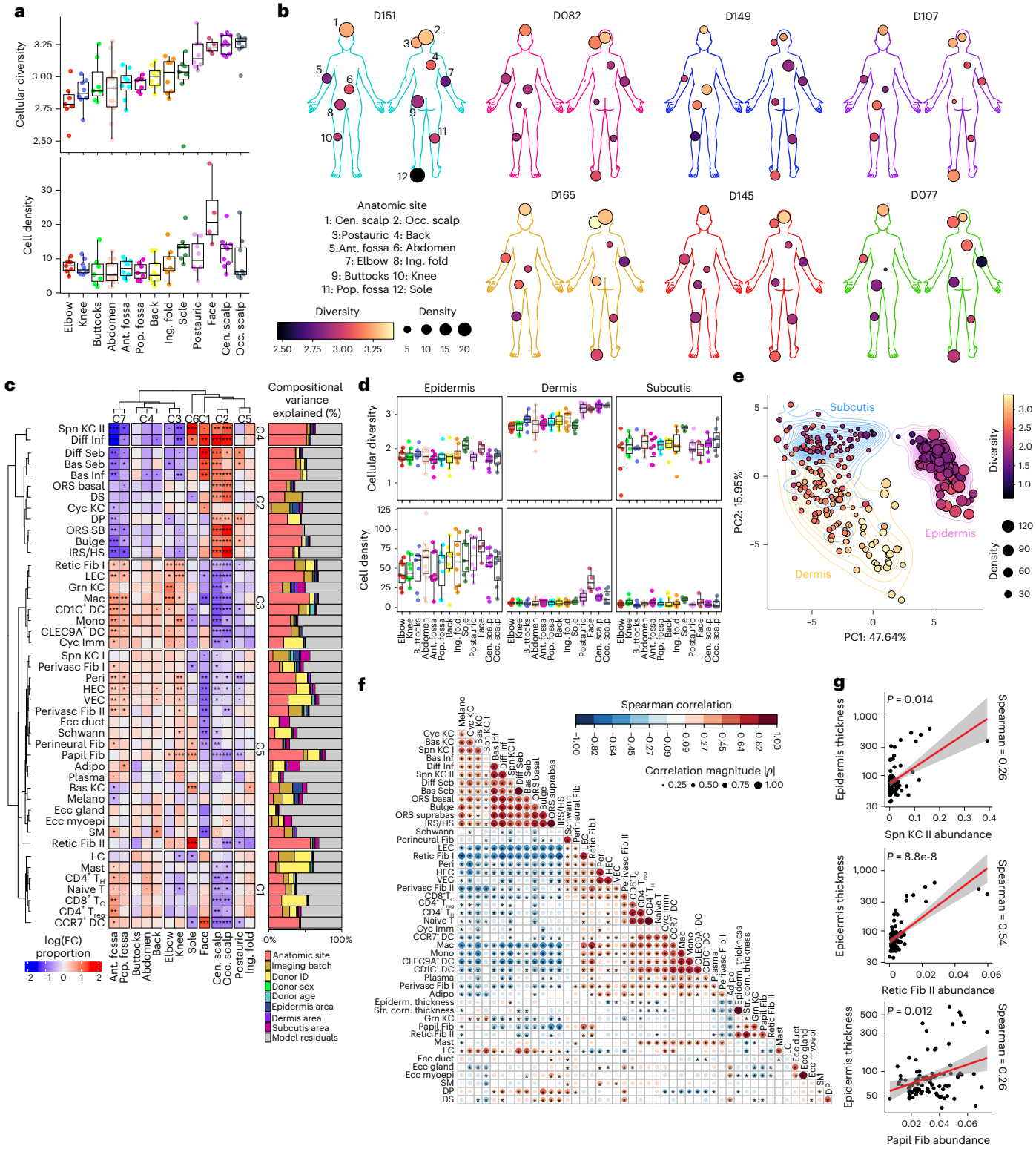

Fig. 2 | Stereotypic patterns of cell-type diversity, density and abundance across sites. a, Box plots of cellular diversity (Shannon index) and density (cells per 100 μm^2) at each anatomic site, ranked from low to high diversity. Box plots are defined as median ± Q1 (bottom bound) or Q3 (upper bound), with whiskers extending to the smallest (bottom) or largest (top) value no further than 1.5× IQR from the lower (Q1) or upper (Q3) hinge representing the IQR (Q3 − Q1). **b**, Body maps depicting diversity and density for seven donors with 12 anatomic sites profiled. **c**, Left: clustered heatmap of cellular abundances across anatomic sites. Right: bar plots of variance portioning showing contributions of each metadata to the variance in data. $P_{adj} < 0.1$; $^*P_{adj} < 0.05$; $^{**}P_{adj} < 0.01$; $^{***}P_{adj} < 0.001$; two-sided moderated *t* test; Benjamini–Hochberg correction for multiple comparisons. **d**, Box plots of cellular diversity (Shannon diversity index) and density (cells per 100 μm^2) at each anatomic site by tissue compartment, ranked from lowest to highest overall diversity, as in **a**. Box plots are defined as in **a**. **e**, PCA of tissue compartment density and diversity. Analyses from **a**, **c**, **d** and **e** are from $n = 109$ samples from $n = 19$ donors. **f**, Spearman correlations of cell-type proportion and epidermal/stratum corneum thickness. $^*P < 0.05$; two-sided ($n = 89$ samples from 7 donors with 12 sites profiled). **g**, Scatterplots and Spearman correlations of epidermal thickness with Spn KC II, Retic Fib II and Papil Fib abundances. Each point is a tissue sample ($n = 89$ samples from $n = 7$ donors with 12 sites profiled). IQR, interquartile range; adj, adjusted.

We then assessed underlying cell-type abundance differences driving diversity across sites with a mixed linear model framework (Methods)[34–36]. Remarkably, anatomic sites exhibited similar compositional patterns within flexural (antecubital and popliteal fossae), extensor (elbow and knee), trunk (back, abdomen and buttocks) or scalp (central and occipital scalp) sites, while face and sole were unique (Fig. 2c), highlighting robust cell abundance patterns across donors. Uniform Manifold Approximation and Projection visualization of cell-type abundances recapitulated these groupings across body categories and density–diversity (Extended Data Fig. 5a–c) and demonstrated donor-level consistency (Extended Data Fig. 5d). High-diversity sites such as the face and scalp featured a wider variance in cell-type abundances, while low-diversity buttocks, abdomen and back exhibited minimal enrichment or depletion of any cell type (Fig. 2c). Variance partitioning of demographic and technical covariates confirmed that the cell-type abundances largely varied by anatomic site, while quantifying other sources of variance (Fig. 2c). Collectively, these data suggest distinct but highly stereotypic cell compositions across sites.

Cell types driving anatomic site differences included, expectedly, Spn KC II (scalp and sole enrichment), Retic Fib II (sole) and hair follicle-associated and sebaceous gland-associated subpopulations (face, scalp and postauricular; Fig. 2c and Extended Data Fig. 3c). Interestingly, innate immune cells such as monocytes, macrophages and DC subsets appeared strongly enriched in the extremities, with similar trends for vasculature-associated cell types such as VECs, HECs, pericytes and Perivasc Fib I/Perivasc Fib II subpopulations, in addition to LECs and Retic Fib I (Fig. 2c). However, T cell subsets were distinctly enriched in antecubital fossa and to a lesser extent the popliteal fossa, divergent from elbow and knee patterns (Fig. 2c). In contrast, epidermal and eccrine gland populations were remarkably stable across sites. Together, these analyses reveal how distinct stable and dynamic cell types form unique coalitions across the skin anatomy.

To assess intratissue heterogeneity, we manually assigned each cell to its classical histological compartment of the epidermis, dermis and subcutis by cross-referencing matched histology (hematoxylin and eosin; Extended Data Fig. 5e). The dermis was most diverse, largely matching overall tissue diversity patterns, followed by the subcutis and epidermis (Fig. 2d). Interestingly, despite high overall diversity in the sole, its dermis was among the least diverse (Fig. 2a). Instead, sole diversity was driven by the epidermis and subcutis, suggesting that site differences can be driven by distinct compartments. In terms of compartment density, the epidermis was highest overall, with the face exhibiting highest epidermal and dermal densities (Fig. 2d). Using principal component analysis (PCA) of population abundance for each compartment, we found that samples clustered primarily by compartment, with density and diversity explaining most PC1 and PC2 variance, respectively (Fig. 2e). Thus, the skin possesses several means to achieve functional specialization, including diversification of distinct compartments.

Pairwise correlations of cell-type proportions reproduced known compartment groupings such as the IFE, pilosebaceous unit and eccrine gland (Fig. 2f). To extend these patterns to tissue properties, we quantified the average epidermal and stratum corneum thicknesses in adjacent histology sections from the seven donors for whom we profiled 12 anatomic sites ($n = 89$ tissues; Extended Data Fig. 5f). Because the sole is thickest, sole-enriched Spn KC II and Retic Fib II were expectedly correlated with both thickness measurements (Fig. 2f,g). However, papillary fibroblasts were the only other cell type that significantly correlated with both (Fig. 2f,g) and were most abundant in the sole, elbow and knee (Fig. 2c)—sites typically exposed to increased mechanical stress. Given the known signaling between Papil Fib and epidermal KCs and recent evidence that sole fibroblasts can thicken nonsole skin[6], our findings suggest that papillary fibroblasts from additional body sites may also potentially facilitate skin thickening.

Defining multicellular neighborhoods in human skin

Given stereotypic cell compositions across sites, we then examined cell-type proximity. Proximity enrichment analysis identified clusters of cell types that mirrored proportion-based correlations (Fig. 2f and Extended Data Fig. 6a). For example, Schwann cells tightly colocalized with perineural fibroblasts, and hair follicle bulge, outer root sheath basal/suprabasal and inner root sheath/HS cells formed cohesive groups, as expected. Orthogonally, spatial clustering with CellCharter[37] identified a maximally stable solution of ten multicellular neighborhoods present across all samples (Fig. 3a,b and Extended Data Fig. 6b,c; Methods). These neighborhoods corresponded to the dermal-epidermal junction (N0—DEJ), superficial and follicle-adjacent perivascular regions (N1—PERIVASC I), differentiated IFE (N2—DIFF IFE), deep perivascular regions (N3—PERIVASC II), reticular dermis (N4—STROMA), upper and lower hair follicle (N5—UPPER HF and N9—LOWER HF), eccrine and sebaceous glands (N6—ECCRINE and N7—SEB GLAND) and subcutis (N8—SUBCUTIS; Fig. 3b and Supplementary Video 2). Although $k = 10$ was the most stable clustering solution, CellCharter results can depend on the length scale of nearest neighbors and gene panel defining cell subpopulations. Alternative solutions (for example, $k = 3$ or 9) produced biologically interpretable neighborhoods that were hierarchically related to one another (Extended Data Fig. 6c,d), but we selected $k = 10$ as the level that best captured biologically meaningful granularity in our data.

Distinct cell types exhibited strong neighborhood preferences (Fig. 3c–e). As expected, epidermal, hair follicle and sebaceous/eccrine gland populations were enriched in their corresponding neighborhoods. Interestingly, most immune cells, including DC subsets, T cells and mast cells, were concentrated in the PERIVASC I neighborhood, alongside Perivasc Fib I/Perivasc Fib II and vasculature-associated cells, forming an organization reminiscent of SALT[38–40], analogous to mucosa-associated lymphoid tissue in other barrier tissues such as the oral cavity (tonsils) or small intestine (Peyer's patches). PERIVASC II, conversely, was relatively immune-poor and comprised primarily vasculature-related cells and lymphatic ECs. Neighborhoods varied in their cellular diversity, from low (SEB GLAND and STROMA) to high (PERIVASC I and DEJ; Fig. 3f). Collectively, these findings delineate recurrent multicellular neighborhoods that define the microanatomy of human skin across body sites.

Neighborhood abundances also varied across sites and reflected major body axes, with scalp/postauricular/face sites, extremities and trunk forming distinct clusters, while sole remained unique (Fig. 4a,b). The three pilosebaceous neighborhoods (UPPER HF, SEB GLAND and LOWER HF) were enriched in the scalp, while STROMA, DEJ, DIFF IFE and PERIVASC II were enriched in elbow and knee. In contrast, PERIVASC I and SUBCUTIS varied minimally across sites (Fig. 4b). Interestingly, the UPPER HF neighborhood was also enriched in sole, consistent with the transcriptional similarity of sole-enriched Spn KC II cells with differentiated infundibulum cells. Given shared expression of the hair follicle master regulator *SOX9* across these populations in our data and prior studies[22] (Extended Data Fig. 3a–c), *SOX9* or other hair follicle regulators in sole KCs may contribute to the sole's increased thickness and unique ability to withstand mechanical stress.

As neighborhoods are contiguous and can interact across boundaries, we evaluated cross-neighborhood coherence and potential neighborhood–neighborhood interactions by calculating pairwise correlations of neighborhood abundances across all samples. DEJ and DIFF IFE were strongly correlated ($\rho = 0.74$), consistent with their coordinated role in driving epidermal thickness, and STROMA correlated with PERIVASC II ($\rho = 0.26$; Fig. 4c). Meanwhile, SEB GLAND correlated with UPPER HF, LOWER HF and PERIVASC I. In contrast, ECCRINE exhibited low correlation with other neighborhoods except PERIVASC II, suggesting more independent or skin-extrinsic regulation.

Finally, our dataset enabled assessment of age-associated changes in neighborhood and cell-type abundance. To avoid anatomic

https://doi.org/10.1038/s41588-026-02552-8

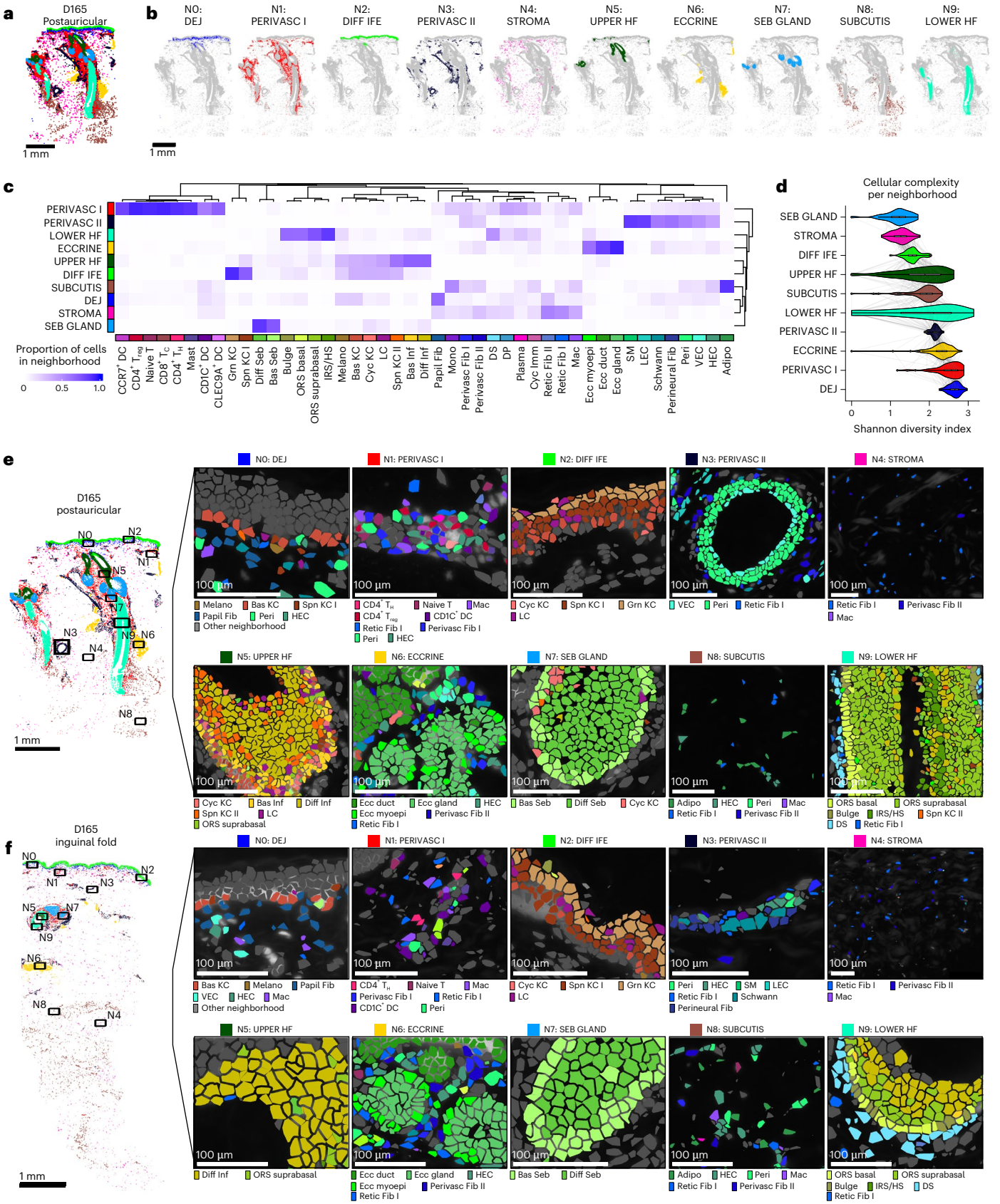

Fig. 3 | Multicellular spatial neighborhoods define cell composition and localization. a, D165 postauricular sample labeled by ten multicellular neighborhoods. **b**, Spatial highlights of cells labeled by each neighborhood from **a**. **c**, Clustered heatmap of cell-type proportions across neighborhoods. Each column sums to 1. **d**, Violin plots of Shannon diversity indices in each neighborhood ($n = 114$ samples each). Box plots inside each violin are defined as in Fig. 2a. **e**, D165 postauricular neighborhood zoom-ins for cell-type constituents of each neighborhood. Cell types are colored by same colors as in **c**. **f**, Same as **e**, but for D165 inguinal fold sample.

https://doi.org/10.1038/s41588-026-02552-8

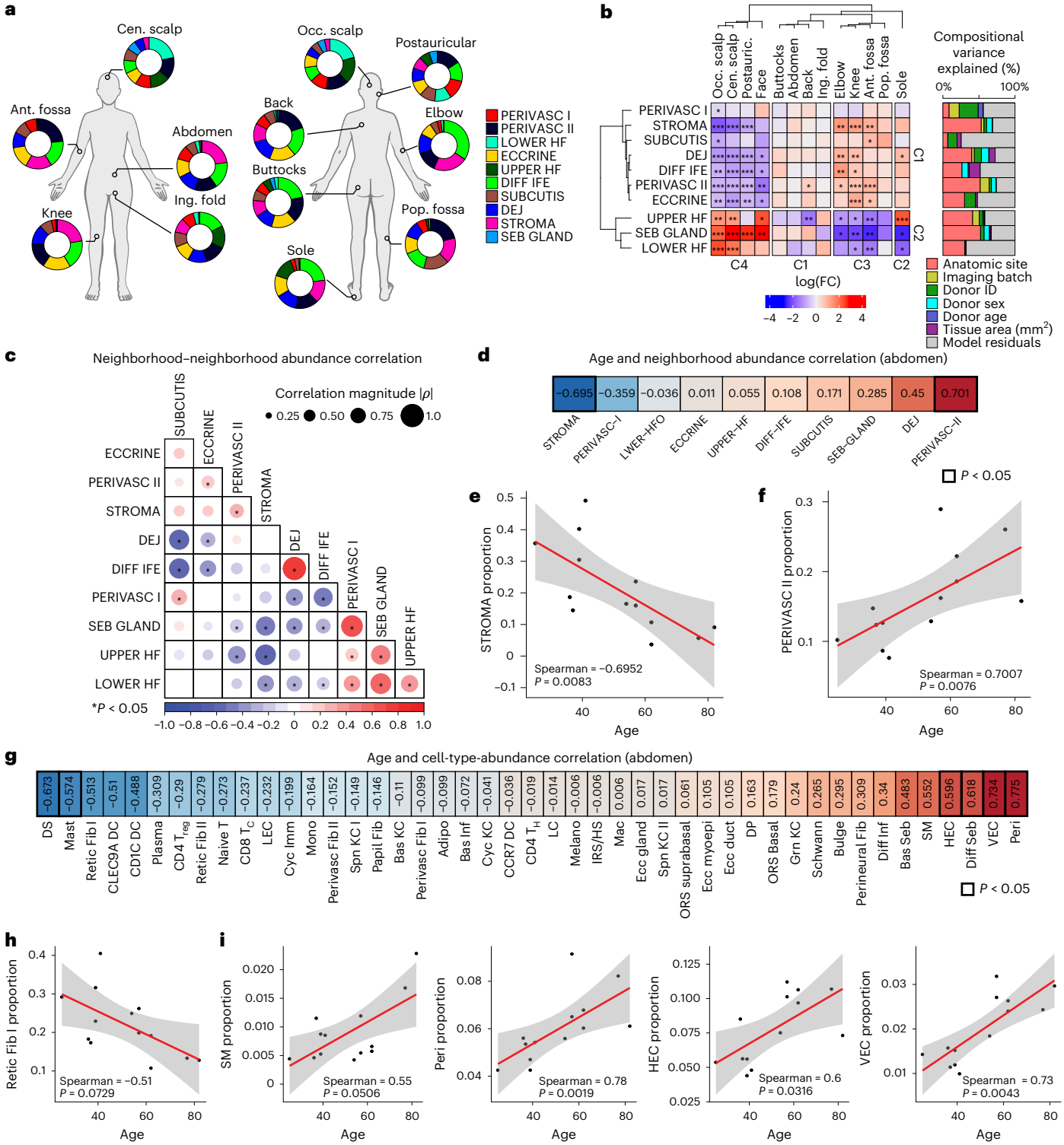

Fig. 4 | Neighborhood abundances vary across anatomic sites and during aging. a, Donut plots of neighborhood proportions at each site across $n = 7$ donors with 12 anatomic sites each. **b**, Right: bar plots of variance portioning showing contributions of each metadata to the variance in data. Left: clustered heatmap of neighborhood abundances across anatomic sites. $P_{adj} < 0.1$; $^{*}P_{adj} < 0.05$; $^{**}P_{adj} < 0.01$; $^{***}P_{adj} < 0.001$; moderated t test, two-sided; Benjamini–Hochberg correction for multiple comparisons. **c**, Clustered pairwise neighborhood abundance Spearman correlations across 13 abdomen samples ($n = 13$ donors). Dot size corresponds with correlation magnitude |Spearman ρ|. **d**, Ranked Spearman correlation coefficients (r) of neighborhood abundance with age from abdomen samples ($n = 13$ samples from $n = 13$ donors). **e**, Scatterplot of STROMA neighborhood proportion and age from $n = 13$ donors. Red line indicates linear model fit; gray area indicates 95% confidence interval. P values are associated with the two-sided Spearman correlation. **f**, Same as **e**, but for PERIVASC II proportion and age. **g**, Ranked Spearman correlation coefficients (r) and associated two-sided P value of cell-type abundance with age from abdomen samples ($n = 13$ samples from $n = 13$ donors). **h**, Scatterplot of Retic Fib I proportion and age from $n = 13$ donors. Red line indicates linear model fit; gray area indicates 95% confidence interval. **i**, Same as **h**, but for additional cell-type proportion and age—SM, Peri, HEC and VEC. Red line indicates linear model fit; gray area indicates 95% confidence interval. All P values for Spearman correlations are calculated as two-sided.

site confounding, we restricted this analysis to abdomen samples (n = 13 donors). Age was associated with decreased STROMA (Spearman = −0.70) and increased PERIVASC II (Spearman = 0.70) abundance (Fig. 4d), indicating a shift from STROMA toward PERIVASC II-dominant tissue composition. Correspondingly, Retic Fib I—the major STROMA constituent—declined with age, whereas HEC, VEC, pericytes and smooth muscle cells, which define PERIVASC II, increased (Fig. 4e). Our data suggest that Retic Fib I loss may underlie the known observation of collagen production decrease with age[41]. While prior studies have shown reduction in gene expression by fibroblast subsets with age in the inguinal fold[12], here we observe a numerical reduction in Retic Fib I in abdominal skin. In sum, we identified ten multicellular neighborhoods that define the cellular spatial organization in skin, including an immune-enriched PERIVASC I, cross-neighborhood relationships and their dynamism across anatomic sites and during aging.

Neighborhoods define distinct spatial communication networks

We then mapped L–R interactions within each MERFISH neighborhood using CellChat (Supplementary Table 6), focusing on PERIVASC I, given its high diversity and immune enrichment (Fig. 3c,d and Supplementary Table 6). Within PERIVASC I, the most interactive cell types were VECs, $CD4^+$ cytotoxic T cells, $CD8^+$ cytotoxic T cells and Perivasc Fib I/Perivasc Fib II (Fig. 5a). To extend beyond MERFISH panel genes and mitigate data sparsity, we simulated neighborhoods using our integrated scRNA-seq reference and repeated CellChat L–R inference (Extended Data Fig. 7a and Supplementary Table 7; Methods). This approach revealed MIF, midkine (MK), CXCL/CCL chemokines and TNF as top PERIVASC I pathways (Fig. 5b) and PERIVASC I and ECCRINE as having the highest interaction volume (Extended Data Fig. 7a,b). PERIVASC I also exhibited the greatest number of unique pathways, reflecting extensive heterotypic immune–stromal signaling (Extended Data Fig. 7c). Top stromal-to-immune L–R pairs in PERIVASC I included CXCL12–CXCR4, MDK–CD74, MIF–CD74 and PTN–NCL (Fig. 5c). Conversely, immune-to-stromal interactions involved PPIA-BSG, NAMPT-ITGA5 + ITGB1 and TNF–TNFRSF1A (Fig. 5d). Scoring MERFISH panel PERIVASC I L–R pairs, such as CXCL12–CXCR4, between anchoring cell types and their nearest neighbors confirmed perivascular coexpression (Fig. 5e; Methods).

We then leveraged Tangram[42] imputation to visualize spatial expression of L–R pairs predicted by scRNA-seq but absent from our MERFISH panel (Extended Data Fig. 8a; Methods). Imputed values correlated with ground truth MERFISH expression for major cell-type lineages and L–R pairs included in our MERFISH panel (Fig. 5e,f and Extended Data Fig. 8b). This enabled visualization of additional interactions, such as TNF–TNFRSF1A, which showed high PERIVASC I, DEJ and SUBCUTIS scores consistent with scRNA-seq (Fig. 5g and Extended Data Figs. 7c and 8c). TNF signaling was particularly intriguing given its pleotropic roles across homeostasis and disease and its ability to upregulate *CCL19* in cultured synovial or skin fibroblasts[28,43,44]. To functionally validate homeostatic TNF signaling in perivascular fibroblasts, we cultured human skin explants from two independent donors in the presence of recombinant TNF or vehicle control (Fig. 5h; Methods). Explants maintained normal histology with minimal apoptosis (Extended Data Fig. 9a,b). RNA FISH of Perivasc Fib I markers *PDGFRA* and *CCL19* on explant tissue sections demonstrated that TNF treatment increased $PDGFRA^+/CCL19^+$ cell numbers by approximately fourfold in superficial perivascular regions ($P < 0.001$; Fig. 5i,j), restoring or exceeding their numbers in fresh skin (Extended Data Fig. 9c,d). Consistent with this, qPCR confirmed a similar approximately fourfold increase in *CCL19* expression in TNF-treated explants ($P = 0.028$; Extended Data Fig. 9e). Notably, $PDGFRA^+$ papillary fibroblasts identified by RNA FISH did not express *CCL19* even after TNF treatment (Extended Data Fig. 9c). Therefore, TNF signaling can maintain $CCL19^+$ Perivasc Fib I identity in the PERIVASC I neighborhood.

Finally, imputation further enabled analysis of molecular variation across sites. Differential L–R scoring identified top PERIVASC I interactions at each location (Fig. 5k and Supplementary Table 8; Methods). In the antecubital fossa, elevated MHC class I (HLA-A/HLA-B/HLA-C/HLA-E) and II (HLA-DRA/HLA-DMA/HLA-DQA1) expression colocalized with CD8/CD4, alongside upregulated chemokine (CXCL2/CXCL3/CXCL8/CXCL12, CCL2) and MIF–CD74 + CD44 coexpression, potentially enhancing T cell retention and recruitment (Fig. 5k,l and Supplementary Tables 8 and 9). At the DEJ, colocalized extracellular matrix (ECM) components and cognate receptors (COL1A1/COL1A2, COL6A1/COL6A2, COMP) were consistently elevated in sole, knee and elbow in both MERFISH and imputed data, suggesting a molecular basis for increased epidermal thickness (Extended Data Fig. 8d–h and Supplementary Tables 8 and 9). Together, MERFISH directly resolved top cell–cell interactions within our panel, while imputation expanded transcriptome-wide analyses, revealing site-specific molecular programs that may underlie anatomical variation, including enhanced immune presence at flexural sites.

Skin neighborhood remodeling in disease

To determine how microanatomic neighborhood dynamics change in disease, we curated publicly available 10x Visium datasets from NS (n = 44) and five skin diseases—atopic dermatitis (AD; n = 7), psoriasis (PP; n = 4), hidradenitis suppurativa (HS; n = 10), basal cell carcinoma (BCC; n = 8) and squamous cell carcinoma (SCC; n = 8), comprising 81 samples across 63 donors (Fig. 6a,b and Supplementary Table 10; Methods)[14,17,20,22,28,45,46]. We identified 20 spot clusters spatially and transcriptionally corresponding to epithelial, immune and stromal compartments (Extended Data Fig. 10a,b). We then mapped these clusters to the ten MERFISH-defined NS neighborhoods and observed strong correspondence across epithelial, stromal, adnexal and perivascular domains (Extended Data Fig. 10a,b and Supplementary Table 11; Methods). Neighborhood signature scoring and cell2location[47] deconvolution further supported these annotations (Fig. 6c–e and Extended Data Fig. 10c,d; Methods). Together, these analyses show that Visium data recapitulate MERFISH-defined homeostatic neighborhoods, enabling interrogation of disease-associated changes.

We annotated four disease-enriched clusters (C1, C3, C4, C15; Extended Data Fig. 10a,e). C1 represented a 'KC stress' neighborhood, with *KRT6A/KRT6B/KRT6C* and *S100A8/S100A9* upregulation (Extended Data Fig. 10f and Supplementary Table 11), and was expanded across SCC, AD epidermis, BCC nontumor areas, indicating a pervasive KC stress response. C4 (DIFF IFE) and C15 (PERIVASC I) both mapped strongly to NS neighborhoods but showed disease-associated expansion. C4 increased in SCC, HS and AD, and was enriched in keratinization genes, consistent with inflammation-driven differentiation (Supplementary Table 11). C15 was expanded in SCC, BCC, HS and AD, and had elevated T cell, perivascular fibroblast and chemokine expression (Extended Data Fig. 10g). C3, enriched in SCC, BCC, HS and AD, had strong plasma cell marker and chemokine expression and corresponded to histologic tertiary lymphoid structures (TLS; Extended Data Fig. 10h). Among disease-specific clusters, C13 and C19 were restricted to BCC tumor areas (Extended Data Fig. 10f,i and Supplementary Table 11). C8, C14 and C0 favored HS (Extended Data Fig. 10e); C14 corresponded with epithelial tunnels, a hallmark of advanced HS (Extended Data Fig. 10f,j). C8 expressed weaker TLS-like signatures and localized near tunnels (Extended Data Fig. 10j,k); and C0 (DEJ and STROMA) was expanded in HS, suggesting stromal contributions to tunnel formation. Together, these annotations define both shared and disease-specific programs (Fig. 6f,g) and spatially compartmentalized pathologic responses.

Among all Visium-annotated neighborhoods, PERIVASC I (C15) was the most broadly expanded neighborhood in disease (Fig. 6h,i), resembling inducible SALT (iSALT), first defined by perivascular aggregates of DCs and T cells in contact hypersensitivity mouse models[48,49]. Biologically relevant processes within this neighborhood revealed

https://doi.org/10.1038/s41588-026-02552-8

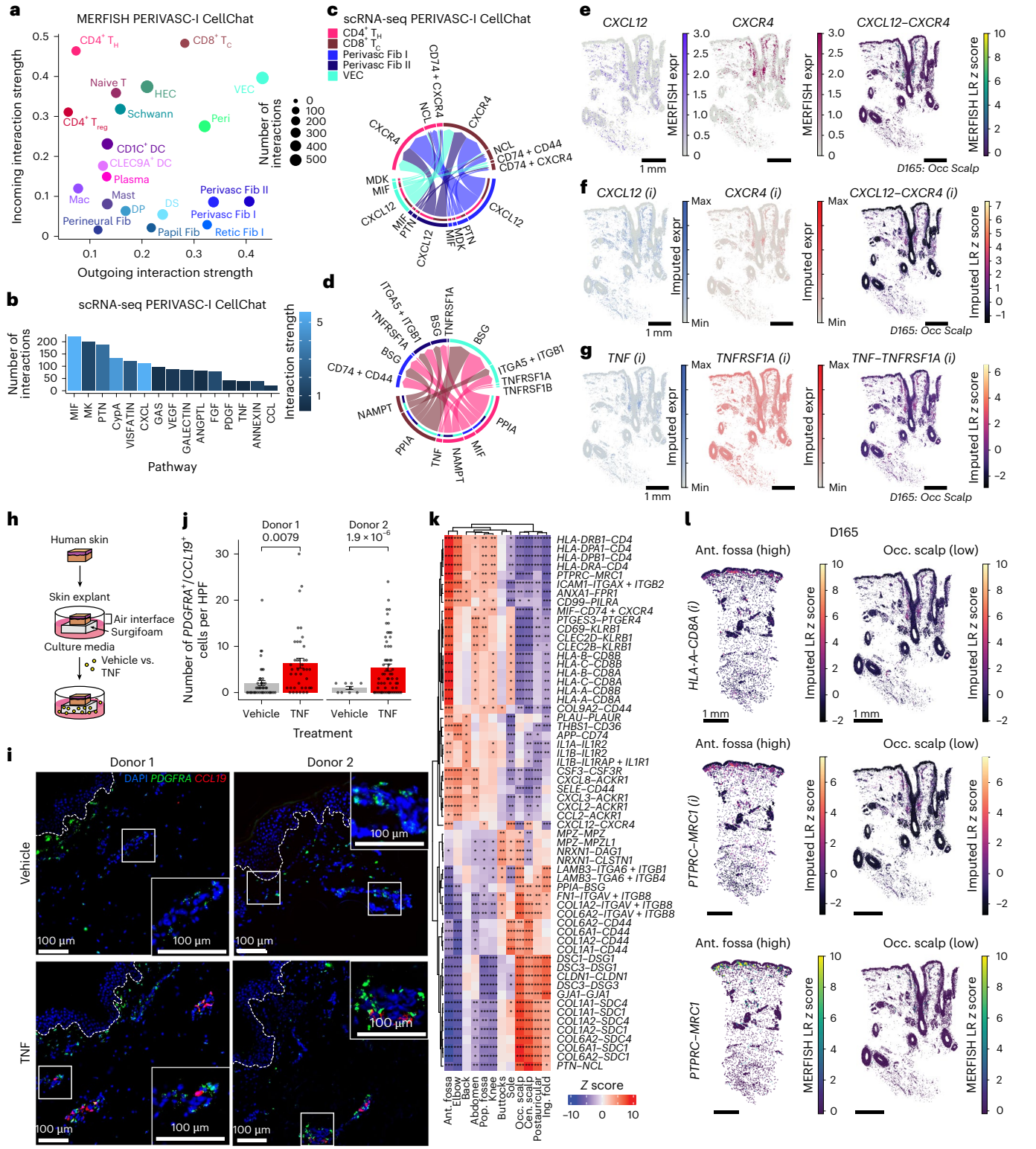

Fig. 5 | Neighborhoods define cell–cell communication within and across sites. a, Scatterplot of incoming and outgoing interaction strengths per cell type in PERIVASC I neighborhood predicted by CellChat on MERFISH data. **b**, Highest predicted active pathways in PERIVASC I neighborhood from scRNA-seq data. **c**, Circos plot of top stromal-to-immune L–R communication pairs in scRNA-seq of cell types in PERIVASC I neighborhood. **d**, Circos plot of top immune-to-stromal L–R communication pairs in scRNA-seq of cell types in PERIVASC I neighborhood. **e**, *CXCL12* and *CXCR4* expression in MERFISH data, and L–R score from D165 occipital scalp (Methods). **f**, Imputed (i) *CXCL12* and *CXCR4* expression, and imputed L–R score from D165 occipital scalp (Methods). **g**, Imputed (i) *TNF* and *TNFRSF1A* expression, and imputed L–R score from D165 occipital scalp. **h**, Schematic of human skin explant workflow. **i**, Representative RNA FISH images of *CCL19* (red), *PDGFRA* (green) and DAPI (blue) from human skin explant sections treated with vehicle control or recombinant TNF (10 ng ml^{-1}) with zoom-in insets of perivascular areas. Dashed white line indicates basement membrane of epidermis. **j**, Quantification of the number of *PDGFRA*$^+$/*CCL19*$^+$ cells per HPF. Data are mean ± s.e.m. *P* values shown were calculated by two-tailed *t* test. **k**, Heatmap of top L–R scores at each site in the PERIVASC I neighborhood. $^*P_{adj} < 0.05$; $^{**}P_{adj} < 0.01$; $^{***}P_{adj} < 0.001$; two-sided moderated *t* test; Benjamini–Hochberg correction for multiple comparisons. **l**, Imputed (i) L–R scores from D165. Nonimputed MERFISH L–R scores are also shown for *PTPRC–MRC1* (a pair included in the MERFISH panel). HPF, high-powered field.

https://doi.org/10.1038/s41588-026-02552-8

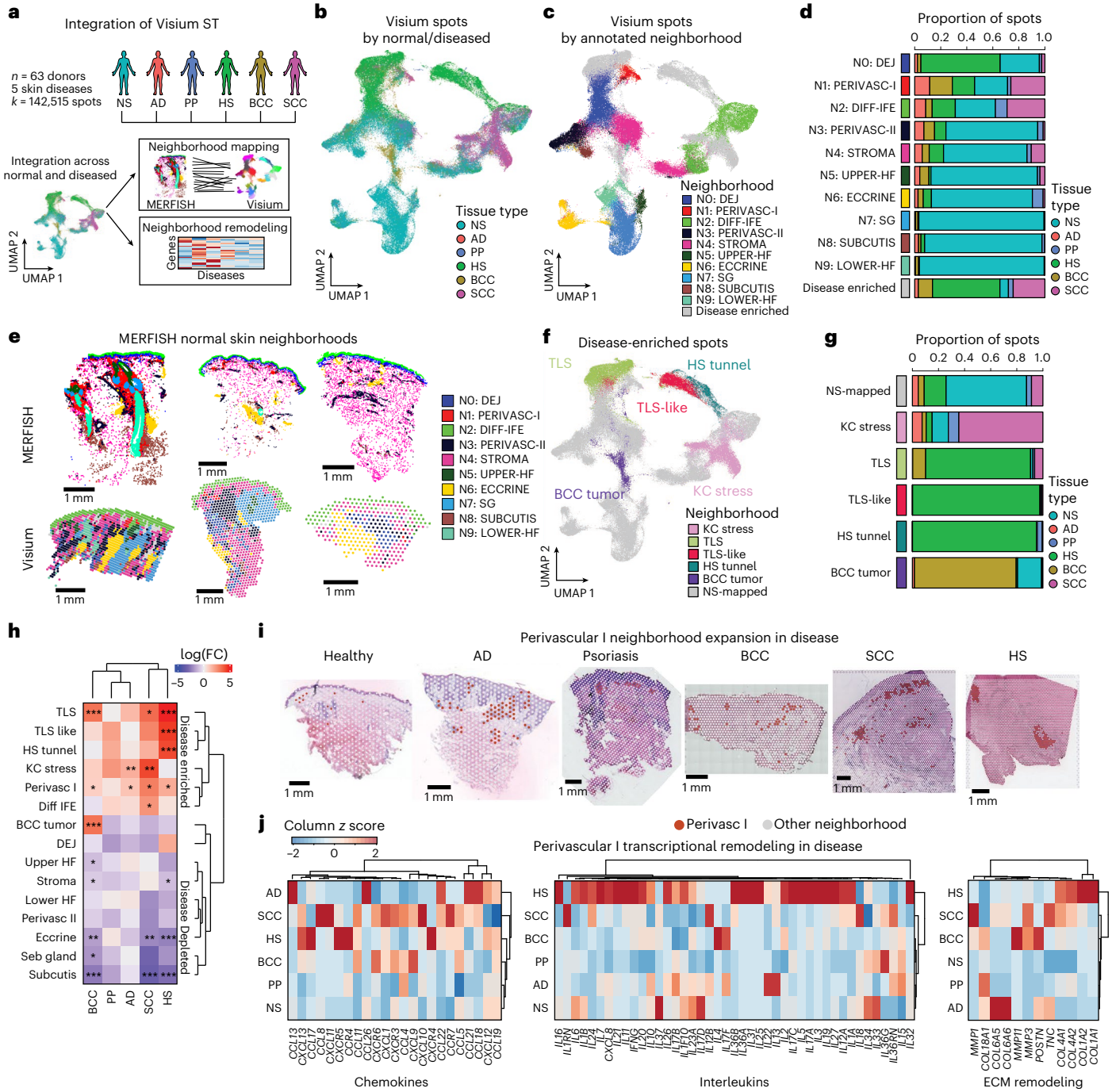

Fig. 6 | Neighborhoods define microanatomic remodeling in human skin disease. a, Workflow for Visium data integration and downstream analysis encompassing neighborhood mapping and neighborhood transcriptional remodeling. **b**, UMAP of integrated Visium spots across all samples labeled by normal and diseased tissue state ($n = 142{,}515$ spots). **c**, Same UMAP as **b**, labeled by the annotated MERFISH neighborhoods they resemble the most. **d**, Bar plots of tissue state proportions comprising each mapped MERFISH neighborhood. **e**, Spatial plots of neighborhoods in representative samples of each tissue state. **f**, Same UMAP as **b** and **c**, labeled by annotated disease-enriched neighborhoods. **g**, Bar plots of tissue state proportions comprising each disease-enriched neighborhood. **h**, Heatmap of Visium neighborhood abundances of diseased tissues relative to NS. $^{*}P_{adj} < 0.05$; $^{**}P_{adj} < 0.01$; $^{***}P_{adj} < 0.001$; two-sided moderated *t* test; Benjamini–Hochberg correction for multiple comparisons. **i**, Perivascular I neighborhood cluster highlight across representative samples of each tissue state. **j**, Heatmap of chemokine (left), interleukin (middle) and ECM remodeling (right) gene expression in Perivascular I spots across tissue states.

disease-associated increases in chemokines, cytokines and ECM factors, highlighting a key spatial domain for immune recruitment and activity (Fig. 6j). Collectively, these analyses suggest that skin disease is associated with both expansion and transcriptional remodeling of homeostatic neighborhoods, with the PERIVASC I neighborhood acting as an immunomodulatory spatial domain. Focused efforts to modulate perivascular cellular dynamics could therefore improve therapeutic targeting.

Discussion

The spatial organization of human skin underlies its form and function within local microenvironments and across the body plan. Our atlas reveals robust spatial gradients across centrifugal, anterior–posterior and craniocaudal body axes, identifying immune cell enrichment in flexural regions and papillary fibroblast enrichment in the sole, knee and elbow, correlating with increased epidermal thickness. This resource

provides multiscale insights into skin organization and serves as a foundation for further mechanistic exploration.

Fibroblast subpopulations displayed strikingly distinct spatial niches. For example, Retic Fib II were dispersed throughout the sole dermis, distinct from superficial papillary fibroblasts. Because sole fibroblasts can induce epidermal thickening in humans[6], determining whether Papil Fib, Retic Fib II or both mediate this effect could refine this promising cellular therapy. Our data align with a recent meta-analysis of human skin fibroblasts[32], which outlined six major subsets that correlate with our populations. While we localized these subsets and quantified their abundance shifts across sites, future work is required to resolve potential cell-state differences across anatomic sites, as MERFISH provides sparse transcriptomes.

MERFISH identified ten multicellular neighborhoods representing core architectural units of skin, each marked by characteristic cell–cell interactions that vary across the body and with age. As ex vivo tissue engineering efforts, including human skin organoids[50], strive to incorporate immune and adnexal structures, our data provide a blueprint for reconstructing essential tissue units, aiding efforts for regenerative therapies and disease modeling. The PERIVASC I neighborhood provides a physical correlate for the conceptual framework of SALT. Within, we show that TNF is key to sustaining *CCL19* expression in perivascular fibroblasts. Reciprocally, Perivasc Fib I may help maintain immune recruitment and residence (for example, through *CCL19*–*CCR7* and *CXCL12*–*CXCR4*), analogous to fibroblastic reticular cells in secondary lymph nodes[51]. Additional anatomic site variation of immune–stromal crosstalk molecules, such as increased MHC class I/class II and *CD4*/*CD8* in the antecubital fossa, suggests adaptive inflammatory setpoints at distinct body sites. While these may have evolved to protect more vulnerable locations, they potentially lower the threshold for chronic inflammation and disease susceptibility. Follow-up work could focus on such hypotheses.

Integration with Visium data revealed disease-associated neighborhood remodeling. *CCL19*$^{+}$ 'pro-inflammatory' fibroblasts in inflammatory skin disease[28,45,52,53] mapped primarily to PERIVASC I, which exhibited pathogenic immune expansion and architectural disruption reminiscent of iSALT. Given that targeting pro-inflammatory fibroblasts alone[54–56] reduces immune infiltration in mouse models of skin inflammation, these stromal cells represent promising therapeutic targets. Fibroblasts also contribute to TLS[57], a distinct disease-associated neighborhood. Whether TLS represents one extreme form of iSALT and whether multiple fibroblast subsets are required for its formation warrant further study. Overall, we highlight immune–stromal interplay in PERIVASC I that may underlie pathogenic inflammation, in addition to identifying other neighborhoods that are remodeled in disease.

Due to MERFISH sensitivity and gene panel limitations, rare innate lymphoid cells and natural killer cells could not be confidently identified, and neurons were not detected. We observed both segmentation 'doublets' and oversegmentation, highlighting opportunities for improving segmentation algorithms, which could be bolstered by imaging data provided in this study. We could not account for all donor-level demographic covariates that could influence analyses, such as environmental sun exposure or protective behaviors. Furthermore, MERFISH imputation from scRNA-seq data should be interpreted cautiously in the absence of ground-truth-matched scRNA-seq validation. Nevertheless, our data should aid future human skin biology and disease studies. Collectively, understanding how skin neighborhoods are constructed, preserved and vary across the human body may be key to maintaining healthy skin throughout our lifespan.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-026-02552-8.

References

1. Grice, E. A. & Segre, J. A. The skin microbiome. *Nat. Rev. Microbiol.* **9**, 244–253 (2011).
2. Byrd, A. L., Belkaid, Y. & Segre, J. A. The human skin microbiome. *Nat. Rev. Microbiol.* **16**, 143–155 (2018).
3. Merleev, A. A. et al. Biogeographic and disease-specific alterations in epidermal lipid composition and single-cell analysis of acral keratinocytes. *JCI Insight* **7**, e159762 (2022).
4. Seth, D., Cheldize, K., Brown, D. & Freeman, E. F. Global burden of skin disease: inequities and innovations. *Curr. Dermatol. Rep.* **6**, 204–210 (2017).
5. Gisondi, P. et al. Quality of life and stigmatization in people with skin diseases in Europe: a large survey from the 'burden of skin diseases' EADV project. *J. Eur. Acad. Dermatol. Venereol.* **37**, 6–14 (2023).
6. Lee, S. S. et al. The use of ectopic volar fibroblasts to modify skin identity. *Science* **385**, eadi1650 (2024).
7. Rinn, J. L., Bondre, C., Gladstone, H. B., Brown, P. O. & Chang, H. Y. Anatomic demarcation by positional variation in fibroblast gene expression programs. *PLoS Genet.* **2**, e119 (2006).
8. Chang, H. Y. et al. Diversity, topographic differentiation, and positional memory in human fibroblasts. *Proc. Natl Acad. Sci. USA* **99**, 12877–12882 (2002).
9. Ascensión, A. M., Fuertes-Álvarez, S., Ibañez-Solé, O., Izeta, A. & Araúzo-Bravo, M. J. Human dermal fibroblast subpopulations are conserved across single-cell RNA sequencing studies. *J. Invest. Dermatol.* **141**, 1735–1744 (2021).
10. Tabib, T., Morse, C., Wang, T., Chen, W. & Lafyatis, R. SFRP2/DPP4 and FMO1/LSP1 define major fibroblast populations in human skin. *J. Invest. Dermatol.* **138**, 802–810 (2018).
11. Philippeos, C. et al. Spatial and single-cell transcriptional profiling identifies functionally distinct human dermal fibroblast subpopulations. *J. Invest. Dermatol.* **138**, 811–825 (2018).
12. Solé-Boldo, L. et al. Single-cell transcriptomes of the human skin reveal age-related loss of fibroblast priming. *Commun. Biol.* **3**, 188 (2020).
13. Yan, Y. et al. Transcriptomic heterogeneity of skin across different anatomic sites. *J. Invest. Dermatol.* **143**, 398–407 (2023).
14. Ganier, C. et al. Multiscale spatial mapping of cell populations across anatomical sites in healthy human skin and basal cell carcinoma. *Proc. Natl Acad. Sci. USA* **121**, e2313326120 (2024).
15. Wiedemann, J. et al. Differential cell composition and split epidermal differentiation in human palm, sole, and hip skin. *Cell Rep.* **42**, 111994 (2023).
16. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).
17. Thrane, K. et al. Single-cell and spatial transcriptomic analysis of human skin delineates intercellular communication and pathogenic cells. *J. Invest. Dermatol.* **143**, 2177–2192 (2023).
18. Farah, E. N. et al. Spatially organized cellular communities form the developing human heart. *Nature* **627**, 854–864 (2024).
19. Cadinu, P. et al. Charting the cellular biogeography in colitis reveals fibroblast trajectories and coordinated spatial remodeling. *Cell* **187**, 2010–2028 (2024).
20. Ji, A. L. et al. Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma. *Cell* **182**, 497–514 (2020).
21. Zou, Z. et al. A single-cell transcriptomic atlas of human skin aging. *Dev. Cell* **56**, 383–397 (2021).
22. Bergenstråhle, L. et al. Super-resolved spatial transcriptomics by deep data fusion. *Nat. Biotechnol.* **40**, 476–479 (2022).
23. Eraslan, G. et al. Single-nucleus cross-tissue molecular reference maps toward understanding disease gene function. *Science* **376**, eabl4290 (2022).

Methods

Experimental model and study participant details

Enrollment and tissue collection. This study strictly complies with all legal, ethical and institutional regulations. Collection of de-identified autopsy samples for this study was performed within the guidelines set by the Mount Sinai Biorepository and Pathology CoRE through an approved protocol (IRB, 12-00145) by the Mount Sinai Institutional Review Board. Collection of surgical discards from de-identified donors for MERFISH profiling in this study was exempted by the Institutional Review Board at the Icahn School of Medicine at Mount Sinai according to the institutional criteria. For autopsies, donors with a PMI of less than 20 h and no history of skin diseases were considered, and samples were clinically and histologically normal. Additional de-identified skin was collected from surgical discards of Mohs reconstruction and panniculectomy procedures. Detailed information on full-cohort demographics is presented in Supplementary Table 1.

For rapid autopsy, 4-mm punch biopsies were collected from skin of up to 12 body sites per donor and placed in 10% buffered formalin (Thermo Fisher Scientific, 22-026-435). After 18–24 h of formalin fixation, samples were transferred to 70% ethanol for a minimum of 1 day before paraffin embedding. De-identified tumor-free tissues from Mohs reconstruction were immediately placed in 10 nM Ribonucleoside-Vanadyl Complex (New England Biolabs, S1402S) in PBS (Gibco, 10010049) on ice. Samples were cut into two sections, one for immediate fixation in 10% buffered formalin for paraffin embedding as above, and the other for embedding in OCT (Thermo Fisher Scientific, 23-730-571). Surgical discards from panniculectomy surgeries were processed similarly. Tissue was placed on ice in the operating room, transported on ice, and 4-mm punch biopsies were both immediately fixed and cryopreserved in OCT.

Experimental methods

Histology. Formalin-fixed paraffin-embedded blocks containing skin tissues were sectioned at 5 µm thickness, stained with hematoxylin and eosin, and imaged at ×20 or ×40 with a Hamamatsu Nanozoomer S210 whole slide scanner and analyzed with QuPath (v0.4.0). Tissue boundaries were drawn using the simple tissue detection feature and manually corrected to accurately capture tissue boundaries.

For average epidermal and stratum corneum thickness measurements, for each tissue sample, ten evenly spaced linear annotations were drawn per tissue sample. Epidermal thickness was measured perpendicularly from the basement membrane to the upper boundary of the stratum corneum, and stratum corneum thickness was measured from its basal to superficial boundary. Mean values across the ten measurements were used for downstream quantitative analyses.

MERFISH sample preparation. Samples were prepared following Vizgen's formalin-fixed paraffin-embedded tissue sample preparation MERSCOPE user guide (Vizgen, 91600112) with targeted modifications to counter high ambient RNAse activity[58] and ECM content in the skin. To improve tissue adherence, samples were dried at room temperature for 30 min immediately before decrosslinking. Decrosslinking incubation was extended to 22 min to intensify this process. Digestion time was also extended to 6 h. We added a concentrated clearing step during the last 6 h of the 24-h 47 °C clearing incubation. This concentrated clearing solution consisted of a 1:4 dilution of proteinase K (New England Biolabs, P8107S) in clearing premix (Vizgen, 20300114). After the concentrated clearing, the sample was returned to regular clearing solution per the MERSCOPE user guide. To reduce autofluorescence, we extended photobleaching to 5 h.

Samples were imaged following Vizgen's MERSCOPE instrument user guide with the following optimizations. To counter tissue lifting from slides, a weighted gel embedding step was added before imaging—we performed normal gel embedding using a 22-mm coverslip, placing a 15-ml conical tube cap on top of the coverslip to weigh down the gel as it polymerized for 1 h and 45 min. DAPI staining was also extended to 20 min.

Human skin explants. We adapted previously published human skin explant culture protocols[59,60]. Briefly, full skin thickness samples were collected from the abdomen of de-identified panniculectomy surgical discard donors (donor 1, 37-year-old female; donor 2, 32-year-old female; donor 3, 41-year-old female; donor 4, 60-year-old female; donor 5, 59-year-old female). After trimming of subcutaneous fat, 8-mm punch biopsies were collected fresh or cultured in triplicates on Surgifoam (McKesson General Medical, 576456) in 24-well plates in the presence of DMEM (Gibco, 11965092) supplemented with 1% FBS (Gibco, 10438026) and 1% penicillin–streptomycin (Gibco, 15140122), maintaining an air–liquid interface for epidermis. After 24 h, culture media were replaced with recombinant human TNF (PeproTech, 300-01A) reconstituted in PBS + 0.1% BSA at 10 ng ml^{-1} (a concentration previously used in vitro[61,62]) or PBS + 0.1% BSA vehicle control in fresh media for 24 h before collecting for histology, RNA FISH, immunofluorescence and qPCR.

RNA FISH. Tissues were prepared according to the RNAscope (ACDBio, 323100) protocol for fixed frozen tissue. Briefly, fresh skin or explants from panniculectomy donors 1 and 2 were fixed in 4% paraformaldehyde for 24 h, then placed in a series of sucrose gradients (10%, 20%, 30% each overnight), followed by freezing in OCT. Cyrostat sections (10 µm) were placed on glass slides (Thermo Fisher Scientific, 1255015). After assay performance validation with positive (*POLR2A*, *PPIB* and *UBC*) and negative (bacterial *dapB*) control probes, target RNA transcripts were detected using Hs-CCL19-C1 (474361) and Hs-PDGFRA-C2 (604481) probes. Fluorescence signal development was performed using RNAscope VIVID dyes, applying TSA-VIVID 650 to the C1 channel and TSA-VIVID 570 to the C2 channel, with DAPI counterstain, and imaged on a Leica upright widefield microscope. Quantification was performed by manually counting the number of *PDGFRA*$^+$/*CCL19*$^+$ cells per ×20 high-powered field of view of areas with visible vasculature in the upper dermis.

Immunofluorescence. Fresh human skin or human skin explants in culture media for 48 h were collected, embedded in OCT and cryosectioned at 10 µm. Sections were postfixed in 4% paraformaldehyde for 10 min, washed in PBS and permeabilized in 0.1% Triton X-100/PBS for 10 min at room temperature. Blocking was performed with 5.0% normal donkey serum (Jackson ImmunoResearch), 1% BSA (Sigma-Aldrich) and 0.3% Triton X in PBS for 1 h. Sections were incubated overnight at 4 °C with anti-activated Caspase-3 (R&D systems, AF835; 1:100), followed by donkey antirabbit Alexa Fluor 647 secondary antibody (Jackson ImmunoResearch; 1:400) for 1 h at room temperature. Nuclei were stained with DAPI (Thermo Fisher Scientific; 1:1,000), and slices were mounted in ProLong Diamond Antifade Mountant (Thermo Fisher Scientific). Images were acquired using a Leica upright widefield microscope with Leica LAS X software and processed in Fiji (National Institutes of Health (NIH)) for level, brightness and contrast adjustment.

qPCR. Fresh human abdominal skin and explants (donor 1, 37-year-old female; donor 3, 41-year-old female; donor 4, 60-year-old female; donor 5, 59-year-old female) tissue replicates ($n = 3$) were embedded in OCT within the same cassette and snap-frozen on dry ice immediately after collection. Approximately 40 frozen sections (16 µm each) were cryosectioned for RNA isolation using QIAzol Lysis Reagent (Qiagen, 79306) and purified with the Quick-RNA Microprep Kit (Zymo Research, R1050) according to the manufacturer's protocols. RNA concentration and purity were assessed using a NanoDrop spectrophotometer (Thermo Fisher Scientific).

One microgram of total RNA was reverse-transcribed using the iScript cDNA Synthesis Kit (Bio-Rad, 1708891). Quantitative PCR was

performed using Luna Universal qPCR Master Mix (New England Biolabs, M3003) supplemented with Rox Reference Dye (Thermo Fisher Scientific, 12223012) in technical quadruplicates ($n = 4$) on a LightCycler 480 System (Roche). ΔCt and ΔΔCt values were calculated manually, and relative gene expression was determined using $2^{-\Delta\Delta Ct}$ method. Primer sequences are provided in Supplementary Table 12.

Computational methods

MERFISH gene panel selection. The MERFISH gene panels were designed through the Vizgen portal (portal.vizgen.com) using sun-exposed NS as total fragments per kilobase million reference. For cell-type identification, five to ten canonical cell-type markers were included per cell type using the integrated scRNA-seq data as a reference. Additionally, highly expressed L–R pairs were included to assess cellular communication. Overall, 304 (54.1%) genes in the panels correspond to cell-type markers, 327 (58.2%) genes overlap with either the CellChatDB ($n = 286$ genes, 50.9%) or Omnipath ($n = 284$ genes, 50.5%) L–R databases and 434 genes (86.2%) were overlapping between the two panels, facilitating integration across panels (Extended Data Fig. 1a). Full gene panels with targeted cell types and ligand/receptor annotations are provided in Supplementary Table 2.

MERFISH cell segmentation. Segmentation was performed with the Vizgen postprocessing tool (VPT; v1.3)[63] using Cellpose with DAPI and Cellbound3 stain (proprietary antibody cocktail stain manufactured by Vizgen) or DAPI alone for benchmarking comparisons[64]. To improve segmentation, the Cellpose/Cellbound3 segmentations were also fed into Baysor (v0.7.1) using a scale parameter of −1 and a prior confidence parameter of 0.99. Baysor prior segmentation confidence parameters were optimized to maximize transcript capture across tissues, and cell boundary outputs were overlaid on DAPI and PolyT images to visually inspect segmentation. Next, cell geometry and transcript matrices from Cellpose and Cellpose/Baysor segmentation were input back into the VPT pipeline (v1.3) developed by Vizgen to partition transcripts into individual cells and create cell-by-gene counts matrices and cell-by-spatial coordinate matrices. VPT was also used to generate VZG files compatible with the MERSCOPE visualizer software using cell boundary and spatial transcript inputs, along with MERSCOPE imaging/staining data. For downstream analysis, the Cellpose/CellBound3 + Baysor segmentation was chosen and cells with volume of <100 μm^3 or containing <10 detected transcripts were removed from further analysis.

MERFISH tissue and compartment area calculation. Tissue compartments were manually annotated using canonical literature markers and the MERSCOPE Visualizer software (v2.4) for the dermis, epidermis and subcutis layers. The epidermis was defined using smFISH of the canonical *KRT5* marker for the basal layer and the presence of *KLF5* transcripts. Areas with *SOX9* expression in infundibular areas of hair follicles were used to determine a cutoff for the transition to the dermis layer. The subcutis layer was defined by the presence of *ADIPOQ* transcripts and adipose tissue, as assessed by DAPI staining. The remaining areas between the epidermis and subcutis were labeled as the dermis. To calculate area of each tissue and compartment, an alphashape of all cellular centroid coordinates was calculated on a per tissue basis and converted into units of mm^2 using the alphashape (v1.3.1) and geopandas (v1.0.1) packages in Python (v3.12.9).

MERFISH clustering and cell-type identification. Global integration of MERFISH samples was performed using scVI (v1.3.0), with the sample barcode as the batch variable and with gene panel and collection source as categorical covariates in a scanpy-based (v1.11.1) Python (v3.12.9) workflow[65,66]. Integration was restricted to the 434 features overlapping across gene panels to reduce panel-specific batch effects. Leiden clustering was performed across a range of resolutions, with a resolution of 1 selected for top-level broad cell-type identification. Cluster markers were identified using the FindAllMarkers() function from Seurat (v5.2.0). At this resolution, 20 clusters were identified and curated into 18 broad cell types based on canonical marker expression and spatial localization.

Cells were grouped into epithelial, stromal and immune compartments, then subset and reprocessed from raw counts using Harmony (v0.0.10) in place of scVI[67,68]. Epithelial populations were reclustered, yielding 30 clusters (resolution 2) that were collapsed into 20 populations; IFE cells were further subclustered, producing 17 clusters (resolution 1). Immune cells were clustered at a resolution of 2, initially generating 30 clusters that were collapsed into 21 populations. Additional subclustering was performed for lymphocytes (resolution 0.8) and macrophages/monocytes (resolution 0.4), resulting in 12 immune cell types. Stromal cells were clustered at a resolution of 2, producing 28 clusters annotated into 17 populations. Fibroblasts were then subset and reclustered at a resolution of 0.8, with annotation based on scRNA-seq markers, literature and spatial localization. Because fibroblast clusters also appeared in immune clustering—likely due to segmentation errors and transcript contamination[69]—fibroblasts from all compartments were reclustered at a resolution of 1, ultimately yielding seven defined fibroblast subpopulations. Clusters indistinguishable by marker expression at this resolution but separable at lower resolutions were merged, while poorly defined or mixed clusters were labeled 'unknown' and removed.

These subtype annotations were combined back into the full MERFISH dataset, yielding 45 distinct cell types for downstream analysis. For visualization purposes only, we reassigned compartments to cells based on these refined cell-type annotations and subset each compartment, followed by scVI integration and Uniform Manifold Approximation and Projection visualization. Further subclustering details are described in the Jupyter notebooks enclosed in the GitHub and Zenodo repositories provided with this paper[70].

Integration of public scRNA-seq data. Publicly available scRNA-seq data were downloaded from 14 studies across 93 samples and 85 donors[10,12,14,15,20,21,23–30]. Cells passing quality control filters (nCount_RNA > 100, nFeature_RNA > 200, pct.mito < 10, pct.ribo < 60, pct.hemo < 5) were retained for further analysis. One sample with <100 cells surviving these quality control filters was deemed low quality and was removed, leaving a total of 285,887 cells for downstream processing. Using Seurat (v5) integration, the data were normalized and scaled while regressing out nCount_RNA and pct.mito, then PCA was performed. Details on the studies and samples integrated are provided in Supplementary Table 3. We used Harmony to integrate the data on a per-donor basis, aiming to preserve biological variation across anatomic sites. Cells were annotated using a similar strategy as MERFISH annotations. Unsupervised clusters were identified and grouped into epithelial, stromal and immune compartments for subclustering using Seurat with resolutions of 0.6, 0.8 and 1, respectively. Clusters for each compartment were mapped to the MERFISH-based annotation structure using a consensus of unsupervised clustering markers and predictions from the Seurat reference mapping workflow. Fibroblasts from the stromal compartment were then further subclustered at a resolution of 0.8 to better resolve subpopulations and map them to the MERFISH populations.

Neighborhood identification. The integrated latent space embeddings of our MERFISH dataset were adjusted with scANVI (v1.3.0) using cell-type annotations before running CellCharter (v0.3.4)[65]. Because CellCharter requires the user to select the k number of neighborhoods, we examined all local maxima of k based on the stability metric produced by the method. Clustering solutions at $k = 3$, $k = 9$ and $k = 10$ were visualized and compared with pyclustree (v0.3.2), a Python implementation of clustree[71]. Ultimately, $k = 10$ was identified as the global maximum of cluster stability and chosen for downstream analysis. Finally,

neighborhoods were annotated and assigned names based on both their cellular composition and spatial localization within the tissues.

Differential abundance. After doublet removal, cell counts for each cell type or neighborhood were input into crumblr (v0.99.11)[69], a mixed linear model framework for compositional data using the centered log ratio[34]. Using Dream (v1.35.5), an extension of the limma (v3.60.6) framework, differentially abundant cell types were determined for each sample on a per-anatomic site basis. Donor sex, donor identification (ID) and imaging batch were included in the model with random effects. Donor age and the area in mm^2 for dermal, epidermal and subcutis compartments for each sample were included as fixed effects. For neighborhood differential abundance, total tissue area was used instead of compartment area to avoid overcorrection, as compartments can be colinear with neighborhoods.

Spatial proximity analysis. To determine which cell types are spatially proximal to each other, we used the nhood_enrichment functions from CellCharter and squidpy (v1.6.5) in Python.

MERFISH missing gene imputation. Tangram (v1.0.4) was used to impute expression of genes not included in our MERFISH panel using the scRNA-seq dataset as the reference in cluster mode[42]. To improve scalability of gene imputation across different samples and improve imputation for populations such as sebocytes and adipocytes that are not well-captured in the scRNA-seq, we first calculated a joint embedding of the scRNA-seq and MERFISH data using only the 434 genes that overlap across both assays and MERFISH gene panels. Unsupervised Leiden clustering was performed with default parameters. These unsupervised cluster labels were then used as the cluster parameter for Tangram in cluster mode. In an effort to minimize computational resource usage while maximizing biological information, the scRNA-seq reference was subset to the union of the following gene sets: (1) the 5,000 most variable genes identified by the scanpy sc.pp.highly_variable_genes function, (2) the 562 genes in the MERFISH panels and (3) 1,623 genes that comprise ligands or receptors in an L–R pair from the CellChat or CellPhoneDB databases[72,73], resulting in a subset of 5,681 genes to be imputed. Tangram was run independently on each MERFISH imaging batch, and the results were concatenated for downstream analysis.

L–R analysis. We applied CellChat (v2.1.2)[72] to the MERFISH dataset on a per-neighborhood basis to identify spatially prioritized L–R interactions, restricting interactions to cell pairs within 200 µm of each other. To compensate for the limited gene panel in MERFISH, we also ran CellChat on the scRNA-seq samples derived from full-thickness biopsies (ten studies). Here we simulated neighborhood interactions on a transcriptome-wide basis with a procedure similar to that described in ref. 18. Briefly, for each neighborhood, a list of resident cell types was identified as cell types that comprised at least 1% of the total cell abundance in that neighborhood. For each neighborhood, the scRNA-seq was subset to only these resident cell types and CellChat was used to identify interactions on only the cell types present in the given neighborhood. Finally, to cross-validate the simulated neighborhood signaling in scRNA-seq, we cross-referenced CellChat hits that overlapped across MERFISH interactions and the scRNA-seq interactions.

MERFISH spatial L–R coexpression scoring. To visualize L–R expression, we used a *k*-nearest-neighbors smoothing approach to coexpression. For each cell, we took the geometric mean of the L–R expression for each centroid cell and its five nearest neighboring cells within 200 µm of spatial distance. This approach combined spatial prioritization with well-established approaches for scoring L–R signaling[74,75]. For ligands or receptors with heteromeric subunits, we used the minimum of the subunits, an approach used in several other tools such as liana[75].

For pseudobulked L–R coexpression analysis on a per-neighborhood basis, the L–R scores for the centroids were averaged across cells in each neighborhood on a per-sample basis. Downstream differential L–R analysis was performed using mixed linear modeling with dream on the pseudobulked L–R scores[35,36]. These approaches were applied to both the Tangram-imputed expression and the actual measured expression for the MERFISH data.

Integration of public Visium data. Publicly available Visium data were downloaded from six studies across 81 samples and 63 donors covering normal/nonlesional skin, BCC, SCC, HS, AD and psoriasis (Supplementary Table 10)[14,17,20,22,28,45,46]. Spots with nCount_Spatial of <10 were filtered out. The data were then processed with the Seurat workflow, regressing out nFeature_Spatial the ScaleData() step. Finally, the samples were integrated with Harmony with sample ID as the batch variable. Unsupervised clustering was performed through the FindClusters() function in Seurat with a resolution of 0.6, and gene expression markers for each cluster were identified with the FindAllMarkers() function.

Visium cell2location. The scRNA-seq object was filtered for genes that are among the 10,000 most highly variable in at least ten donors, resulting in an object with 14,762 genes. The scRNA-seq object was then used to train the cell2location model with 'n_epochs' = 250, and the per-cluster means workflow was used for the model according to the cell2location tutorial[47]. Cell2location (v0.1.4) was then run on each Visium tissue sample separately following default parameters, and the results were concatenated for downstream analysis[47].

Visium neighborhood mapping. To map the multicellular neighborhoods defined by MERFISH onto Visium data, we performed differential expression analysis of the MERFISH data to identify top expressed genes within each neighborhood using scanpy. We took the top 25 genes by log fold change determined by scanpy per neighborhood and applied the AddModuleScore() function in Seurat to calculate a neighborhood score for each Visium spot. We calculated the median neighborhood score per cluster and visualized these results in a clustered heatmap as shown in Extended Data Fig. 10c.

Visium differential abundance. Similar to cell types in MERFISH differential abundance analyses, for Visium, spots annotated by the 20 clusters or annotated neighborhoods were used as input into crumblr[69]. Differentially abundant spot clusters relative to NS samples were determined for each disease. Study and donor ID were also included in the model with random effects to account for batch and donor-related artifacts.

Statistics and reproducibility. Statistical analyses were performed with R (v4.4.0) and Python (v3.12.9). Parameters such as number of replicates, the number of independent experiments, measures of center, dispersion and precision (mean ± s.d. or s.e.m.), statistical test and significance are reported in Figs. 2, 4–6 and Extended Data Figs. 1, 5, 6, 8–10, and figure legends. No statistical method was used to predetermine sample size, and all samples were included in analyses wherever possible. All computational analyses are provided as code vignettes in the Zenodo repository with this study[70].

Reporting summary

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

Data availability

MERFISH and histology image data are available at the EMBL-EBI BioImage Archive (accession S-BIAD2376). Processed MERFISH data are available on Zenodo (https://doi.org/10.5281/zenodo.16795569)[70] and

an interactive website to explore the data is available online at https://rstudio-connect.hpc.mssm.edu/humanskin-spatialcensus/. Single-cell RNA-seq data and Visium data analyzed in this study were acquired from previously published datasets through GEO (accessions GSE144236 (ref. 17), GSE202352 (ref. 15), GSE153760 (ref. 24), GSE173205 (ref. 25), GSE129611 (ref. 30), GSE151091 (ref. 27), GSE156326 (ref. 29), GSE158955 (ref. 28), GSE173651 (ref. 17), GSE144239 (ref. 20), GSE197023 (ref. 46) and GSE225475 (ref. 45)), ArrayExpress (accessions E-MTAB-13085 and E-MTAB-13084)[14], the Genome Sequencing Archive (accession HRA000395)[21], the European Genome-Phenome archive (accession EGAS00001002927)[26] and Mendeley Data (https://data.mendeley.com/datasets/2bh5fchcv6/1)[22]. Previously published scRNA-seq data from ref. 10 were provided courtesy of R. Lafyatis. Previously published spatial data from ref. 28 was provided courtesy of C. Lu. GTEx bulk RNA-seq raw counts matrices of lower-leg and suprapubic human skin, as well as scRNA-seq data for normal human skin, were obtained from the GTEx portal (https://www.gtexportal.org/home/downloads/adult-gtex/bulk_tissue_expression and https://www.gtexportal.org/home/downloads/adult-gtex/single_cell)[23].

Code availability

No original software is used to analyze data in this paper. All data analysis scripts use existing methods and are provided in a GitHub repository provided with this paper (https://github.com/paularstrpo/ns-atlas-analysis) and on Zenodo (https://doi.org/10.5281/zenodo.16795569)[70].

References

58. Köten, B. et al. RNase 7 contributes to the cutaneous defense against *Enterococcus faecium*. *PLoS ONE* **4**, e6424 (2009).
59. Shannon, J. L., Kirchner, S. J. & Zhang, J. Y. Human skin explant preparation and culture. *Bio Protoc.* **12**, e4514 (2022).
60. Neil, J. E., Brown, M. B. & Williams, A. C. Human skin explant model for the investigation of topical therapeutics. *Sci. Rep.* **10**, 21192 (2020).
61. Banno, T., Gazel, A. & Blumenberg, M. Effects of tumor necrosis factor-α (TNF α) in epidermal keratinocytes revealed using global transcriptional profiling. *J. Biol. Chem.* **279**, 32633–32642 (2004).
62. Gubernatorova, E. O. & Tumanov, A. V. Tumor necrosis factor and lymphotoxin in regulation of intestinal inflammation. *Biochemistry* **81**, 1309–1325 (2016).
63. Wiggin, T. & Yu, C. Vizgen/vizgen-postprocessing. *GitHub* https://github.com/Vizgen/vizgen-postprocessing (2024).
64. Pachitariu, M. & Stringer, C. Cellpose 2.0: how to train your own model. *Nat. Methods* **19**, 1634–1641 (2022).
65. Luecken, M. D. et al. Benchmarking atlas-level data integration in single-cell genomics. *Nat. Methods* **19**, 41–50 (2022).
66. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat. Methods* **15**, 1053–1058 (2018).
67. Hao, Y. et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* **42**, 293–304 (2024).
68. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
69. Petukhov, V. et al. Cell segmentation in imaging-based spatial transcriptomics. *Nat. Biotechnol.* **40**, 345–354 (2022).
70. Restrepo, P. A single-cell spatial census of human skin anatomy. *Zenodo* https://doi.org/10.5281/zenodo.16795569 (2025).
71. Zappia, L. & Oshlack, A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *Gigascience* **7**, giy083 (2018).
72. Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).
73. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).
74. Dimitrov, D. et al. Comparison of methods and resources for cell-cell communication inference from single-cell RNA-seq data. *Nat. Commun.* **13**, 3224 (2022).
75. Dimitrov, D. et al. LIANA+ provides an all-in-one framework for cell–cell communication inference. *Nat. Cell Biol.* **26**, 1613–1622 (2024).

Acknowledgements

We thank members of the Ji laboratory for helpful discussions. This work was supported by the NIH (K08CA263187 to A.L.J. and T32AR082315 to A.H. and R.G.) and Damon Runyon Cancer Research Foundation Clinical Investigator award (121-23 to A.L.J.). This research was conducted with the support of the Biorepository and Pathology CoRE at the Icahn School of Medicine at Mount Sinai, the Bioinformatics for Next Generation Sequencing shared resource facility within the Tisch Cancer Institute at the Icahn School of Medicine at Mount Sinai, which is partially supported by NIH (grant P30CA196521), and in part by a pilot grant from the NIAMS/NIH SBDRC at Mount Sinai (grant P30AR079200). This work was supported in part through the computational and data resources and staff expertise provided by Scientific Computing and Data at the Icahn School of Medicine at Mount Sinai and by the Clinical and Translational Science awards (grant UL1TR004419) from the National Center for Advancing Translational Sciences. Research reported in this publication was also supported by the Office of Research Infrastructure of the NIH under awards S10OD026880 and S10OD030463. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The data in this paper were used in a dissertation as partial fulfillment of the requirements for a PhD degree at the Graduate School of Biomedical Sciences at Mount Sinai.

Author contributions

P.R. and A.L.J. conceived the project and performed the bioinformatic analysis. A.W., A.H., H.S.S., A.N., I.I., A.K., L.C., R.G., M.G.H., A.S. and S.M. performed experiments and analyzed the data. A.R., D.D. and D.H. built the web resource. E.G.-K., R.B., B.U., M.K., C.P.L., P.T., J.M.L., S.G. and S.M. interpreted the data. A.L.J. guided experiments, data analysis and interpretation. P.R. and A.L.J. wrote the paper with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41588-026-02552-8.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41588-026-02552-8.

Correspondence and requests for materials should be addressed to Andrew L. Ji.

Peer review information *Nature Genetics* thanks Matthew Vesely and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.