JIDOpen

Integrated Single-Cell Analysis Reveals Spatially and Temporally Dynamic Heterogeneity in Fibroblast States during Wound Healing

Axel A. Almet[1,2], Yingzi Liu[3,4], Qing Nie[1,2,3,5] and Maksim V. Plikus[2,3,4,5]

Wound healing is a dynamic process over temporal and spatial scales. Key to repair outcomes are fibroblasts; yet, how they modulate healing across time and in different wound regions remains incompletely understood. By integrating single-cell RNA-sequencing datasets of mouse skin and wounds, we infer that fibroblasts are the most transcriptionally dynamic skin-resident cells, evolving during postnatal skin maturation and rapidly after injury toward distinct late scar states. We show that transcriptional dynamics in fibroblasts are largely driven by genes encoding extracellular matrix and signaling factors. Lineage trajectory inference and spatial gene mapping reveal that *Prg4*-expressing fibroblasts transiently emerge along early wound edges. Within days, they become replaced by long-lasting and likely noninterconverting fibroblast populations, including *Col25a1*-expressing and *Pamr1*-expressing fibroblasts that occupy subepidermal and deep scar regions, respectively, where they engage in reciprocal signaling with immune cells. Signaling inference shows that fibroblast–immune crosstalk repeatedly uses some signaling pathways across wound healing time, whereas use of other signaling pathways is time and space limited. Collectively, we uncovered high transcriptional plasticity by wound fibroblasts, with early states transiently forming distinct microniches along wound edges and in the fascia, followed by stable states that stratify scar tissue into molecularly dissimilar upper and lower layers.

Keywords: Data integration, Fibroblasts, Scar, Single-cell RNA sequencing, Wound healing

Journal of Investigative Dermatology (2024) ■, ■–■; doi:10.1016/j.jid.2024.06.1281

INTRODUCTION

Skin is the largest organ in the body by surface area and self-renews rapidly. Adult skin is anatomically and functionally complex, consisting of various cell populations and extracellular matrix (ECM) that are in delicate balance to maintain dynamic homeostasis. To respond to sudden environmental insults, such as wounding, several new molecular and cellular processes rapidly activate. These processes, collectively referred to as the wound healing program, replace lost skin with scar tissue (Gurtner et al, 2008). However, from a morphological and functional perspective, scars are a suboptimal substitute for normal skin. Although the epidermal barrier and mechanical integrity of the skin are restored, many functionally important components typically fail to reform. Scars also commonly consist of an excess of deposited ECM, which can lead to skin morbidities. Nonetheless, wound healing with a scar is critical for animal survival, and deeper understanding of the cellular and molecular interplay that produces a scar is required.

A pivotal cell type that drives wound healing are fibroblasts, which transiently activate to proliferate, synthesize new ECM, and exert contractile forces to pull wound edges closer together and fill the remaining defect with new scar tissue. Before wounding, fibroblasts are relatively quiescent because they seldom proliferate, but upon wounding, they activate to participate in repair. Triggered by injury-associated signaling cues, fibroblasts activate a 3-phase program involving (i) rapid migration from wound edges into the site of skin loss, (ii) transient proliferation to multiply in numbers, and (iii) ECM deposition and its remodeling (Darby et al, 2014). Wound-triggered fibroblasts also convert into myofibroblasts, a specialized contractile state. It remains a matter of debate whether myofibroblasts revert to their prewounded state after completion of healing, permanently remain differentiated, or progress to a new dormant state of established scar. A study by Plikus et al (2017) used bulk RNA-sequencing data on sorted fibroblasts sampled from multiple postwounding time points in mice to show that the averaged fibroblast transcriptome is dynamic and that the myofibroblast state appears transiently. What the bulk RNA-

[1]*Department of Mathematics, University of California, Irvine, Irvine, California, USA;* [2]*NSF-Simons Center for Multiscale Cell Fate Research, University of California, Irvine, Irvine, California, USA;* [3]*Department of Developmental and Cell Biology, University of California, Irvine, Irvine, California, USA; and* [4]*Sue and Bill Gross Stem Cell Research Center, University of California, Irvine, Irvine, California, USA*

[5]*These authors contributed equally to this work.*

Correspondence: Qing Nie, Department of Mathematics, University of California, Irvine, Irvine, California, USA. E-mail: qnie@uci.edu and Maksim V. Plikus, Department of Developmental and Cell Biology, University of California, Irvine, Irvine, California, USA. E-mail: plikus@uci.edu

Abbreviations: ARI, Adjusted Rand Index; ECM, extracellular matrix; LW, large wound; NMI, normalized mutual information; PWD, postwounding day; scRNA-seq, single-cell RNA sequencing; SW, small wound

Received 31 January 2023; revised 14 June 2024; accepted 26 June 2024; accepted manuscript published online XXX; corrected published online XXX

© 2024 The Authors. Published by Elsevier, Inc. on behalf of the Society for Investigative Dermatology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

sequencing data could not reveal is whether the observed transcriptome changes were driven by a single fibroblast state converting to another or by changes across several heterogeneous cell states. Because scar tissue is a morphological and functional substitute of normal skin, it is also essential to understand whether scar-resident cells have the same characteristics and functions as normal skin-resident cells or whether they are mere substitute cells. Answering this question can lead to future scar-reducing and even scarless tissue repair therapies.

Single-cell RNA sequencing (scRNA-seq) enables one to characterize the transcriptional heterogeneity of cell populations at unprecedented resolution and identify previously unknown cell states. Several recent scRNA-seq studies investigated transcriptomic changes in murine skin at various time points after full-thickness wounding, revealing insights about the healing process (Abbasi et al, 2020; Gay et al, 2020; Guerrero-Juarez et al, 2019; Haensel et al, 2020; Hu et al, 2023; Konieczny et al, 2022; Mascharak et al, 2022; Phan et al, 2020; Shook et al, 2020; Vu et al, 2022). However, commonly, only one postwounding time point is considered, offering static snapshots of the sequenced cell population. Yet, wound healing is dynamic in time and space, consisting of multiple synergistic and competing processes. Therefore, studying it comprehensively requires sampling from numerous postwounding time points (Mascharak et al, 2022).

In this work, we integrated several scRNA-seq datasets on murine skin to construct a joint timeline of wound healing dynamics (Supplementary Text S1 provide extended discussion). We focused on fibroblasts, which exhibited the most dynamic transcriptional changes across wound healing. We dissected molecular contributors to fibroblast heterogeneity, inferring that fibroblasts exhibit several distinct yet transient states across time and that scar-resident fibroblasts do not return to the prewounded state. We inferred that transcriptional fibroblast heterogeneity is dominantly driven by changes in key fibroblast-specific functionalities, including matrisome (ie, structural ECM factors) and secretome (ie, cell-bound and diffusible signaling factors). Finally, lineage trajectory inference and marker gene staining experiments suggested that early on in wound healing, fibroblasts occupy distinct microniches along wound edges before assuming new stable transcriptional states mid-way through wound healing that then do not interconvert in late wounds.

RESULTS

Shared cell types and diverse cell states characterize wound healing

We analyzed scRNA-seq data of murine skin sampled from different time points across wound healing (Figure 1a). We considered both juvenile, postnatal day 21 and young adult, postnatal day 49 unwounded skin because the former is the earliest recorded age for which scRNA-seq data of stereotypical healing with a scar are available. All wounded conditions were collected from full-thickness excisional wounds, where earlier time points were generated from small wounds (SWs) (<1 cm diameter) and later time points from large wounds (LWs) (>1 cm diameter). For LWs, the increased size slows the biological timescale of repair events, meaning that late calendar times in LWs, from postwounding day (PWD) 12 to PWD18, are comparable with earlier biological times along a SW healing timescale. Therefore, earlier SW and later LW time points can be reasonably integrated. We followed the original study on LW PWD18 and separated cells sampled from regenerative (denoted as REG) outcomes, where wound-induced hair neogenesis occurred, from cells sampled from fibrotic (denoted as FIB) outcomes containing hairless scars (Gay et al, 2020) (details are provided in Datasets and data preprocessing under Supplementary Materials and Methods). We integrated the datasets using Scanorama (Hie et al, 2019). Figure 1b suggested that there is significant transcriptomic heterogeneity. Across the islands of diverse skin cells in the neighborhood embedding, there was a spectrum of mixing across wound healing time. Some islands were well-mixed and contained cells from multiple time points, whereas other islands contained cells from a single time point predominantly.

We clustered the integrated scRNA-seq data (Traag et al, 2019), setting resolution = 0.3, obtaining 20 distinct clusters (Figure 1c). Cell-type annotations were based on differentially expressed genes reported in published murine skin scRNA-seq analyses and consensus skin marker genes from non−scRNA-seq literature. We identified epidermal keratinocytes (Supplementary Figure S1); fibroblasts (Figure 2a); vascular cells, consisting of pericytes (Supplementary Figure S2a–d), endothelial cells (Supplementary Figure S2e–h), lymphatic endothelial cells, and vascular smooth muscle cells (Figure 1c); and dermis-resident immune cells (Supplementary Figure S3). Additional less abundant cell populations included Schwann cells (Supplementary Figure S4), melanocytes, epidermis-associated Langerhans cells, and skeletal muscle cells (Figure 1c) (details are provided in Supplementary Text S2 and Subclustering under Supplementary Materials and Methods).

To evaluate the efficacy of integration, we used scIB (Luecken et al, 2022). The results were plotted in Figure 1d, where metric values closer to 1 indicate good performance (details are provided in Data integration under Supplementary Materials and Methods). Several metrics assessing batch effect removal—Average Silhouette Width batch score (0.93) and graph connectivity score (1.0)—and those assessing biological conservation—normalized mutual information (NMI) cluster label score (0.92) and isolated label F1 score (0.98)—performed well, indicating that cells within the same types are well-mixed and that rarer cell types are preserved across conditions. However, low kBET (k-Nearest Neighbor Batch Effect Test) score (0.14) indicated disparities between local and the neighborhood structure, whereas moderate principal component regression batch, Average Silhouette Width label, and isolated label silhouette scores (0.52, 0.57, and 0.56, respectively) reflected variability in cell-type integration across batches. Plotting cell-type composition across wound healing time (Figure 1e) showed that few cell types were universally present across all times. This likely resulted from both the real biological differences in wound's cell composition across healing time and from technical issues arising from the lack of a universal cell isolation protocol, whereby certain cell types are captured for scRNA-seq with low efficiency, depending on the protocol's details.

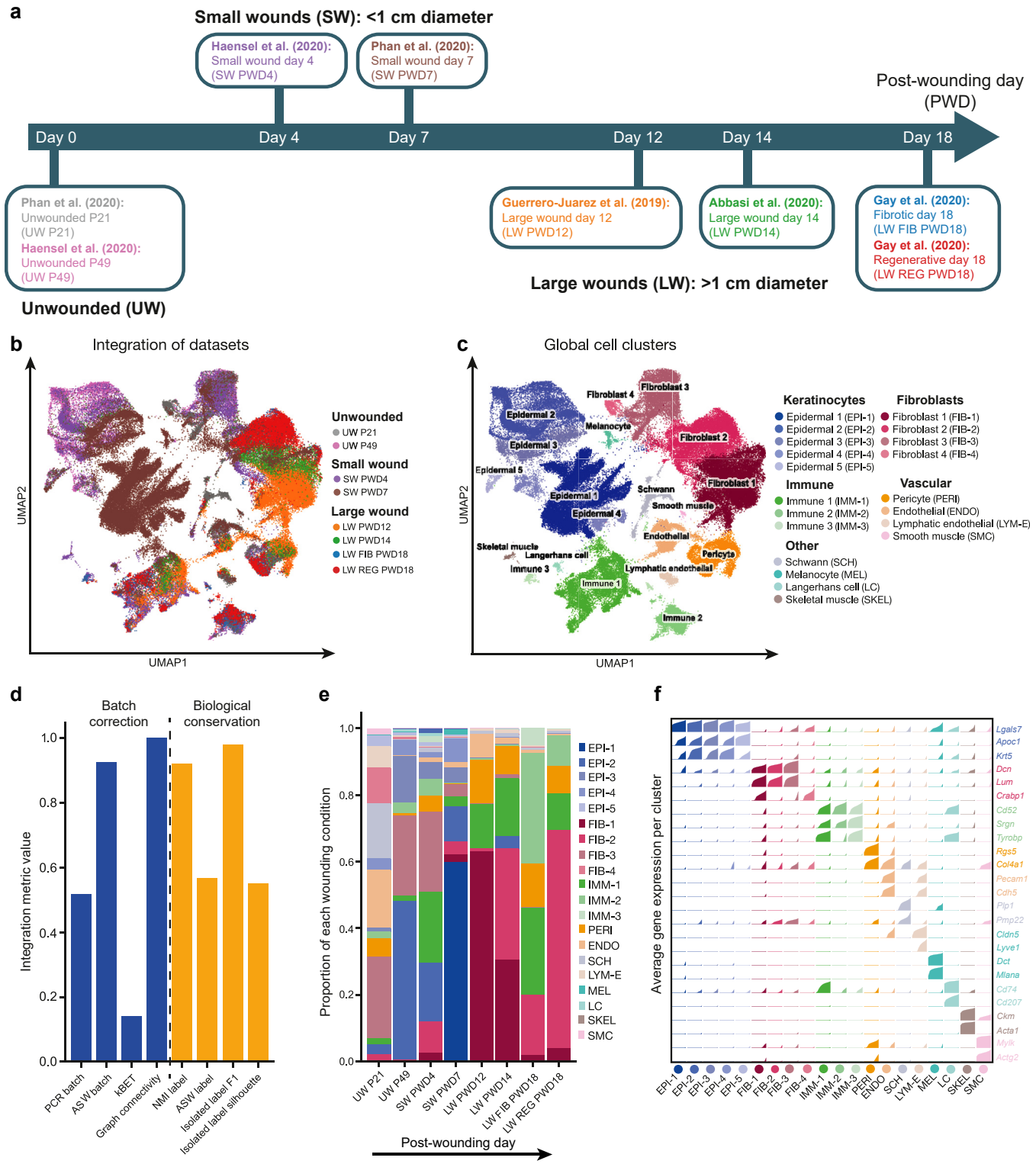


Figure 1. Integrated analysis of scRNA-seq datasets on unwounded mouse skin and wound tissue across biological timescale of healing. (**a**) Schematic representation of the bioinformatic analysis timeline employed in this study, showing relative position of all integrated scRNA-seq datasets along the biological timescale of skin wound healing. Literature sources of all scRNA-seq datasets are references. (**b**) UMAP plot of the integrated scRNA-seq datasets. Cells are color coded and annotated by their source of origin (unwounded skin, small wound, large wound) and biological time point (listed on the right). (**c**) UMAP plot of the initial round of clustering and assigned cell type labels. We categorized all cells in the integrated scRNA-seq dataset into 5 major types: keratinocytes, fibroblasts, immune cells, vascular cells, and other skin-resident cells. Cells are color coded according to annotations shown on the right. (**d**) Metrics calculated to assess the quality of scRNA-seq data integration by batch correction (blue bars on the left) and biological conservation (orange bars on the right). (**e**) Relative cell-type composition of individual scRNA-seq datasets arranged according to study timeline from **a** and annotated along the x-axis. Stacked bars show relative abundance of distinct cell types and subtypes, as clustered on **c**. Bar segments are colored and annotated according to **c**. (**f**) Track plots of the discriminatory marker genes for each annotated cell type and subtype from **c**. Color coding is according to **c**. UW denotes unwounded skin. ASW, Average Silhouette Width;

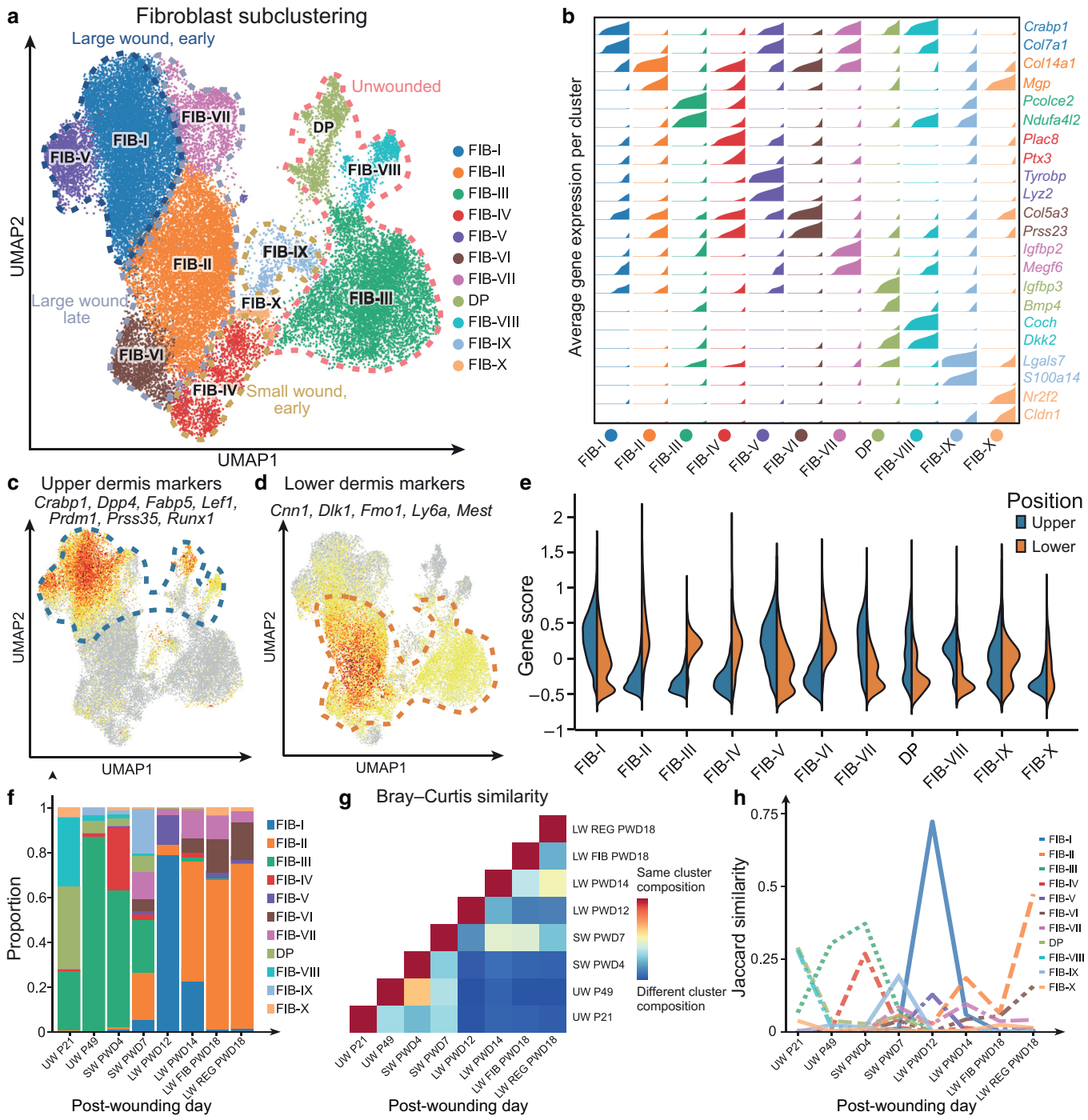


Figure 2. Fibroblast heterogeneity is dominantly driven by wound condition. (**a**) UMAP plot of the identified fibroblast subclusters. Subclusters are color coded according to annotations shown on the right. Groups of subclusters on the UMAP plot are additionally annotated with the predominant source of origin (unwounded skin; small wound, early; large wound, early and late). (**b**) Track plots of unique differentially expressed genes for each fibroblast subcluster from **a**. Annotation and color coding are according to **a**. (**c**) Feature plot of the upper dermis gene score expression. Orange indicates gene score–expressing fibroblasts. Genes that constitute upper score are listed at the top. (**d**) Feature plot of the lower-dermis gene score expression. Orange indicates gene score–expressing fibroblasts. Genes that constitute lower score are listed at the top. (**e**) Split violin plot indicating that several fibroblast subclusters are enriched primarily for either upper- or lower-dermis gene score identity. (**f**) Relative fibroblast subcluster composition within individual scRNA-seq datasets arranged according to biological time point from Figure 1b and annotated along the x-axis. Stacked bars show relative abundance of distinct fibroblast subclusters, as detailed on **a**. Bar segments are colored and annotated according to **a**. (**g**) Calculated Bray–Curtis similarity scores between wound healing conditions according to Figure 1b (x-axis), measuring the compositional overlap with respect to fibroblast subclusters as defined along y-axis according to **a**. (**h**) Jaccard similarity between fibroblast subcluster labels, color coded and annotated according to **a**, and wound healing condition according to Figure 1b (x-axis). UW denotes unwounded skin. LW, large wound; PWD, postwounding day; scRNA-seq, single-cell RNA sequencing; SW, small wound; UMAP, Uniform Manifold Approximation and Projection.

LW, large wound; PWD, postwounding day; scRNA-seq, single-cell RNA sequencing; SW, small wound; UMAP, Uniform Manifold Approximation and Projection.

Considering large number of integrated cells, we used this opportunity to identify groups of differentially expressed markers that distinguish between skin cell lineages. Figure 1f showed the binned gene expressions of such markers, revealing genes that can be used for reliable annotation of major mouse skin cell lineages. For example, we found that *Lgals7*, *Apoc1*, and *Krt5* were shared across all epidermal clusters, EPI-1 through 5, and were specific to those clusters. On the other hand, canonical mesenchymal markers, such as *Col1a2* or *Col3a1*, were not specific to all fibroblasts. Rather, *Dcn* and *Lum* were differentially expressed in FIB-1 and 3, and *Crabp1* was shared across both FIB-1 and FIB-4. In addition, *Cd52*, *Cd74*, and *Tyrobp* were differentially expressed for immune clusters (IMM-1 through 3) and Langerhans cell, but *Cd207* was expressed only in Langerhans cells, distinguishing dermis-resident from epidermis-resident immune clusters. We propose that the gene set shown in Figure 1f can be used for improved annotation of mouse skin cell lineages.

Fibroblasts are transcriptionally dynamic across wound healing time

Of all cell lineages, fibroblasts integrated the least across conditions (Figure 1b), suggesting that they exhibited the greatest degree of heterogeneity across wound healing. All fibroblasts divided into 11 subclusters, including hair follicle–associated dermal papilla cells (Figure 2a and also Subclustering under Supplementary Materials and Methods). These subclusters were confirmed as distinct by identification of marker genes (Figure 2b). We then sought to discern the driving factors of fibroblast heterogeneity, including (i) spatial location in the dermis, (ii) biological wound healing time, and (iii) cell cycle state. First, it has been suggested that normal dermal fibroblasts' identity is driven by whether they are located in upper papillary or lower reticular dermis (Driskell et al, 2013). This notion is supported by the fact that certain marker genes indeed distinguish upper from lower fibroblasts well (Philippeos et al, 2018; Plikus et al, 2021). Therefore, we compiled the lists of literature-supported upper and lower cutaneous fibroblast marker genes that were also present in our integrated data (Supplementary Table S1 and Supplementary Figure S5). We then calculated an upper and lower gene score for every fibroblast (Satija et al, 2015). Cells with higher upper scores were enriched in subclusters FIB-I, -V, -VII, and -VIII and dermal papilla (Figure 2c and e), corresponding predominantly to unwounded skin and LW PWD12 wound, whereas cells with higher lower scores were enriched in FIB-II, -III, -IV, and -VI (Figure 2d and e), corresponding predominantly to unwounded skin and both early-stage SW PWD4 and late-stage LW PWD18 wounds. An extended discussion of spatial gene scores is provided in Supplementary Text S3. We conclude that a spatial gene program significantly contributes to fibroblasts' identity, both prior to and after wounding.

We also observed significant compositional changes in fibroblasts across unwounded skin maturation and wound healing time (Figure 2f)—dynamics that were not paralleled in any other cell lineage present in the data (Supplementary Figures S1–4). For example, lower FIB-III fibroblasts increased in abundance upon maturation of unwounded skin (26.5% at postnatal day 21 and 87.1% at postnatal day 49) but subsequently decreased during early wound healing (61.2% in SW PWD4 and 23.6% in SW PWD7) and became rare at later postwounding time points (~0.1% in LW PWD18). We calculated pairwise Bray–Curtis scores (Figure 2g) to measure compositional similarity between wound healing times with respect to fibroblast subcluster labels. For a given pair of wound healing times, a Bray–Curtis score of 1 indicates that the 2 conditions have the same subcluster composition, whereas a score of 0 indicates a completely dissimilar composition. We observed that LW PWD12 was more similar in composition to LW PWD14 (Bray–Curtis score = 0.23) than to LW PWD18 fibroblasts (0.11 and 0.12 for fibrotic and regenerative PWD18 conditions, respectively). The highest Bray–Curtis similarity of 0.63 between unwounded postnatal day 49 and SW PWD4 samples likely arose from the fact that the latter included postnatal day 49 + 4-day wound edge skin (Haensel et al, 2020). The observed changes in Bray–Curtis scores suggested dynamic transcriptional states by fibroblasts across wound healing time. We also calculated Jaccard similarity scores between fibroblast subclusters and wound condition labels (Figure 2h). The highest score occurred between FIB-I and LW PWD12 (0.72), indicating that FIB-I subcluster consisted of almost entirely LW PWD12 fibroblasts. Almost all other fibroblast subclusters followed this trend, albeit not as dramatic, having significant overlap with effectively only 1 wound healing time. The sole exception to this trend was FIB-III, which has relatively high Jaccard similarities with both postnatal day 49 skin (0.3) and SW PWD4 wound (0.37) owing to the earlier-mentioned tissue sampling protocol. Subclusters that peaked in Jaccard similarities with the same wound condition—for example, FIB-I and FIB-V with LW PWD12 fibroblasts (0.72 and 0.13, respectively)—tended to be closer in the neighborhood embedding (Figure 2a), suggesting transcriptomic similarity. The isolated peaks in Jaccard similarity across wound healing time suggested that the latter was indeed the primary driver of fibroblast heterogeneity. We note that the dominant early wound cluster, FIB-IV (28.0% of SW PWD4), consisted primarily of lower identity fibroblasts. However, after re-epithelialization, LW PWD12 consisted primarily of upper identity fibroblasts (FIB-I and FIB-V cells comprising 91.0%). Toward late wound healing, upper identity fibroblasts were superseded by lower-identity FIB-II and FIB-VI fibroblasts (89.7%).

We also observed that certain fibroblast states were enriched for cell cycle phase scores associated with active cell division (Satija et al, 2015)—significant portions of FIB-I (45.7%), -IV (32.2%), -VI (51.1%), and -IX fibroblasts (52.6%) were predicted to be in S and G2M phases (Supplementary Figure S6a and b). Surprisingly, the fibroblast subclusters mentioned earlier were not exclusively associated with early wound healing times, suggesting late proliferative events (Supplementary Figure S6c). For instance, FIB-I cells, 90.0% of which originated from LW PWD12, were sampled after wound re-epithelialization, whereas 75.7% of highly proliferative FIB-VI cells originated exclusively from late wounds. Next, we examined proliferative patterns by fibroblasts in experimentally reproduced full-thickness wounds in C57BL/6J mice by staining scar tissues at matching postwounding

time points (details are provided in Replication wounding studies under Supplementary Materials and Methods) for an established proliferation marker *Mki67* and panfibroblast marker *Pdgfra* (Supplementary Figure S7). As expected, in early SW PWD4 wounds, we observed abundant *Mki67*$^+$ fibroblasts at wound edges (Supplementary Figure S7a). However, we also observed prominent clusters of *Mki67*$^+$/*Pdgra*$^+$ fibroblasts in late wounds, including in lower scar regions of LW PWD12 and PWD18 wounds (Supplementary Figure S7c and d). We posit that the late expansion of wound fibroblasts with lower spatial identity observed from integrated analysis (Figure 2f) could be conceivably attributed to late proliferative events by FIB-VI cells (Supplementary Figure S6) taking place in the lower scar region (Supplementary Figure S7c and d).

Thus far, our analysis broadly indicates that wound healing time is the major driver of fibroblast heterogeneity and that fibroblast states are transient during healing and do not return to the prewounding state. We partially validated these conclusions using a 2-pronged bioinformatic approach. First, we reanalyzed spatial transcriptomic datasets (10X Visium) collected on 6-mm dorsal mouse skin wounds at PWD2, 7, and 14 (Foster et al, 2021) (details are provided in the Analysis of fibroblast distribution from Foster et al [2021] under Supplementary Materials and Methods). Using keratinocyte markers *Lgals7* and *Krt5*, we validated the presence of wound epidermis, confirming that re-epithelialization was ongoing at PWD2 and 7 and was completed by PWD14, whereas scoring spatial data against fibroblast markers *Dcn*, *Lum*, and *Pdgfra* confirmed their broad presence in wounds throughout the entire healing time course (Figure 3a). Next, we scored spatial data for upper fibroblast marker *Crabp1* and lower fibroblast marker *Ly6a* (Figure 3b and c). *Crabp1* signal distribution revealed that at PWD2, upper identity fibroblasts are practically nonexistent but prominently emerge by PWD7 in upper, subepidermal wound space and persist in PWD14 scars (Figure 3b). In contrast, *Ly6a* signal was prevalent in wound tissue already at PWD2 and remained restricted to the lower scar region at PWD7 and 14 (Figure 3c). These reanalyses confirm that wound fibroblasts indeed assume distinct upper and lower transcriptional and corresponding spatial identities during wound healing. They also reveal that upper-identity fibroblasts emerge in SWs as early as PWD7 before re-epithelialization is complete (Figure 3b).

Next, we stained wound tissues for selected fibroblast marker genes (details are provided in Replication wounding studies under Supplementary Materials and Methods). We performed 2 sets of experiments. For one set, we experimentally reproduced mouse age and wound size from which our integrated scRNA-seq datasets were sampled. Second, we performed wounding experiments for a consistent mouse age (postnatal day 49) and wound size (6 mm). We limited our primary analysis to 2 fibroblast states—FIB-IV cells, which appear early and transiently after wounding, and FIB-I and -VII cells, which appear later and persist in healing wounds. We found that *Prg4* was a selective marker of FIB-IV fibroblasts (Figure 3d) and expressed only in early SW PWD4 tissues in the lower dermis, as confirmed from our analysis of consistent wound size experiments (Figure 3e and Supplementary Figure S8a). We also observed analogous dynamics when we replicated mouse age and wound size from the original scRNA-seq studies (Supplementary Figure S9). At PWD4, *Prg4*$^+$ cells were present in 2 locations—in the reticular dermis, immediately adjacent to the wound edge, and in the fascia, below the wound bed. This *Prg4* pattern is consistent with the possibility that FIB-IV cells are early wound response fibroblasts with lower identity. To this end, we also reanalyzed recently published scRNA-seq data collected on 6-mm dorsal mouse skin wounds at PWD2, 7, 14, and 30 (Mascharak et al, 2022) (details are provided in Analysis of fibroblasts from Mascharak et al [2022] under Supplementary Materials and Methods). Because only a total of 364 fibroblasts were recovered in the data mentioned earlier, we analyzed them separately to minimize clustering bias (Supplementary Figure S10a). *Prg4* was almost exclusively expressed by early PWD2 fibroblasts (Supplementary Figure S10b), most of which also featured lower identity (Supplementary Figure S10b). Of the 227 differentially expressed genes expressed by these PWD2 fibroblasts, nearly half were also differentially expressed genes of *Prg4*$^+$ FIB-IV fibroblasts from our original analysis, including numerous inflammatory pathway factors (Supplementary Figure S10d and e). Therefore, we posit that the early healing response involves transient emergence of inflammation-primed fibroblasts that occupy distinct spatial microniches around the wound bed periphery. Next, we examined the staining pattern of collagen *Col25a1*, which was enriched in late FIB-VII fibroblasts (Figure 3f). We observed from our staining analysis of consistent wound size experiments that *Col25a1*$^+$ fibroblasts emerged after re-epithelialization during the remodeling phase of wound healing starting from PWD8, primarily in the upper dermis (Figure 3g and Supplementary Figure S8b). From the wound replication study, we observed that they were absent in unwounded skin and in SW PWD4 wounds but appeared at PWD7 and persisted afterward in the upper wound compartment, including at PWD14 (Supplementary Figure S11). As an additional proof of principle, we also stained tissues from the consistent wound size experiment for *Pamr1*, which we observed from scRNA-seq analysis was largely restricted to late wound fibroblast states FIB-II and FIB-VI. Consistently, *Pamr1*$^+$ fibroblasts were detected starting from PWD8 and largely in the lower scar (Supplementary Figure S8c and d). These staining results strengthen our bioinformatic prediction that latter-born subsets of wound fibroblasts assume distinct anatomical positions within forming scar tissue, including *Col25a1*$^+$ upper cells, which are likely part of broader population of upper *Crabp1*$^+$ fibroblasts (Figure 3b) (Guerrero-Juarez et al, 2019).

Wounding induces changes in key fibroblast functions

Next, we examined how heterogenous fibroblasts differed with respect to key functional outputs: (i) secretome, which included signal ligand and signal receptor genes, and (ii) matrisome, which included ECM synthesis genes encoding for structural ECM proteins and ECM-modifying genes encoding for ECM-degrading and -remodeling enzymes. Each function was defined by a gene set manually curated from pre-existing databases and primary literature (details are

provided in Biased clustering under Supplementary Materials and Methods). We reanalyzed fibroblasts for each of the gene categories mentioned earlier and across all functional output genes, generating functionally biased clusters. We calculated the Adjusted Rand Index (ARI) and NMI to measure similarity between the functionality-based annotations and the original subcluster annotations (Figure 4a). The ARI and NMI scores of 0.68 and 0.67, respectively, showed that the clusters obtained on functional genes alone captured much of the original fibroblast heterogeneity, indicating that fibroblast states were driven primarily by changes in the functions mentioned earlier. Indeed, the neighborhood embedding produced on only functional output genes resembled the original neighborhood embedding (Figure 4b). Of the functional annotations, clustering on ECM synthesis genes had the highest overlap, with ARI (NMI) score of 0.60 (0.63). On this gene set alone, the neighborhood embedding retained many of the key structures observed in the original subclustering (Figure 4c). The next highest overlap was by signal ligand genes, with ARI (NMI) score of 0.48 (0.48), showing coarser neighborhood embedding (Figure 4d), whereas ECM-modifying and signal receptor genes produced low overlaps, with ARI (NMI) scores of 0.26 (0.37) and 0.25 (0.30), respectively, suggesting relatively static gene expression. Thus, we conclude that transcriptional heterogeneity of wound fibroblasts is largely driven by their gene product outputs into scar's matrisome and secretome.

To determine how fibroblasts varied with respect to the 3 most informative functions—ECM synthesis, signal ligand, and ECM-modifying genes—we used Intercode (Rybakov et al, 20201)[1] to learn latent representations of the gene expression matrix with respect to decomposed factors defined by the annotated functional gene sets (details are provided in Biased clustering under Supplementary Materials and Methods). We plotted fibroblasts with respect to these factor loadings in Figure 4e–g. Labeling fibroblasts by the original subcluster labels revealed varying degrees of separation with respect to each functionality pair. Of the pairs, fibroblast population separation with respect to ECM synthesis and signal ligand was most similar to that observed for the full gene set. To validate that changes in ECM synthesis and signal ligand gene expression correspond to changes in fibroblast states, we performed differential expression analysis using MAST (Finak et al, 2015). We then plotted gene expression of differentially expressed genes across time, finding several universal ECM synthesis genes (Figure 4h) and signaling ligand genes (Figure 4i) as well as dynamically expressed genes. Many collagen genes were expressed by all fibroblasts; yet, certain ones were unique to a specific time or a subset of times, such as *Col23a1* in postnatal day 21 skin; *Col7a1* in PWD12 wounds; and *Col8a1*, *Col12a1*, *Col14a1*, and *Col25a1* in late PWD12 through 18 wounds. We also observed significant expression changes in several non-collagen ECM synthesis genes—cochlin in postnatal day 21 skin, *Rspo1* in postnatal day 49 skin, *Prg4* in SW PWD4, *Tnn* in LW PWD12, and *Matn4* in LW PWD18. Expression of signal ligand genes was also dynamic (Figure 4i), notably across chemokines (*Ccl* and *Cxcl*) between unwounded and wounded skin as well as other ligands implicated in tissue scarring (Supplementary Text S4), including *Bmp7* in postnatal day 21 skin, *Ccl19* in postnatal day 49 skin, *C4b* in SW PWD4, *Cdh1* in SW PWD7, *Il1b* in LW PWD12, and *Vegfd* in LW PWD18 fibroblasts.

Highly dynamic transcriptional state changes by wound fibroblasts suggested dynamic underlying gene regulatory networks. To infer significant gene regulatory networks, we performed SCENIC analysis (Van de Sande et al, 2020) (details are provided in Regulon inference under Supplementary Materials and Methods), which separated fibroblasts into 12 regulon-based clusters (Supplementary Figure S12a). Many of these clusters significantly overlapped with the original transcriptomic-based subclusters (Supplementary Figure S12b), whereas others better correlated with the wound healing time (Supplementary Figure S12c). Furthermore, many of the differentially expressed functional output genes discussed earlier, both ECM and signaling factors, were predicted to be under shared control by the same regulons (Supplementary Figure S12d and e and extended discussion in Supplementary Text S5). Collectively, we conclude that wound-triggered fibroblast state changes are largely driven by a combination of changes in ECM production and secreted signals that in turn are driven by unique gene regulatory network states.

Pseudotime analysis predicts parallel transitions into divergent states by wound fibroblasts

The analyses provided earlier showed that fibroblasts are transcriptionally dynamic over wound healing time and spatially segregate upon scar maturation. It remains unresolved whether distinct transcriptional fibroblast states originate from distinct progenitor sources or whether, instead, they are continuously lineage connected. To comprehensively answer this question will require genetic lineage tracing, an effort that is ongoing in the field. However, the available single-cell data offer an opportunity to predict possible lineage relationship between fibroblast states using bioinformatic inference tools (Figure 5a). In this study, we performed pseudotime analysis on all fibroblasts using scVelo (Bergen et al, 2020) and PAGA (Wolf et al, 2019). We projected the estimated velocities and directional flows onto the original Uniform Manifold Approximation and Projections, labeling both fibroblast subclusters (Figure 5b) and wound healing times (Figure 5c). RNA velocity streams in Figure 5b suggest transitions from FIB-III state to dermal papilla and FIB-VIII and from FIB-II state and FIB-VII to FIB-I. We deemed these predictions to be likely erroneous because they represent transitions from postnatal day 49–enriched to postnatal day 21–enriched fibroblasts and from PWD18-enriched to PWD12-enriched fibroblasts (Figure 5c). Next, we calculated PAGA graphs with partitions defined either by fibroblast subclusters (Figure 5d) or by wound healing times (Figure 5e). High connectivity of the former graph suggested possible interconversion between fibroblast states, whereas the latter graph indicated that transitions were strongest between conditions connected along the biological timescale. We then used RNA velocity–inferred pseudotime (details are

[1] Luecken MD, Büttner M, Chaichoompu K, Danese A, Interlandi M, Mueller MF, et al. Benchmarking atlas-level data integration in single-cell genomics. bioRxiv 2020:1–61.

provided in Trajectory analysis under Supplementary Materials and Methods), which estimated the following transition directions as most plausible (Figure 5f): (i) transition from FIB-III fibroblasts (lower fibroblasts abundance in unwounded skin) to FIB-IX and then to *Prg4*$^{+}$ FIB-IV fibroblasts (both transient early wound fibroblasts), (ii) from FIB-IX to FIB-X fibroblasts (spatially ambiguous fibroblasts of late wounds), (iii) from FIB-IX to FIB-VII and then to FIB-I and FIB-V fibroblasts (upper wound fibroblasts that appear early during healing), and (iv) independently from FIB-IX to FIB-II and then to FIB-VI fibroblasts (lower wound fibroblasts that appear early in healing and expand in number over time). Importantly, no transitions were inferred between early upper wound clusters (FIB-I, -V, -VII) and lower wound clusters (FIB-II, -VI). This strengthens the theory that (i) upper and lower wound fibroblasts originate from a common early wound progenitors (FIB-IX, -IV), which themselves likely originate from lower dermis fibroblasts, and that (ii) once upper and lower fibroblasts quickly establish in forming scar, they remain separate over the healing process. Together, the bioinformatic analyses and staining assays mentioned earlier suggest that skin injury triggers emergence of early fibroblast progenitors along wound edges, which are transient and become quickly superseded by long-lasting and non-interconverting fibroblast states of upper subepidermal and lower anatomically deep scar layers.

Wound fibroblasts engage in dynamic signaling crosstalk with immune cells

We hypothesized that the observed transcriptional dynamics of wound fibroblasts could be regulated by dynamic signaling interactions with wound-resident immune cells. Therefore, we employed CellChat (Jin et al, 20232, 2021)[2] to infer the most significant cell−cell communication activities between fibroblast subclusters (Figure 2a) and immune subclusters (Supplementary Figure S3a) across wound healing time (Figure 6a and Cell−cell communication analysis under Supplementary Materials and Methods). We distinguished signaling pathways that were directed from immune cells to fibroblasts (Figure 6b), signaling ligands secreted by fibroblasts to other fibroblasts (Figure 6c), and those exclusively driven by fibroblasts to immune cells (Figure 6d). We identified 5 unique immune-to-fibroblast signaling pathways, which exhibited 3 types of signaling patterns (Figure 6b). In contrast, we identified only one pathway that was exclusively driven from fibroblasts to other fibroblasts, PROS, which is dynamically bimodal, activating in early wounds and reactivating in late wounds (Figure 6c). Categorizing active signaling pathways into 3 temporal modes based on when they were active and for how long revealed that fibroblasts and immune cells secreted distinct sets of signaling ligands with different dynamics across wound healing (Supplementary Figure S13 and extended discussion in Supplementary Text S6). This CellChat analysis provides rich resources of predictive data on potential new signaling regulators of wound healing that will require confirmation in future functional experiments.

[2] Sun X, Annusver K, Dalessandri T, Kasper M. Rare Gli1+ perivascular fibroblasts promote skin wound repair. bioRxiv 2022.

DISCUSSION

Scar tissue is morphologically and functionally distinct from normal unwounded skin. Yet, at the cellular scale, the degree to which scar-resident cells differ from unwounded skin-resident cells remains poorly understood. Furthermore, the transcriptional trajectories that normal skin cells take en route their scar-resident progenies during healing remain largely unknown. To start addressing these questions, we constructed an integrated scRNA-seq timeline of wound healing in murine skin, focusing on fibroblasts, the essential cell type in cutaneous repair. Previous scRNA-seq studies of fibroblasts have mainly focused on identifying generalizable fibrotic states across diverse tissues and organs (Buechler et al, 2021; Forte et al, 2022) and on fibroblast heterogeneity across aging (Solé-Boldo et al, 2020). In the context of skin wound healing, Phan et al (2021) have studied fibroblast state differences between SWs and LWs, whereas Haensel et al (2020) and Joost et al (2018) examined early wound-triggered dynamics of basal epidermal keratinocytes and hair follicle−derived keratinocytes, respectively. Mascharak et al (2022) generated scRNA-seq data from multiple time points across SW healing to examine the effect of inhibiting mechanotransduction through YAP on scarring. Our work evaluates single-cell fibroblast heterogeneity across numerous consecutive time points of the same fibrotic process of wound-triggered skin scarring and profiles their transcriptional dynamics in comparison with those of other major wound-associated skin cell lineages across the same timeline.

Of all cell lineages, fibroblasts were the most dynamic across wound healing, both in their transcriptional state and in their gene regulatory state (extended discussion is provided in Supplementary Text S5). This may be caused by dynamic signaling interactions between immune cells and fibroblasts, where many pathways are reused over wound healing time (Figure 6 and Supplementary Figure S13; extended discussion provided in Supplementary Text S6). We inferred that fibroblast state heterogeneity in wounds is driven by several factors: the anatomical space (upper vs lower wound domain); cell cycle status; and, critically, biological time along the healing program. Importantly, fibroblasts assume transient transcriptional states during wound healing—different fibroblast states emerge as the dominant state, only to be rapidly superseded by other states as the healing program progresses. In situ expression validated our bioinformatic identification of an early wound fibroblast state, FIB-IV, which was characterized by expression of *Prg4* and a number of chemokines and inflammatory response genes. A study by Krawetz et al (2022) showed that *Prg4* accelerates wound closure and promotes regeneration in ear cartilage wounds in mice. In a scRNA-seq study on human diabetic foot ulcers, a *PRG4*$^{+}$ fibroblast state expressing chemokines and inflammatory response genes was also identified (Theocharidis et al, 2022). We note that FIB-IV fibroblasts share similarities with CD201^{+} fascia progenitors, described by Correa-Gallegos et al (2023), which can transition to proinflammatory fibroblasts in early wounds. Thus, FIB-IV−like proinflammatory fibroblasts may be important for repair in a broad range of wounds in both mice and humans. Furthermore, reanalysis of published data by Mascharak et al (2022) suggested that FIB-IV cells emerge in mouse wounds as early as PWD2. Only

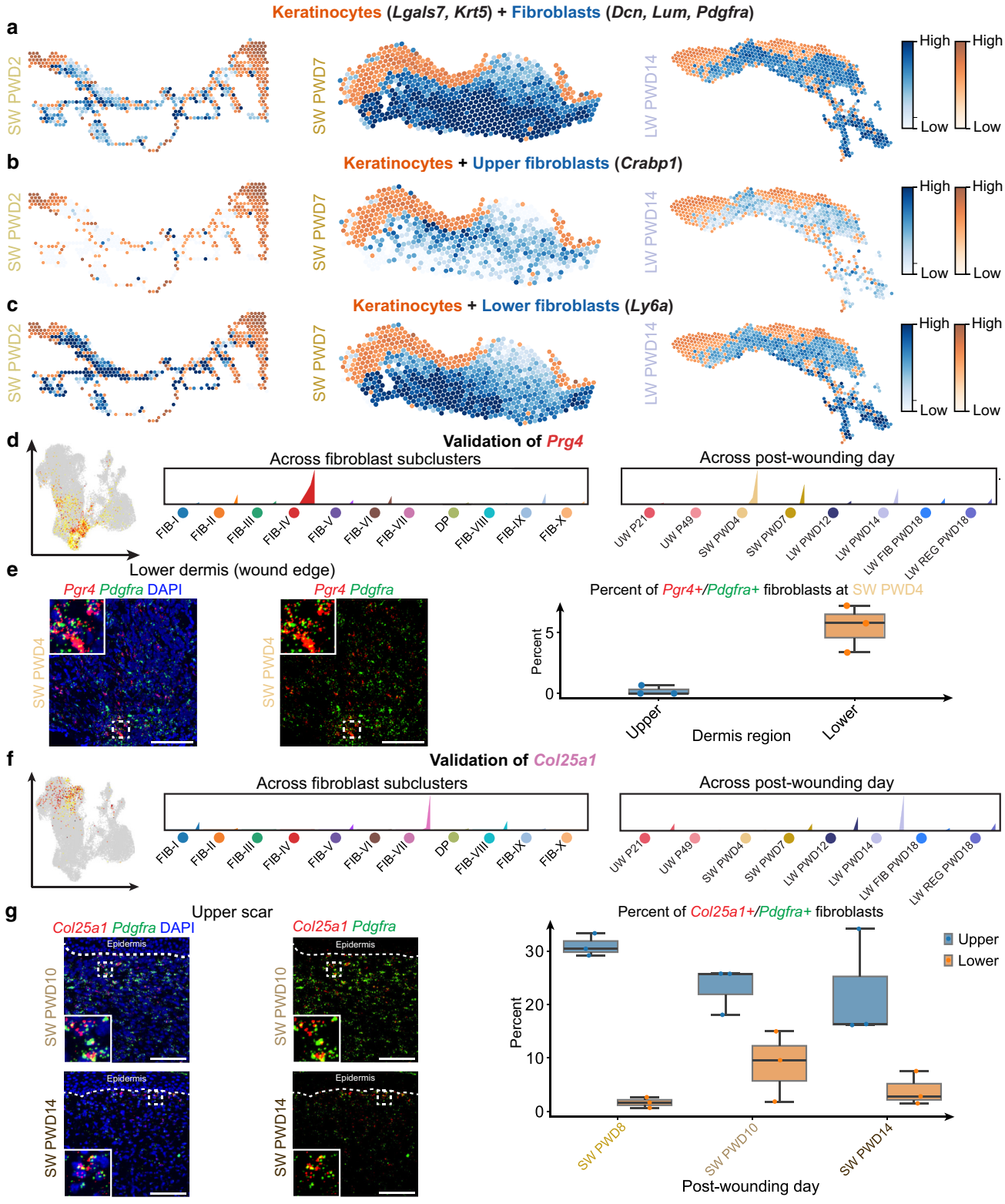


Figure 3. Spatial transcriptomic and in situ expression analyses validate spatial–temporal dynamics of wound fibroblasts. (**a**–**c**) Reanalysis of keratinocyte and fibroblast distributions in published spatial transcriptomic datasets (10X Visium platform) generated on 6-mm mouse skin wounds, sampled at 3 consecutive time points: PWD2, 7, and 14. Distributions of the keratinocyte score (orange), calculated from *Lgals7* and *Krt5* marker genes, and distribution of general fibroblast score (blue), calculated from *Dcn*, *Lum*, and *Pdgfra* genes, are shown in **a**. Distributions of keratinocyte score (orange) and upper fibroblast marker gene *Crabp1* (blue) are shown in **b**, whereas that of keratinocyte score (orange) and lower fibroblast marker gene *Ly6a* (blue) are shown in **c**. Strength of gene or gene score expression in **a**–**c** is color coded according to color heatmaps shown on the right. Red color denotes high expression levels. (**d**) Left panel shows feature plot of *Prg4*, a DEG for fibroblast subcluster FIB-IV. Orange indicates expressing fibroblasts. Center panel shows track plots of *Prg4* expression across fibroblast subclusters, annotated and color coded according to Figure 2a, whereas right panel shows track plots of *Prg4* expression across biological time points, annotated and color coded according to Figure 1b. (**e**) In situ expression of *Pdgfra* (green) and FIB-IV–specific marker *Prg4* (red) in 6-mm wound sampled

after construction of a joint transcriptomic timeline could one begin to appreciate the dynamic nature of fibroblasts—other wound-resident cell lineages do not parallel these rapid dynamics. Importantly, within the observed timeline, we show that late wound-resident fibroblasts are transcriptionally distinct from unwounded skin fibroblasts, suggesting that fibroblast within scars may be permanently altered.

Pseudotime analysis did not identify a consistent lineage model between fibroblast states. There are 2 reasons for why trajectory inference methods may have been only partially effective in their application to fibroblasts. First, current trajectory inference methods assume the same common timescale of dynamics over hours, which does not hold for wound healing that occurs over several days and weeks. Second, unlike some other cell types, fibroblast differentiation is marked by an expulsion of differentiation product, the ECM factors and modifying enzymes, into the extracellular space (Plikus et al, 2021). This likely enables fibroblasts to transition back and forth between proliferative and differentiated ECM-making states, especially during wound healing. In contrast, keratinocytes accumulate their differentiation product, such as keratin filaments, within their cytoplasm, and this specialization becomes incompatible with the ability to divide. Although genetic lineage-tracing studies will be necessary to unequivocally resolve sources of wound-resident fibroblasts, our bioinformatic analysis lays foundation for such future works and predicts that (i) there can be a common progenitor population for most wound fibroblasts, transiently emerging around the wound margin (in skin proper and deeper fascia); (ii) later-born wound fibroblast populations quickly occupy distinct upper and lower layers of the wound, after which they likely do not interconvert; and (iii) across time of healing, there can be several successive waves of wound fibroblasts, when one population supersedes another through proliferative expansion. Successive waves of wound fibroblast proliferation have been observed in an earlier study by Driskell et al (2013), including relatively late proliferation by upper wound fibroblasts during re-epithelialization.

There are several limitations to this largely bioinformatic inference-based work. First, we integrated scRNA-seq datasets on skin from mice of different biological ages, both postnatal day 21 juveniles and postnatal day 49 young adults. Although postnatal day 21 and postnatal day 49 mice are commonly used for wound healing studies (Ito et al, 2007; Plikus et al, 2017) and skin at both ages bears great morphological resemblance, including telogen hair follicles, we revealed significant age-dependent transcriptional differences in several cell lineages. Unlike keratinocytes or pericytes, dermal fibroblasts and vascular endothelial cells showed significant cell cluster composition differences between postnatal day 21 and postnatal day 49 unwounded skin samples. This suggests that (i) distinct cutaneous cell lineages undergo asynchronous postnatal maturation and that (ii) the initial time point upon which wounding is induced may play a more significant role in scRNA-seq data interpretation than previously thought. However, our staining data collected on 6-mm wound model suggest that fibroblast dynamics predicted by our analysis are largely conserved across wounds of different sizes. Second, specific to dermal fibroblasts, we showed that genes encoding matrisome and secretome factors alone can explain a large part of their transcriptional heterogeneity (Figure 4a–d). However, biological significance of numerous dynamically changing specific ECM (Figure 4h) and signaling factors (Figure 4i) remains largely unexplored. For example, we do not know the function played by a membrane-tethered collagen COL25A1, which becomes specifically expressed by upper wound fibroblasts (Figure 3g and h). In developing skeletal muscle, myotube-produced COL25A1 is essential in guiding adhesion-dependent establishment of neurointramuscular junctions (Munezane et al, 2019). Conceivably, in a similar manner, late wound fibroblast–produced COL25A1 may aid in scar tissue reinnervation. Future works will be required to test the function of numerous ECM and signaling genes, whose dynamic expression patterns in wound fibroblasts are comprehensively cataloged in this study (Figure 4h and i). Third, transcriptomic data, captured by scRNA-seq technique, offer just one layer of information on cells. However, it is insufficient to fully characterize them. Additional assays, such as protein staining, will be necessary to validate predicted wound fibroblast states, and lineage tracing will be pivotal to confirm fibroblast genealogy in wounds. In a similar vein, multiomic experiments will be needed to validate our predictions about immune–fibroblast signaling crosstalk dynamics. Along these lines, recent single-cell Assay for Transposase-Accessible Chromatin sequencing profiling of postnatal day 0 neonatal mouse skin fibroblasts revealed that numerous genes, shown as strongly differentially expressed on scRNA-seq, often lack corresponding chromatin accessibility differences (Thompson et al, 2022). At the same time, when compared between postnatal days 0 and 4, mouse skin fibroblasts prominently evolve in their chromatin accessibility states (Kim et al, 2022). These emerging data suggest that (i) epigenetically similar cells in a given tissue can diverge toward significantly different transcriptional states, likely as the result of postepigenetic regulation, and that (ii) cells' chromatin state significantly changes across biological timescales, such as upon postnatal

at PWD4 at the wound edge. Left panel shows tissue counterstained with DAPI (blue). Right plot shows the percentage of *Prg4*$^{+}$/*Pdgfra*$^{+}$ fibroblasts in upper and lower dermis at PWD4. White solid boxes show the magnified versions of the regions marked by the white dashed boxes. (**f**) Left panel shows feature plot of *Col25a1*, a DEG for mid-late upper wound–specific fibroblasts. Orange indicates expressing fibroblasts. Center panel shows the track plots of *Col25a1* expression across fibroblast subclusters, annotated and color coded according to Figure 2a, whereas the right panel shows the track plots of *Col25a1* expression across biological time points, annotated and color coded according to Figure 1b. (**g**) In situ expression of *Pdgfra* (green) and upper wound fibroblast–specific marker *Col25a1* (red) in 6-mm wound sampled at PWD10 (top panels) and PWD14 (bottom panels) in the upper scar region. The left columns show tissues counterstained with DAPI (blue). White dashed lines mark wound epidermis. White solid boxes show the magnified version of the regions marked by the white dashed boxes. Right plot shows the percentage of *Col25a1*$^{+}$/*Pdgfra*$^{+}$ fibroblasts in upper and lower scar regions at PWD8, 10, and 14. Bars in **e** and **g** = 100 μm. UW denotes unwounded skin. DEG, differentially expressed gene; DP, dermal papilla; LW, large wound; PWD, postwounding day; SW, small wound.

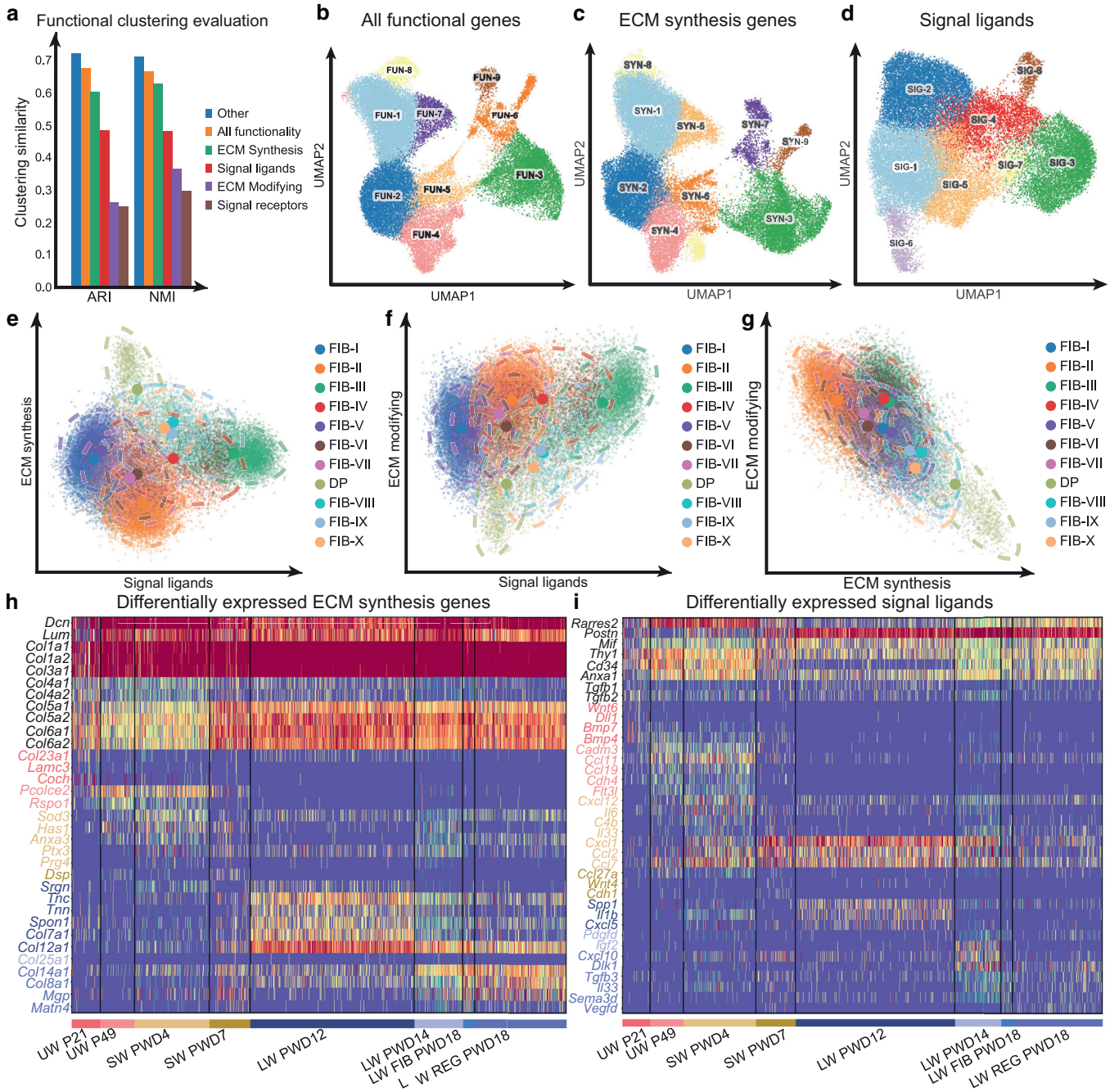


Figure 4. Functional exploration reveals key driver gene categories of fibroblast heterogeneity. (**a**) ARI (left) and NMI (right) for assessing the similarity between different functional cluster annotations are shown. Higher values along y-axis indicate higher similarities. (**b**–**d**) UMAP plots revealing the neighborhood embeddings of the fibroblast populations when integrating only on (**b**) all functionally annotated genes, (**c**) ECM synthesis genes only, or (**d**) signal ligand genes only. In **b**–**d**, fibroblast cluster labels are overlayed onto UMAPs. (**e**–**g**) Scatter plot representation of the fibroblast population, after training on an interpretable autoencoder where the latent dimensions correspond to the functional gene sets, denoted along x- and y-axes. Dashed ellipses correspond to 95% confidence areas. (**h**) Shared and biological time point–specific ECM synthesis genes inferred by differential expression tests. (**i**) Shared and biological time point–specific signal ligand genes inferred by differential expression tests. ARI, Adjusted Rand Index; ECM, extracellular matrix; NMI, normalized mutual information; UMAP, Uniform Manifold Approximation and Projection.

maturation. Future application of single-cell Assay for Transposase-Accessible Chromatin sequencing and other emerging single-cell multiomic profiling techniques to wounded skin will better access the degree to which distinct cell lineages reprogram or resist reprogramming in face of injury-triggered tissue repair or regeneration programs.

MATERIALS AND METHODS

Dataset curation

We integrated 8 previously published scRNA-seq datasets on full-thickness excisional wounds sampled from different wound healing times and from both SWs and LWs (details are provided in Datasets in Supplementary Materials and Methods).

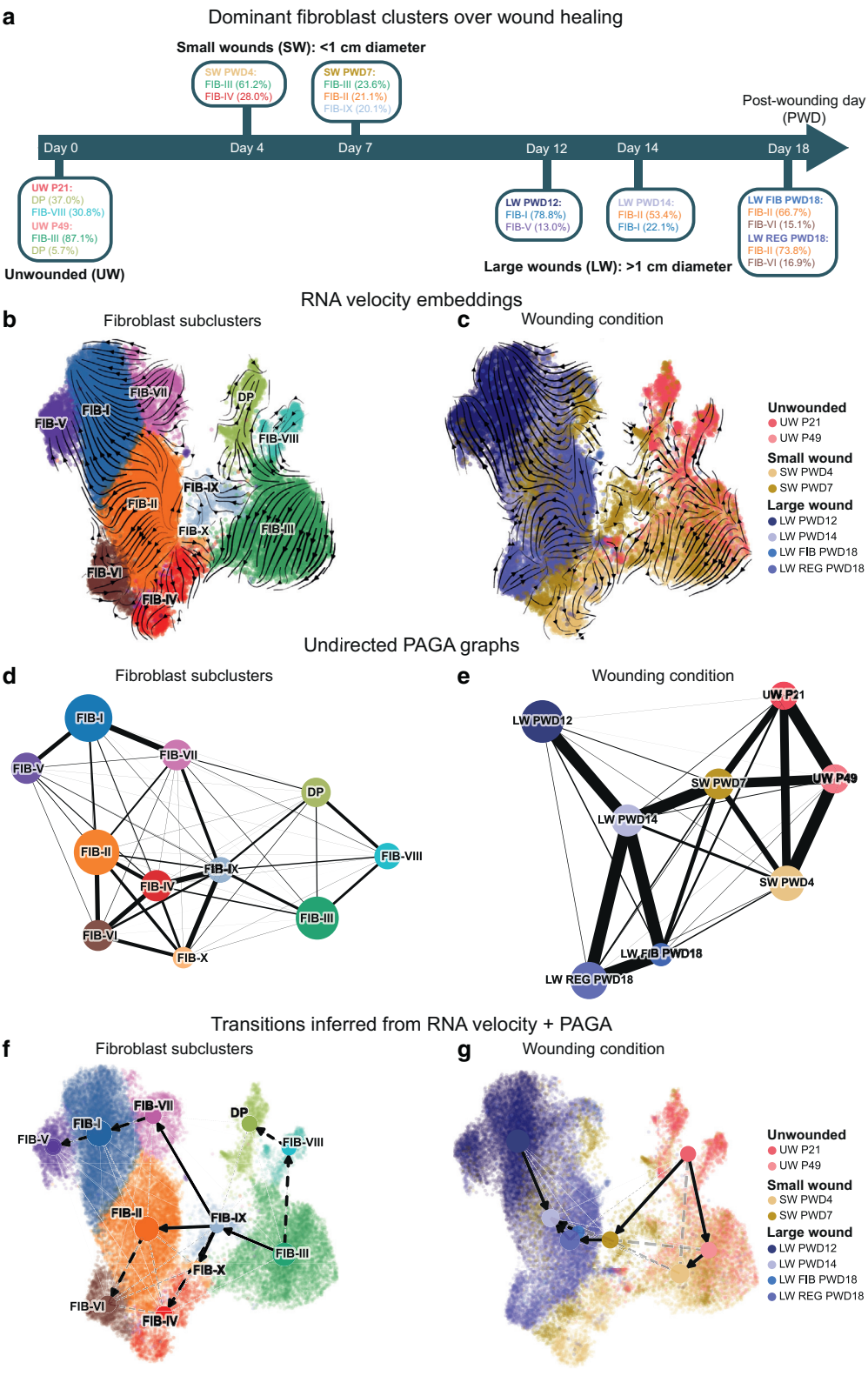


Figure 5. Joint trajectory analysis on unwounded skin and wound fibroblasts reveals their potential lineage relationship. (**a**) Summary of the dominant fibroblast subclusters at each time point along the biological timescale of skin wound healing, as originally defined in Figure 1a. (**b**–**c**) RNA velocity streamline plots of fibroblasts, where cells are annotated and color coded according to subclusters, originally defined in Figure 2a (in panel **b**) or, alternatively, according to biological time points, originally defined in Figure 1b (in panel **c**). (**d, e**) PAGA graphs constructed on fibroblast subclusters, originally defined in Figure 2a (in panel **d**), or on biological time points, originally defined in Figure 1b (in panel **e**). (**f, g**) Directed PAGA graphs with the most likely transitions between fibroblast subclusters, color coded and annotated according to Figure 2a (in panel **f**), or biological time points defined in Figure 1b (in panel **g**). Solid arrows in **f** and **g** denote most likely lineage relationships between fibroblasts, whereas dashed arrows represent transitions, deemed as conflicting with prior knowledge on normal skin development and wound healing. UW denotes unwounded skin. DP, dermal papilla; LW, large wound; PWD, postwounding day; SW, small wound.

Data integration

We preprocessed all curated scRNA-seq datasets using Scanpy (details are provided in Data preprocessing in Supplementary Materials and Methods). We integrated the data across different wound healing time points using Scanorama (Hie et al, 2019) and assessed the quality of integration using scIB (Luecken et al, 2022) (details are provided in Data integration in Supplementary Materials and Methods).

Data analysis

We performed analysis of functional fibroblast genes using Intercode (Rybakov et al, 2020[1]) and MAST (Finak et al, 2015) (details are

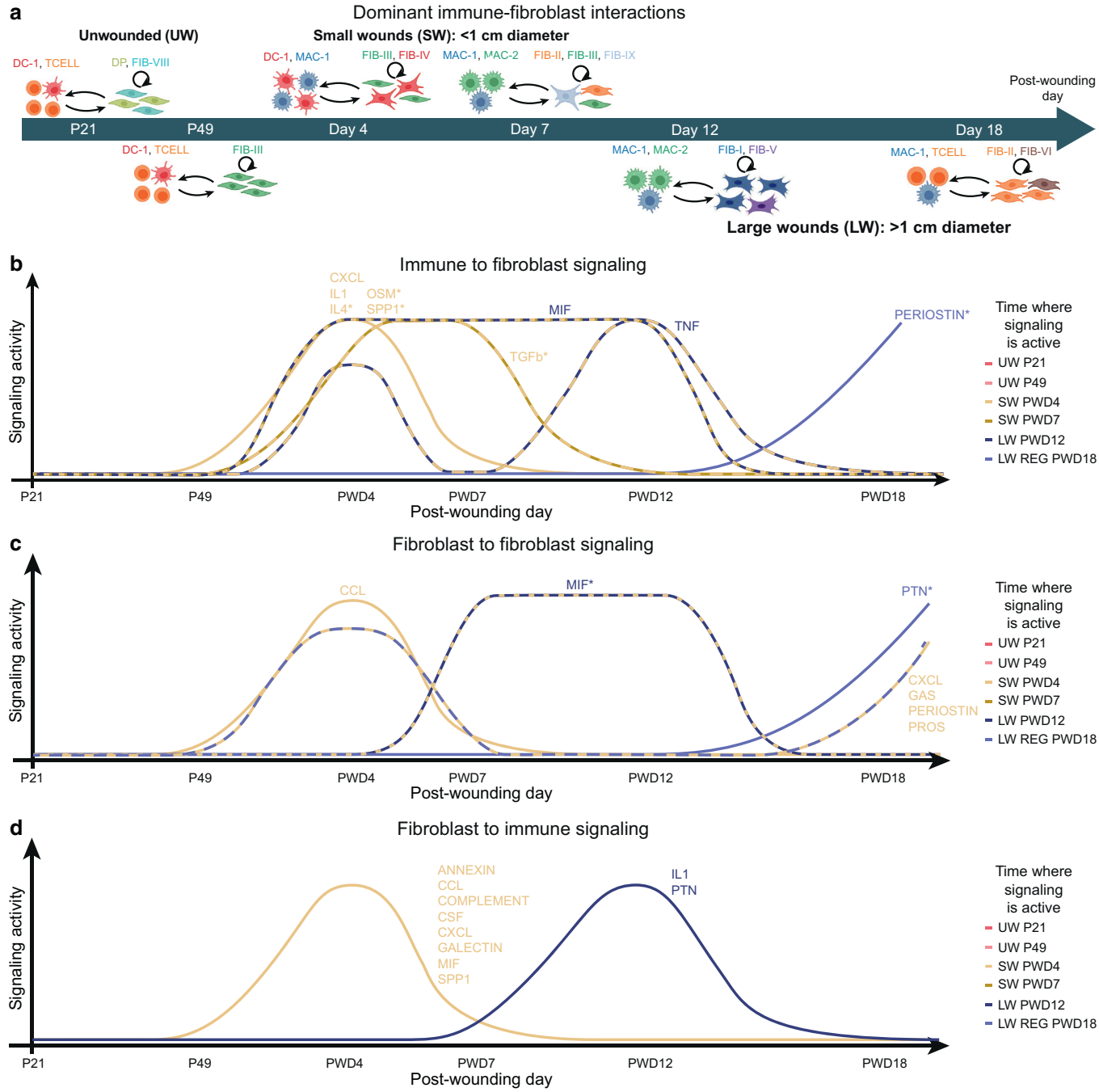


Figure 6. Wound healing–associated fibroblast–immune signaling communications predicted by CellChat analysis. (**a**) Schematic representation of dominant immune cell types and fibroblast states at each time point and their possible interactions. Summary of (**b**) paracrine signaling pathways with ligands secreted from immune cells to fibroblasts, (**c**) autocrine signaling pathways with ligands secreted from fibroblasts to other fibroblasts, and (**d**) paracrine signaling pathways with ligands secreted from fibroblasts to immune cells. In **b**–**d**, pathways marked with an asterisk were predicted to be likely mitogenic. Extended discussion is provided in Supplementary Text S6 and associated Supplementary Figure S13. UW denotes unwounded skin. LW, large wound; P21, postnatal day 21; P49, postnatal day 49; PWD, postwounding day; SW, small wound.

provided in Biased clustering in Supplementary Materials and Methods). Trajectory inference was performed using scVelo (Bergen et al, 2020). We used kallisto|bustools (Melsted et al, 2019) to generate the unspliced and spliced mRNA count matrices that were used for RNA velocity analysis (details are provided in Trajectory analysis in Supplementary Materials and Methods). Regulon inference was performed using pySCENIC (Van de Sande et al, 2020). We inferred significant cell–cell communication activities using CellChat (Jin et al, 2021) (details are provided in Cell–cell communication analysis in Supplementary Materials and Methods).

Wound replication studies

We performed 2 series of wounding experiments. First, we recapitulated the wounding experiments used to generate each curated

[3] Rybakov S, Lotfollahi M, Theis FJ, Wolf FA. Learning interpretable latent autoencoder representations with annotations of feature sets. bioRxiv 2020:2–6.

dataset in C57BL/6J mice, where mice were matched for age and wound size (details are provided in Replication wounding studies in Supplementary Materials and Methods). Second, we generated 6-mm excisional wounds in postnatal day 49 C57BL/6J mice, studying them on PWD4, 8, 10, and 14. The wounding procedures were approved by the Institutional Animal Care and Use Committee at the University of California, Irvine (Irvine, CA).

Spatial gene expression profiling

We generated in situ expression data using RNAscope with the following mouse probes: *Lum* (480361), *Prg4* (437661-C2), *Mki67* (416771), *Pdgfr-C3* (480661-C3), *Pamr1* (491431-C2), and *Col25a1* (538511-C2), whereas DAPI was used for nuclei counterstaining (details are provided in Replication wounding studies in Supplementary Materials and Methods).

Full information about the methods used can be found in the Supplementary Materials and Methods.

DATA AVAILABILITY STATEMENT

All original data used in this study can be accessed from National Center for Biotechnology Information Gene Expression Omnibus Repository at the following accession numbers: GSE153596, GSE142721, GSE113854, GSE108677, and GSE141814. The integrated datasets—one containing all cell lineages and the other containing the subclustered fibroblast population—with cell-type annotation are available at the following Zenodo repository: https://doi.org/10.5281/zenodo.10198892. We have also made 2 web resources to query the 2 integrated datasets at https://skingenes.biochem.uci.edu/wound-healing/ and https://skingenes.biochem.uci.edu/wound-fibroblasts/.

CODE AVAILABILITY STATEMENT

All codes used to generate figures and analysis have been made available at the following GitHub repository: https://github.com/axelalmet/FibroblastAnalysis_2023.

ORCIDs

Axel A. Almet: http://orcid.org/0000-0001-9173-8278
Yingzi Liu: http://orcid.org/0000-0002-8704-4756
Qing Nie: http://orcid.org/0000-0002-8804-3368
Maksim V. Plikus: http://orcid.org/0000-0002-8845-2559

CONFLICT OF INTEREST

The authors state no conflict of interest.

ACKNOWLEDGMENTS

This work was supported by a National Science Foundation grant DMS1763272, National Institutes of Health grants U01-AR073159 and R01-AR079150, Chan Zuckerberg Initiative grant AN-0000000062, and a grant from the Simons Foundation (594598). MVP is also supported by LEO Foundation grants LF-AW-RAM-19-400008 and LF-OC-20-000611, National Institutes of Health grant R01-AR079470, W. M. Keck Foundation grant WMKF-5634988, and Horizon Europe grant 101137006. YL is supported by the clinical fellowship from the California Institute for Regenerative Medicine training grant (EDUC4-12822). We would like to thank Driskell for sharing a missing dataset from National Center for Biotechnology Information GEO Datasets (GSE153596) and Januszyk and Longaker for sharing the data from their original study. MVP is the guarantor of this work.

AUTHOR CONTRIBUTIONS

Conceptualization: AAA, QN, MVP; Data Curation: AAA, QN, MVP; Formal Analysis: AAA, QN, MVP; Funding Acquisition: QN, MVP; Investigation: AAA, YL; Methodology: AAA, YL, QN, MVP; Project Administration: QN, MVP; Resources: QN, MVP; Software: AAA; Supervision: QN, MVP; Validation: YL; Visualization: AAA, YL; Writing - Original Draft Preparation: AAA, QN, MVP; Writing - Review and Editing: AAA, YL, QN, MVP

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at www.jidonline.org, and at 10.1016/j.jid.2024.06.1281.

REFERENCES

Abbasi S, Sinha S, Labit E, Rosin NL, Yoon G, Rahmani W, et al. Distinct regulatory programs control the latent regenerative potential of dermal fibroblasts during wound healing. Cell Stem Cell 2020;3:396–412.e6 [published correction appears in Cell Stem Cell 2021;28:581–3].

Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol 2020;38:1408–14.

Buechler MB, Pradhan RN, Krishnamurty AT, Cox C, Calviello AK, Wang AW, et al. Cross-tissue organization of the fibroblast lineage. Nature 2021;593:575–9.

Correa-Gallegos D, Ye H, Dasgupta B, Sardogan A, Kadri S, Kandi R, et al. CD201+ fascia progenitors choreograph injury repair. Nature 2023;623:792–802.

Darby IA, Laverdet B, Bonté F, Desmoulière A. Fibroblasts and myofibroblasts in wound healing. Clin Cosmet Investig Dermatol 2014;7:301–11.

Driskell RR, Lichtenberger BM, Hoste E, Kretzschmar K, Simons BD, Charalambous M, et al. Distinct fibroblast lineages determine dermal architecture in skin development and repair. Nature 2013;504:277–81.

Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol 2015;16:278.

Forte E, Ramialison M, Nim HT, Mara M, Li JY, Cohn R, et al. Adult mouse fibroblasts retain organ-specific transcriptomic identity. eLife 2022;11:1–26.

Foster DS, Januszyk M, Chinta MS, Yost KE, Gulati GS, Nguyen AT, et al. Integrated spatial multiomics reveals fibroblast fate during tissue repair. Proc Natl Acad Sci U S A 2021;118:e2110025118.

Gay D, Ghinatti G, Guerrero-Juarez CF, Ferrer RA, Ferri F, Lim CH, et al. Phagocytosis of Wnt inhibitor SFRP4 by late wound macrophages drives chronic Wnt activity for fibrotic skin healing. Sci Adv 2020;6:eaay3704.

Guerrero-Juarez CF, Dedhia PH, Jin S, Ruiz-Vega R, Ma D, Liu Y, et al. Single-cell analysis reveals fibroblast heterogeneity and myeloid-derived adipocyte progenitors in murine skin wounds. Nat Commun 2019;10:650.

Gurtner GC, Werner S, Barrandon Y, Longaker MT. Wound repair and regeneration. Nature 2008;453:314–21.

Haensel D, Jin S, Sun P, Cinco R, Dragan M, Nguyen Q, et al. Defining epidermal basal cell states during skin homeostasis and wound healing using single-cell transcriptomics. Cell Rep 2020;30:3932–47.e6.

Hie B, Bryson B, Berger B. Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. Nat Biotechnol 2019;37:685–91.

Hu KH, Kuhn NF, Courau T, Tsui J, Samad B, Ha P, et al. Transcriptional space-time mapping identifies concerted immune and stromal cell patterns and gene programs in wound healing and cancer. Cell Stem Cell 2023;30:885–903.e10.

Ito M, Yang Z, Andl T, Cui C, Kim N, Millar SE, et al. Wnt-dependent de novo hair follicle regeneration in adult mouse skin after wounding. Nature 2007;447:316–20.

Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. Nat Commun 2021;12:1088.

Joost S, Jacob T, Sun X, Annusver K, La Manno G, Sur I, et al. Single-cell transcriptomics of traced epidermal and hair follicle stem cells reveals rapid adaptations during wound healing. Cell Rep 2018;25:585–97.e7.

Kim JY, Park M, Ohn J, Seong RH, Chung JH, Kim KH, et al. Twist2-driven chromatin remodeling governs the postnatal maturation of dermal fibroblasts. Cell Rep 2022;39:110821.

Konieczny P, Xing Y, Sidhu I, Subudhi I, Mansfield KP, Hsieh B, et al. Interleukin-17 governs hypoxic adaptation of injured epithelium. Science 2022;377:eabg9302.

Krawetz RJ, Abubacker S, Leonard C, Masson AO, Shah S, Narendran N, et al. Proteoglycan 4 (PRG4) treatment enhances wound closure and tissue regeneration. NPJ Regen Med 2022;7:32.

Luecken MD, Büttner M, Chaichoompu K, Danese A, Interlandi M, Mueller MF, et al. Benchmarking atlas-level data integration in single-cell genomics. Nat Methods 2022;19:41–50.

Mascharak S, Talbott HE, Januszyk M, Griffin M, Chen K, Davitt MF, et al. Multi-omic analysis reveals divergent molecular events in scarring and regenerative wound healing. Cell Stem Cell 2022;29:315–27.e6.

Melsted P, Booeshaghi AS, Liu L, et al. Modular, efficient and constant-memory single-cell RNA-seq preprocessing. Nat Biotechnol 2021;39:813–8.

Munezane H, Oizumi H, Wakabayashi T, Nishio S, Hirasawa T, Sato T, et al. Roles of collagen XXV and its putative receptors PTPσ/δ in intramuscular motor innervation and congenital cranial dysinnervation disorder. Cell Rep 2019;29:4362–76.e6.

Phan QM, Fine GM, Salz L, Herrera GG, Wildman B, Driskell IM, et al. Lef1 expression in fibroblasts maintains developmental potential in adult skin to regenerate wounds. eLife 2020;9:e60066.

Phan QM, Sinha S, Biernaskie J, Driskell RR. Single-cell transcriptomic analysis of small and large wounds reveals the distinct spatial organization of regenerative fibroblasts. Exp Dermatol 2021;30:92–101.

Philippeos C, Telerman SB, Oulès B, Pisco AO, Shaw TJ, Elgueta R, et al. Spatial and single-cell transcriptional profiling identifies functionally distinct human dermal fibroblast subpopulations. J Invest Dermatol 2018;138:811–25.

Plikus MV, Guerrero-Juarez CF, Ito M, Li YR, Dedhia PH, Zheng Y, et al. Regeneration of fat cells from myofibroblasts during wound healing. Science 2017;355:748–52.

Plikus MV, Wang X, Sinha S, Forte E, Thompson SM, Herzog EL, et al. Fibroblasts: origins, definitions, and functions in health and disease. Cell 2021;184:3852–72.

Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 2015;33:495–502.

Shook BA, Wasko RR, Mano O, Rutenberg-Schoenberg M, Rudolph MC, Zirak B, et al. Dermal adipocyte lipolysis and myofibroblast conversion are required for efficient skin repair. Cell Stem Cell 2020;26:880–95.e6.

Solé-Boldo L, Raddatz G, Schütz S, Mallm JP, Rippe K, Lonsdorf AS, et al. Single-cell transcriptomes of the human skin reveal age-related loss of fibroblast priming. Commun Biol 2020;3:188.

Theocharidis G, Thomas BE, Sarkar D, Mumme HL, Pilcher WJR, Dwivedi B, et al. Single cell transcriptomic landscape of diabetic foot ulcers. Nat Commun 2022;13:181.

Thompson SM, Phan QM, Winuthayanon S, Driskell IM, Driskell RR. Parallel single-cell multiomics analysis of neonatal skin reveals the transitional fibroblast states that restrict differentiation into distinct fates. J Invest Dermatol 2022;142:1812–23.e3.

Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. Sci Rep 2019;2:5233.

Van de Sande B, Flerin C, Davie K, De Waegeneer M, Hulselmans G, Aibar S, et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nat Protoc 2020;15:2247–76.

Vu R, Jin S, Sun P, Haensel D, Nguyen QH, Dragan M, et al. Wound healing in aged skin exhibits systems-level alterations in cellular composition and cell-cell communication. Cell Rep 2022;40:111155.

Wolf FA, Hamey FK, Plass M, Solana J, Dahlin JS, Göttgens B, et al. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. Genome Biol 2019;20:59.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/

SUPPLEMENTARY TEXT

Supplementary Text S1: Value of data integration to understand wound healing and computational challenges

Because more and more single-cell RNA-sequencing (scRNA-seq) studies of the same biological system subject to various conditions and perturbations have emerged, it has become of great interest to construct joint single-cell maps using data integration methods. By integrating multiple datasets across different conditions, the philosophy is that the whole is greater than the sum of its parts, allowing one to learn transcriptomic characteristics that are either shared across conditions or that are condition specific. However, a number of biological and technical factors can influence gene expression profiles, rendering scRNA-seq data integration a difficult task (Luecken et al, 2022). Biological factors can include animal species model used, which can vary by age, sex, gender, genotype, and housing environment; the wounding model, which can include wound size (small vs large), wound depth, wound geometry, and physical location on the body; the sample collection protocol, which can include or exclude edge skin along with the wound proper; the dissection protocol; the dissociation protocol of the sample for sequencing; and many more. Technical factors can include the sequencing technology and the resultant sequencing depth, which can affect the number of reads per cell capture. Furthermore, gene expression data are subject to intrinsic noise, and separating biological noise from technical noise is still a source of great debate (Sarkar and Stephens, 2021). All these factors make it difficult to determine how cells from different datasets should be placed on an integrated transcriptomic landscape—which cells are the same or similar cell type from different conditions, and which are unique to specific conditions? The difficulty of the data integration task is further exacerbated by lack of consensus best method. Rather, the best method depends on the datasets and biological system in question (Luecken et al., 2022). Skin itself presents a unique challenge because it contains a remarkable diversity of cell lineages across several compartments: epidermis, dermis, ectodermal appendages, and dermal adipose tissue. Some cell types are present only under certain biological conditions (before wound vs after wound), whereas others undergo remarkable changes in their transcriptional state across wound healing. Therefore, when constructing an integrated single-cell map, all of the factors mentioned earlier need to be considered.

Supplementary Text S2: Each wound-resident cell lineage demonstrates a distinct temporal pattern of cell state dynamics

We examined whether major skin cell lineages undergo compositional changes over wound healing time. We subclustered epidermal cells (Supplementary Figure S1), pericytes, and endothelial cells, which are the dominant vascular cell lineages present (Supplementary Figure S2); skin-resident immune cells (Supplementary Figure S3); and Schwann cells (Supplementary Figure S4) separately and examined them in a manner similar to that of fibroblasts. We did not analyze epidermal cells in detail because they were not present across all wound conditions (Supplementary Figure S1c and d), a likely technical artefact stemming from differences in cell isolation protocols between studies.

Further analysis of pericytes revealed 5 distinct subclusters (Supplementary Figure S2a–d). Subcluster PERI-3 significantly expresses myeloid markers such as *Fcer1g* and *Tyrobp* (Supplementary Figure S2b). PERI-3 is primarily present in postwounding day (PWD)7 and PWD12 skin (Supplementary Figure S2c), suggesting that this subpopulation of pericytes aids in wound repair. In previous works, myeloid-derived pericytes have been observed in developing skin (Yamazaki et al, 2017), although we did not observe PERI-3 in juvenile skin at postnatal day 21. We also observed a subcluster of proliferative pericytes, PERI-4 (Supplementary Figure S2b), which only appeared in PWD4, PWD12, and PWD14, suggesting that pericyte proliferation is activated during wound repair. This has been observed in studies of fibrosis in other organs, such as lung or kidney (Birbrair et al, 2014). Compositional changes in pericytes revealed dynamics that partially overlapped with fibroblast dynamics (Supplementary Figure S2d). The dominant pericyte subcluster in unwounded skin, PERI-5, was rapidly reduced in proportion after wounding and was no longer present at late wound times. On the other hand, PERI-1 cells transiently increased to become the dominant cluster in early wounds but then disappeared in late wounds, whereas the latter were dominated by PERI-2 cells that were effectively absent from unwounded skin. Taken together, we conclude that pericytes partially parallel fibroblast dynamics and do not return to their unwounded state within a timeframe represented in our analysis.

A natural question is why pericytes mirror fibroblast dynamics in wound healing when they serve different functions—fibroblasts deposit extracellular matrix (ECM) to seal the wound, whereas pericytes aid in angiogenesis and maintenance of vasculature. There are some clues offered in Goss et al (2021), who studied the developmental origins of pericytes. They found that fibroblasts could give rise to pericytes in the skin, and both upper and lower fibroblasts gave rise to pericytes, proposing four distinct lineages according to expressions of *Ng2*, *Dlk*, and *Lrig*, on the basis of fibroblast origins. In unwounded skin, pericytes could equally arise from both fibroblast lineages, but in wounded skin, one lineage remained restricted to the lower dermis. If this is true, then the dominant pericyte state would also be associated with the dominant fibroblast state. Therefore, given that we have observed that late wound fibroblasts are dominated by lower fibroblasts, the dominant observed pericyte subcluster in late wounds would indeed be distinct from unwounded or early wound skin. Furthermore, it has been established that pericytes have different functions at different stages of wound healing (Bodnar et al, 2016). In early wounds, pericytes recruit endothelial cells through the expression of chemokine CXCL signaling to promote angiogenesis. In later states of wound healing, the role of pericytes switches more to remodeling vasculature. In some ways, this does mirror fibroblasts. However, although pericytes parallel fibroblasts dynamics partially, Sun et al (2022)[2] showed that GLI1$^+$ pericytes are not the sources of new GLI1$^+$ fibroblasts entering the wound area during healing, suggesting that pericyte dynamics occur in parallel with fibroblast dynamics.

In contrast to fibroblasts and pericytes, endothelial cells did not show dynamic changes during wound healing (Supplementary Figure S2e–h). Rather, the dominant driver of their state appeared to be animal age. For all conditions except postnatal day 21 and PWD7 (postnatal day 21 + 7 days), the mouse model was aged between postnatal day 42 to postnatal day 63; here, ENDO-1 was the dominant subcluster. On the other hand, in juvenile unwounded (postnatal day 21) and juvenile wounded (PWD7) skin, the dominant subcluster was ENDO-2, which was virtually absent from all other times. We note that there is a small ENDO-4 subpopulation of cells with myeloid gene expression features, which is only present in wounded PWD4 to PWD12 skin (Supplementary Figure S2g), mirroring the appearance of fibroblasts and pericytes with myeloid gene characteristics. Therefore, we conclude that endothelial cells prominently change upon postnatal mouse skin maturation (between postnatal day 21 and postnatal day 42) but remain stable in response to wounding.

Analysis of immune cell populations revealed distinct dynamics from those seen in fibroblasts and pericytes (Supplementary Figure S3). After subclustering, we identified 14 immune subclusters (Supplementary Figure S3a–d), including three clusters of dendritic cells (DC1–4), five macrophage clusters (MAC1–5), one T-cell cluster, one natural killer cell subcluster, one basophil subcluster, one neutrophil subcluster, one plasmacytoid dendritic cell subcluster, and one mast cell subcluster. Further subclustering revealed four key subpopulations of T cells (Supplementary Figure S3e–h), including regulatory T cells (Tregs), natural killer T cells, proliferating T cells, T-cell–presenting innate lymphoid cells, and a small subset of activated T cells marked by differential expression of *Spp1* (osteopontin). Of these T-cell subpopulations, the Treg and natural killer T cell subclusters were the only subclusters present across all wounding conditions. Natural killer T ceskinlls remained relatively unchanged in proportion across wound healing time, whereas Tregs increased over time in proportion. From juvenile postnatal day 21 to postnatal day 49 skin, there is a relative increase in Tregs. In wounded skin, Tregs represented the largest T-cell subtype, particularly during PWD12 skin time point. Innate lymphoid cells were present in unwounded skin and late wound tissue and decreased in proportion as unwounded skin matured (from postnatal day 21 to postnatal day 49). In unwounded postnatal day 49 skin, the actively proliferating proliferating T cell subcluster was virtually absent, but in wounded PWD7 to PWD14 skin, this population expanded, before decreasing in proportion by late wound time again.

Finally, we subclustered the Schwann cell population (Supplementary Figure S4), which revealed significant changes in unwounded skin and early wounds. From postnatal day 21 to postnatal day 49, there was an expansion of the subcluster SCH-4 but a depletion of SCH-1 and SCH-2 cells. Upon wounding, SCH-3 emerged as the dominant Schwann cell state, and SCH-2 was absent from all wounds. From PWD12 onward, the dominant subcluster was SCH-1, with no significant change in Schwann cell composition in later wounds.

Thus, our combined analysis on integrated data reveals the following, to our knowledge, previously unreported insight into skin cells. First, both fibroblasts and vascular cell types undergo prominent age-dependent changes that take place after age of 3 weeks. Such changes are cell type specific because pericytes appear to already be transcriptionally mature in postnatal day 21 animals. Second, different skin cell lineages demonstrate divergent transcriptional dynamics in response to wounding. Both fibroblasts and pericytes enter early postwounding states and then transition into late postwounding states, both of which are distinct from their prewounding homeostatic states. On the other hand, endothelial cells and T cells are not significantly transcriptionally plastic in response to wounding.

Supplementary Text S3: Analysis of genes that correlate significantly with upper and lower fibroblast identity scores

To discern whether upper and lower fibroblasts are distinguished, we performed differential expression with respect to dermal fibroblast upper and lower scores, separately (Supplementary Figure S5). From this analysis, we inferred that the top five genes most correlated with the upper score were *Il1rn*, *Ccl9*, *Lilr4b*, *Ly6c2*, and *Cytip*, whereas the top five genes least correlated with the upper score were *Adamts10*, *Adamtsl5*, *Angptl1*, *Trim12a*, and *Ccpg1* (Supplementary Figure S5a). Of the top genes, *Il1rn*, *Ccl9*, *Lilr4b*, *Ly6c2*, and *Cytip* were primarily expressed in the activated large wound (LW) PWD12 FIB-V subcluster (labeled purple in Supplementary Figure S5a).

In contrast, the top five genes most correlated with the lower score were *Sema3c*, *Adamts15*, *Sp110*, *Sfrp4*, and *Tcea3*, whereas the top five genes least correlated with the lower score were *Apela*, *Adam8*, *Wnt10a*, *Clec4d*, and *Wif1* (Supplementary Figure S5b). Of the genes most correlated with the lower score, *Sema3c*, *Adamts15*, and *Sfrp4* were primarily expressed in the PWD18 FIB-II subcluster (labeled orange in Supplementary Figure S5b), whereas *Tcea3* was expressed by both postnatal day 49 FIB-III and PWD18 FIB-II fibroblasts (labeled green in Supplementary Figure S5b), and *Sp110* was expressed by the small wound (SW) PWD4 FIB-IV subcluster (labeled red in Supplementary Figure S5b). We note that *Sfrp4* has been implicated previously in fibrotic outcomes in wound healing (Gay et al, 2020), in line with our observation that late LW PWD18 wounds are dominated by fibroblasts of primarily lower identity. We found that the top 5 genes that were most correlated with the lower score and significantly expressed in the unwounded FIB-III cluster were *Tcea3*, *Eln*, *Pcsk6*, *Sult5a1*, and *Osr1* (labeled green in Supplementary Figure S5b). All of these genes were also expressed by FIB-II fibroblasts, suggesting that the lower identity is shared across unwounded and wounded skin. Furthermore, *Osr1* has been shown to mark fibroadipogenic progenitors during muscle development (Vallecillo-García et al, 2017) and is expressed in adult fibroadipogenic progenitors activated owing to injury (Stumm et al, 2018), suggesting that expression of *Osr1* in lower fibroblasts in the skin may similarly point to their role as a reservoir for fibroblast activation in wound healing.

Supplementary Text S4: Significance of selected differentially expressed signaling ligands in fibroblasts

Some of the differentially expressed fibroblasts' signaling ligands have already been implicated in scar formation. In

juvenile unwounded skin, we observed differential expression of *Apela*. *Apela* has been shown to promote fibroblast migration (Doğan, 2019) and to inhibit TGFβ-induced fibrosis (Yokoyama et al, 2018). On the other hand, PWD12 fibroblasts significantly expressed hepatocyte GF gene *Hgf*, which downregulates collagen production in the skin (Qin et al, 2017) and helps to reduce scarring (Xiao and Xi, 2013).

Supplementary Text S5: Regulon inference reveals wound healing-associated fibroblast state regulation

We clustered fibroblasts on the basis of the regulon enrichment scores (AUCell), setting resolution = 0.6, which produced 12 gene regulation–based clusters (Supplementary Figure S12a). Calculating the Jaccard similarity between the original transcriptomic-based fibroblast subclusters and the inferred regulon clusters revealed significant overlap between the 2 sets of cluster labels (Supplementary Figure S12b). Many regulon clusters showed high overlap with a single subcluster and low overlap for all other subclusters; for example, REG-1 had high overlap with FIB-I (Jaccard similarity = 0.75), suggesting that many of the original subclusters also correspond to unique gene regulatory states. However, some regulon clusters, such as REG-4 and REG-5, had significant overlap across multiple transcriptomic-based subclusters, such as REG-4, for whom the highest overlaps occurred across FIB-III (0.11), FIB-VIII (0.17), and FIB-IX (0.14), corresponding to fibroblasts from unwounded skin and wounded SW PWD7 skin, suggesting a shared regulatory state across these fibroblast subclusters. To further examine for drivers of gene regulator states, we calculated the Jaccard similarity between regulon clusters and wound healing times (Supplementary Figure S12c). Examining the Jaccard similarities revealed that some regulon clusters that did not overlap significantly with specific fibroblast subclusters did overlap with a specific wound healing time. For example, REG-4 consisted of postnatal day 21 unwounded (0.29 overlap) and SW PWD7 wound (0.54) fibroblasts. In addition, REG-5 consisted primarily of LW PWD14 fibroblasts (0.75 overlap), which were shared across the subclusters, FIB-II (0.18) and FIB-VII (0.15). One of the smallest regulon clusters, REG-11, had a maximum overlap with LW FIB PWD18 wound (0.05) but did not significantly overlap with any fibroblast subcluster (maximum overlap of 0.02), suggesting subtleties in gene regulation in fibroblasts that are not captured by just gene transcription. We then examined how these fibroblast regulon states likely influence fibroblast functions across wound healing. For each of these functional categories, we determined the top differentially expressed genes (DEGs) where an upstream transcription factor (TF) was inferred, synthesizing this information across wound healing time with respect to the curated fibroblast functions (Supplementary Figure S12d). We observed that many of the key DEGs relating to fibroblast functions (and more) are downstream targets of significant TFs. For example, in juvenile postnatal day 21 unwounded skin, we found that *Coch*, which we showed was a significant differentially expressed ECM synthesis gene, is regulated by both *Egr3* and *Hoxd9*, whereas another significant ECM synthesis gene, *Col23a1*, is a downstream target of *Egr3* and several other TFs, including *Alx4*, *Cepba*, *Irf8*, *Klf11*, *Lef1*, *Mef2d*, *Msx*, *Tcf7l1*, *Tfap2c*, *Thrb*, and *Trps1*. In young adult postnatal day 49 unwounded skin, we observed that of the key signaling DEGs, *Ccl19* is a downstream target of *Trp63*, whereas *Cadm3* is a target of each of *Ar*, *Fosb*, *Irx2*, *Rarg*, and *Srebf1*. In early SW PWD4 wound, we observed that *Cxcl13*, a differentially expressed signaling ligand, is a downstream target of *Pparg*. In LW PWD12 wounds, the significant ECM synthesis gene *Col7a1* is a downstream target of numerous TFs, including *Bach1*, *Elk3*, *Kdm5a*, *Kdm5b*, *Maz*, *Nfatc1*, *Nfkb1*, *Smarca4*, and *Zfp281*. In late LW PWD18 wounds, the differentially expressed ECM genes *Col8a1* is targeted by *Nuak1*, *Tcf7l2*, *Tfdp1*, and *Zfhx3*, whereas *Col14a1* is targeted by *Hif1a*, *Tcf7l2*, *Tfdp1*, and *Zfp148*. Performing the same synthesis for the top DEGs that represented other functionality, we found that many of these genes are also regulated by the same sets of TFs that regulate specific fibroblast functionalities (Supplementary Figure S12e). For example, *Dkk2*, which is a significant DEG for juvenile postnatal day 21 skin, is regulated by *Egr3* and *Hoxd9*. In young adult postnatal day 49 skin, *Ndufa4l2* is a downstream target of *Ar*, *Rarg*, and *Trp63*. In early SW PWD4 wounds, *Plac8* is inferred to be regulated by *Irf4* and *Irf9*. In LW PWD12 wounds, we observed that the top DEG, *Crabp1*, was regulated by *Elk3*, whereas important upper identity markers *Fabp5* and *Prss35* were regulated by *Elk3*, *Etv3*, *Kdm5b*, *Maz*, *Nfatc1*, *Rest*, *Smarca4*, and *Zfp281* and *Elk3*, *Kdm5b*, and *Zfp281*, respectively. In late LW PWD18 wound, the top DEG, *C1qtnf3*, was regulated by *Cpeb1*, *Hoxa9*, and *Zfhx3*. Many of these inferred TFs have appeared to have significant regulatory roles for key fibroblast functions. In juvenile postnatal day 21 skin, REG-4 and REG-9 are the dominant regulatory clusters, the former of which is also the dominant fibroblast regulatory state in wounded SW PWD7 (postnatal day 21 + 7-day) skin. REG-4 fibroblasts differentially expressed *Egr3* and *Klf5*. *Egr3* is induced by TGFβ signaling and promotes fibrosis (Fang et al, 2013), whereas *Klf5* enables excessive ECM synthesis upon systemic sclerosis (Noda et al, 2014). REG-9 fibroblasts, which significantly overlap with dermal papilla cells, differentially express *Lef1*, *Prdm1*, *Tcf7l1*, and *Trps1*. In particular, *Lef1* is key to fibroblast development and promotes wound-induced hair neogenesis (Phan et al, 2021), whereas *Prdm1* and *Trps1* both regulate hair follicle formation (Fantauzzo and Christiano, 2012; Telerman et al, 2017). In unwounded postnatal day 49 skin and wounded SW PWD4 (postnatal day 49 + 4-day) skin, REG-2 was the dominant fibroblast regulatory state. REG-2 fibroblasts were marked by expression of *Ar* and *Tcf7*. In early wounded skin, REG-7 fibroblasts, which express *Gata6* and *Pparg*, emerged as an additional gene regulatory state. In lung pulmonary fibrosis, *Gata6* regulates quiescence of myofibroblasts (Leppäranta et al, 2010), whereas *Pparg* is regulated by TGFβ and helps to inhibit fibroblast activation and mediates fibrosis in systemic sclerosis (Wei et al, 2010). REG-1 fibroblasts were the dominant gene regulatory state in LW PWD12 wounds and significantly expressed *Nfatc1* and *Nfkb1*. In PWD14 fibroblasts, the dominant regulon cluster was REG-5, for which *Nuak1* is differentially active. Recently, it was shown that *Nuak1* promotes YAP and TGFβ signaling, which can exacerbate scarring (Zhang et al, 2022). LW PWD18 wounds consisted primarily of REG-3 and REG-6 fibroblasts. REG-3 showed

significant expression of *Cpeb1* and *Tcf7l2*. *Cpeb1* is a downstream target of TGFβ and promotes fibrotic markers such as α-smooth muscle actin, type-I collagen, and fibronectin in postburn scars (Cui et al, 2020). REG-6 fibroblasts significantly expressed *Barx2*, which regulates hair follicle remodeling (Olson et al, 2005), and *Snai2*. As such, from the integrated SCENIC and differential expression analysis, we predict that the observed distinct fibroblast functionalities across wound healing time are driven by dynamic changes in fibroblast gene regulatory behavior. From these results, we conclude that fibroblast gene regulatory states and the observed transcriptional states overlap significantly with additional condition-driven nuances, primarily, animal age and postwounding time.

Supplementary Text S6: Immune–fibroblast signal interactions exhibit a diverse range of temporal modes

It is known that the immune cell environment changes significantly over wound healing to address different functional needs (Eming et al, 2017), which was confirmed by subclustering of the immune cell population (Supplementary Figure S3). We then asked whether these changes in the immune cell population contributed to the observed changes in fibroblast states through cell–cell signaling. At each time point, we inferred the most significant cell–cell communication pathways between the immune states present and the dominant fibroblast clusters using CellChat (Supplementary Figure S13). We then summarized the signaling pathways by their dynamics with respect to signals secreted from immune cells to fibroblasts (Supplementary Figure S13a), signals secreted from fibroblasts to other fibroblasts (Supplementary Figure S13b)—which we refer to as autocrine signaling—and signals secreted from fibroblasts to immune cells (Supplementary Figure S13dc. To focus on specific wound-induced signaling pathways, for each of Supplementary Figure S13a–c, we excluded pathways that were present in postnatal day 21 and postnatal day 49 unwounded skin. Motivated by previous results (Supplementary Figure S6), we also inferred which of these pathways may be mitogenic and help to maintain proliferative competency.

Examining immune-to-fibroblast signaling (Supplementary Figure S13a), we observed that most signaling pathways tended to peak at a single time point and thus were temporally unimodal in activity. Most unimodal pathways peaked in early PWD4 wounds, including CXC chemokines (CXCL), IL-1, IL-4, oncostatin M, and osteopontin (SPP1). Some pathways activated and then remained active across consecutive time points, such as MIF (PWD7 through PWD12) and TGFβ (PWD4 through PWD7); we referred to these pathways as being prolonged in activity. One pathway, TNF, demonstrated what we refer to as a bimodal signaling pattern: it emerged at one time point, PWD4, before disappearing and then re-emerging at a later time point, PWD12. Examining autocrine fibroblast-to-fibroblast signaling (Supplementary Figure 13b), we observed that only CC fibroblast-driven chemokine (CCL) was dynamically unimodal, emerging in early PWD4 wounds. Fibroblast-driven MIF signaling behaved in parallel to its immune-driven counterpart and was prolonged over consecutive time points, albeit starting later in PWD7 wounds. We observed that fibroblast-driven CXCL, periostin, and PROS exhibited bimodal signaling patterns, emerging in early PWD4 wounds before subsiding and not emerging again until late PWD18 wounds. We observed only unimodal signaling patterns for pathways secreted from fibroblasts to immune cells (Supplementary Figure 13c). Most pathways, including the annexin A1 pathway, CCL, complement, CXCL, galectin, MIF, and SPP1, were active in early PWD4 wounds. The remaining pathways, IL-1 and PTN, were active in later PWD12 wounds. Of the observed pathways, annexin A1 and complement were uniquely secreted from fibroblasts to immune cells.

SUPPLEMENTARY MATERIALS AND METHODS

Datasets

We integrated the following conditions of sorted and unsorted scRNA-seq data of murine skin fibroblasts, all of which were prepared for sequencing using 10X Genomics Chromium 3′ v2 kits and sequenced on the Illumina HiSeq 4000 platform:

1. Unwounded postnatal day 21 (Phan et al., 2020). This dataset consists of 2 replicates of unsorted, unwounded juvenile postnatal day 21 murine skin. Data were accessed from National Center for Biotechnology Information (NCBI) Geo-Datasets (GSE153596).
2. Unwounded postnatal day 49 (Haensel et al, 2020). This dataset consists of 2 replicates of unsorted, unwounded young adult postnatal day 49 murine skin. Data were accessed from NCBI Geo-Datasets (GSE142721).
3. SW PWD4 (Haensel et al, 2020). This dataset consists of 3 replicates of unsorted SW postnatal day 49 + 4 (6-mm diameter, sampled from wound and adjacent tissue) murine skin. Data were accessed from NCBI Geo-Datasets (GSE142721).
4. SW PWD7 (Phan et al, 2020). This dataset consists of 3 replicates of unsorted SW postnatal day 21 + 7 (2-mm diameter, sampled from wound edge) murine skin. Data were accessed from NCBI Geo-Datasets (GSE153596). One sample was provided by Ryan Driskell through e-mail correspondence.
5. LW PWD12 (Guerrero-Juarez et al, 2019). This dataset consists of 12 pooled replicates of unsorted LW (1.5-cm square, sampled from whole wound) postnatal day 42–56 + 12 murine skin dermis. Data were accessed from NCBI Geo-Datasets (GSE113854).
6. LW PWD14 (Abbasi et al, 2020). This dataset consists of FACS-sorted LW (1.5-cm diameter) postnatal day 28 + 14 murine skin fibroblasts, sampled specifically from the center and periphery of the wound. Data were accessed from NCBI Geo-Datasets (GSE108677).
7. LW FIB PWD18 (Gay et al, 2020). This dataset consists of unsorted (1.2-cm diameter) postnatal days 42–56 + 18 large fibrotic wounds sampled from murine skin dermis, where no hair neogenesis was observed. Data were accessed from NCBI Geo-Datasets (GSE141814).
8. LW REG PWD18 (Gay et al, 2020). This dataset consists of unsorted (1.2-cm diameter) postnatal days 42–56 + 18 large regenerative wounds sampled from murine skin dermis, where wound-induced hair neogenesis was

observed. Data were accessed from NCBI Geo-Datasets (GSE141814).

These studies have demonstrated the usefulness of scRNA-seq in revealing insights about wound healing. However, one limitation of these studies is that only one postwounding time point is considered (in addition to an unwounded skin condition). These time points can be combined using integration, which is a powerful computational technique that has been used to successfully reconstruct developmental trajectories in embryogenesis (Petitpré et al, 2022; Qiu et al, 2022), reveal novel transition states in tissue regeneration (McKellar et al, 2021), establish a consensus cell-type reference across organs (Domínguez Conde et al, 2022; Elmentaite et al, 2021), characterize disease pathogenesis (Melms et al, 2021; Wendisch et al, 2021), and more.

Data preprocessing

We process all scRNA-seq data using the Scanpy toolkit (Wolf et al, 2018). To save time on preprocessing, we matched the generated RNA velocity counts to the published gene expression matrices provided by the authors at the dataset repositories mentioned earlier. Potential doublets were removed from each dataset by running the doublet detection tool Scrublet (Wolock et al, 2019) on the raw, prenormalized counts. We then filtered each dataset separately using the following quality control parameters:

1. Unwounded postnatal day 21 (Phan et al, 2020). We kept cells with 250−5500 genes and 300−25,000 counts. We kept genes expressed in at least 3 cells. After quality control, we retained 5921 cells.
2. Unwounded postnatal day 49 (Haensel et al, 2020). We kept cells with 500−5000 genes and 300−25,000 counts. We kept genes expressed in at least 3 cells. After quality control, we retained 10,572 cells.
3. SW PWD4 (Haensel et al, 2020). We kept cells with 400−5000 genes and 300−40,000 counts. We kept genes expressed in at least 3 cells. After quality control, we retained 16,106 cells.
4. SW PWD7 (Phan et al, 2020). We kept cells with 400−4500 genes and 300−30,000 counts. We kept genes expressed in at least 3 cells. For this condition, we observed abnormally high expression of ribosomal protein−related genes, particularly in keratinocytes. To improve integration and clustering, we filtered out the ribosomal protein−related genes for this dataset. After quality control, we retained 32,330 cells.
5. LW PWD12 (Guerrero-Juarez et al, 2019). We kept cells with 450−3000 genes and 300−10,000 counts. We kept genes expressed in at least 3 cells. After quality control, we retained 22,129 cells.
6. LW PWD14 (Abbasi et al, 2020). We kept cells with 800−5000 genes and 300−5000 counts. We kept genes expressed in at least 3 cells. After quality control, we retained 5793 cells.
7. LW FIB PWD18 (Gay et al, 2020). We kept cells with 250−5000 genes and 300−30,000 counts. We kept genes expressed in at least 3 cells. After quality control, we retained 4613 cells.
8. LW REG PWD18 (Gay et al, 2020). We kept cells with 250−5000 genes and 300−30,000 counts. We kept genes expressed in at least 3 cells. After quality control, we retained 10,314 cells.

We also removed cells with a mitochondrial gene expression percentage above 10% for each dataset. After filtering, we combined the datasets into a single annotated data frame, taking the union of all identified genes. Each condition was labeled using the observation variable sample, representing, whereas replicates for conditions were labeled using the observation variable sub_sample. The merged gene expression data are then row normalized, such that each cell has 10,000 total reads, and log transformed (natural logarithm) with a pseudocount of 1 added to each entry prior to transformation.

Subclustering

We now describe the parameters and methods to obtain the subclusters for the identified major cell lineages, including keratinocytes fibroblasts, immune cells, vascular cells, and Schwann cells.

Keratinocytes. We isolated the initial keratinocyte clusters, EPI-1 through EPI-5, for further subclustering. When integrating the epidermal cell population, we calculated the top 2,000 highly variable genes using the seurat flavor (rather than the cell_ranger flavor) owing to computational issues arising from keratinocytes not being present across all samples. We performed 2 rounds of data integration and subclustering to remove erythrocytes. For the first round of Leiden clustering, we set resolution = 0.4, which identified 10 epidermal subclusters. After removing the erythrocyte population and reintegrating the population, we set the Leiden clustering parameter to resolution = 0.6, obtaining 14 epidermal subclusters. From marker genes, we identified 5 subclusters of basal epidermal cells (BAS-1 through BAS-5), 2 subclusters of cycling epidermal cells (CYC-1 through CYC-2), 2 subclusters of hair follicle cells (HF-1 through HF-2), one spinous epidermal subcluster, and 4 subclusters of wound-associated epidermal cells present in SW PWD7 wounds (WO-1 through 4) (Supplementary Figure S1).

Fibroblasts. We isolated the initial fibroblast clusters, FIB-1 through FIB-4, for further subclustering. We followed the same data integration process as described previously when integrating all cell types, but for fibroblasts, we only calculated the top 2,000 highly variable genes. We found that we had to perform 2 rounds of data integration and clustering. After the first round of Leiden clustering, where we set setting resolution = 0.5, we identified 12 fibroblast subclusters. However, upon inspection of marker genes, we observed that 1 subcluster had significant expression of hemoglobin-associated genes, such as *Hba-a1*, indicating that this subcluster represented erythrocytes rather than actual fibroblasts. We therefore removed this subcluster and reintegrated the remaining population of fibroblasts. For the second round of Leiden clustering, we set resolution = 0.7, obtaining 11 fibroblast subclusters. We also found a subcluster of dermal papilla cells, a specialized type of fibroblasts that reside at the base of hair follicles and were originally identified in the

study of postnatal day 21 skin, on the basis of expressions of *Crabp1*, *Corin*, *Lef1*, and *Rspo3* (Phan et al, 2020; Plikus et al, 2021).

Vascular cells. We isolated and subclustered the 2 predominant vascular cell types, endothelial cells and pericytes. We did not reintegrate lymphatic endothelial cells nor smooth muscle cells owing to their small population sizes. We integrated and subclustered each cell type only once, calculating the top 2,000 highly variable genes. When subclustering endothelial cells, we set the Leiden clustering parameter resolution = 0.3, obtaining 4 subclusters. For pericytes, we set resolution = 0.4, obtaining 5 pericyte subclusters (Supplementary Figure S2).

Immune cells. We isolated the initial immune clusters, IMM-1 through IMM-3, for further subclustering. We followed the same data integration process when integrating all cell types, calculating the top 2000 highly variable genes. As for fibroblasts, we had to integrate and subcluster the population twice to remove a population of erythrocytes. For the first round of Leiden clustering, we set the parameter resolution = 0.4, identifying 15 immune subclusters. For the second round of Leiden clustering, we set resolution = 0.3, obtaining 14 immune subclusters. To identify immune cell types more specifically, we inputted the top 50 marker genes from each subcluster into the MyGeneset tool from ImmGen (http://rstats.immgen.org/MyGeneSet_New/index.html) and Enrichr (Xie et al, 2021) and compared the suggested cell types. We identified one basophil subcluster, 4 dendritic cell subclusters (DC-1 through DC-4), 5 macrophage subclusters (MAC-1 through MAC-5), one T-cell cluster, one NK cell cluster, one plasmacytoid dendritic cell cluster, one mast cell subcluster, and one neutrophil subcluster.

We then further subclustered the T-cell population using the same integration procedure but with a cluster resolution of resolution = 0.3. We identified 5 T-cell subclusters of Tregs, NK T cells, T-lymphocyte presenting innate lymphoid cells, and a small population of activated T cells, marked by significant expression of *Spp1* (Supplementary Figure S3).

Schwann cells. We isolated the initial Schwann cell cluster for further subclustering. We recalculated the top 2000 highly variable genes using the cell_ranger flavor. We performed one round of integration and subclustering. We set the Leiden cluster parameter to resolution = 0.4, resulting in 5 Schwann cell subclusters (Supplementary Figure S4).

Data integration

For the initial round of clustering, we identified 4000 highly variable genes across all datasets using the cell_ranger flavor. Here, the sample observations are the batches to be integrated. To account for the multiple batches, we identified the highly variable genes in an iterative process (Luecken et al, 2022). We started by identifying 4000 highly variable genes per batch. Each gene was then ranked by the number of batches across which the gene is highly variable and then by the mean dispersion parameter across batches. We first considered genes that are highly variable across all n batches (n = 8) and then genes that are highly variable across n − 1 batches and so on, until we have obtained 4000 highly variable genes.

We integrated the datasets using Scanorama (Hie et al, 2019), which integrates datasets using panoramic stitching, a technique from image analysis, and approximates nearest neighbor searching in compressed gene space (low-dimensional embeddings) to match similar cells across datasets. We chose Scanorama for its ability to distinguish rarer cell types, albeit at the cost of reduced batch correction, compared with other methods such as BBKNN (Polański et al, 2020) and Harmony (Korsunsky et al, 2019). We found no substantial difference between the integrated embedding produced by Scanorama and the integrated embedding produced by scVI (Lopez et al, 2018), which integrates data using more sophisticated generative modeling and variational inference. Thus, we chose the embedding generated from Scanorama because the computational time was dramatically lower.

To prepare the combined dataset for Scanorama, we split the data back into the original batches. We then ran the function integrate_scanpy, which returns a list of low-dimensional embeddings, and X_scanorama, which are used in lieu of embeddings produced by standard principal component analysis. To improve data integration, we ran Scanorama on a subsetted copy of the original dataset, where only the 4000 highly variable genes were considered (Luecken et al, 2022). We then merged the resultant list of embeddings and attached them to the original combined dataset. Then, we generated the kNN graph, setting k = 30, and clustered the population using the Leiden algorithm, setting resolution = 0.3. The clustering resolution was chosen to balance between the need to identify rarer cell types present in one or 2 wounding conditions and the precaution against overclustering by identifying cell states that may occur owing to variation in nonbiological factors.

When integrating scRNA-seq data, there are 2 commonly desired outcomes: first, to remove technical batch effects that confound cell-type identification, and second, to conserve biological variation by properly distinguishing between cell types shared across many conditions and cell types that are unique to only some conditions. Whether one outcome is desired more so than the other depends on scientific question and features of the considered data because different integration methods tend to be better at only one of batch effect removal or biological variance conservation (Luecken et al, 2022). The challenge with skin data is that its high cell-type diversity requires the chosen data integration method to achieve a balance of both batch effect removal and biological variance conservation. To assess quality of integration, we used the package scIB (Luecken et al, 2022) to calculate several metrics to assess the quality of batch correction and biological conservation. Metrics under the former category gauge how well cell types from different datasets are integrated, whereas metrics under the latter category gauge how similar cells of the same cell label (eg, cluster) are. For all metrics, values closer to 1 indicate good performance, whereas values closer to 0 indicate poor performance.

To assess batch correction, we calculated the Average Silhouette Width (ASW) batch score, graph connectivity score, the principal component regression batch score, and

the kBET (k-Nearest Neighbor Batch Effect Test) score. The ASW batch score measures the amount of batch mixing per cell-type label, where an ASW score of 1 indicates ideal batch mixing. The graph connectivity score measures the degree to which all cells with the same cell-type label are connected, that is, that more similar cells should be closer on the neighborhood graph. The principal component regression batch score measures the variance contribution to gene expression data contributed by batch effects, where a smaller principal component regression score corresponds to reduced contribution by batch effects. The kBET score measures the degree to which the cell-type label composition of a cell's local kNN graph neighborhood corresponds to the global label composition of the entire kNN graph.

To assess biological conservation, we calculated the normalized mutual information label score, the isolated label F1 score, the ASW label score, and the isolated silhouette score. The normalized mutual information label score measures the overlap between the original clustering labels and the optimal Louvain clustering. The isolated label F1 score measures the degree to which cell types that are rare across batch are successfully grouped together in the neighborhood embedding. The ASW label score measures how well cells from the same type are mixed and how well cells from different cell-type labels are separated. The isolated silhouette score measures the relationship between the ASW of isolated cell-type labels that are present in the fewest number of batches after integration and the ASW of nonisolated cell-type labels. For all clustering and subclustering analysis, we identified all cell lineages and cell states using marker genes inferred using Mann–Whitney *U* test (method = Wilcoxon in Scanpy).

Analysis of gene correlations with upper and lower fibroblast scores

To further discern which genes distinguished fibroblasts of primarily upper identity from fibroblasts of primarily lower identity, we used the tradeSeq package (Van den Berge et al, 2020) to fit the unnormalized counts of the 2000 highly variable fibroblast genes (as inferred during integration) as continuous functions of the lower and upper fibroblast scores. Gene expression was fitted to the scores using cubic smoothing splines, where each gene was modeled using a linear combination of $K = 6$ cubic basis functions. From the 2000 highly variable genes, we performed a Wald test to determine which genes varied in a statistically significant manner with either the upper or lower score, that is, the associationTest implemented by tradeSeq. We corrected the Wald test–derived *p*-values for multiple hypothesis testing using Bonferroni correction, and then only selecting for genes where on the adjusted *p*-value $Q < .05$, we inferred 1571 genes as differentially associated with either the upper or lower score. We then ranked these genes by the Spearman correlation between their predicted gene expressions with interpolated grids constructed from the upper and lower scores separately.

Replication wounding studies

We performed 2 sets of wounding experiments on C57BL/6J mice. We first recapitulated the wounding experiments as they were described in the original integrated studies. The wounding procedures involved conducting 2-mm punch wounds on mice aged 3 weeks, 6-mm punch wounds on mice aged 7 weeks, 1.5-cm excisional wounds on mice aged 4 weeks, and 1.2-cm and 1.5-cm excisional wounds on mice aged 7 weeks. Postwound skin tissues were collected at specific time points indicated in the original studies. The second set of wounding experiments were performed using consistent 6-mm punch wounds on C57BL/6J mice at postnatal day 49. Postwound skin tissues were collected at PWD4, 8, 10, and 14. The wounding procedures were approved by the Institutional Animal Care and Use Committee at the University of California, Irvine (Irvine, California). We prepared sections from fresh-frozen tissues and processed them with RNAscope multiplex fluorescent detection kit, version 2 (323100, Advanced Cell Diagnostics), following the manufacturer's instrutions with minor adjustments. The following mouse probes were used in this study: *Lum* (480361), *Prg4* (437661-C2), *Mki67* (416771), *Pdgfr-C3* (480661-C3), *Pamr1* (491431-C2), and *Col25a1* (538511-C2). DAPI was used for nuclei counterstaining. Images were captured on the Olympus FV3000 confocal laser scanning microscope and the Keyence BZ-X810 Microscope.

Analysis of fibroblast distribution from Foster et al (2021)

To partially validate our predictions that fibroblasts assume distinct upper and lower spatial identity during wound healing, we briefly reanalyzed the spatial transcriptomics data generated in Foster et al (2021). In this study, 10X Visium spatial transcriptomics datasets were generated on silicone ring-splinted 6-mm dorsal skin wound samples collected at PWD2, 7, and 14. For each PWD time point, we then removed genes that were expressed in fewer than 3 cells and removed spots that did not express at least 10 genes. We then normalized the gene expression counts such that the total counts summed to 10,0000 and log transformed the normalized data, adding a pseudocount of 1. For each time point, we used Squidpy to generate a spatial graph, where each spot was connected to its 6 nearest neighbors. Using these spatial graphs, we then performed Leiden clustering (resolution = 0.0001 for PWD2, resolution = 0.3 for PWD7, no clustering for PWD14) to isolate the tissue slices analyzed in the original Foster et al (2021)'s study. For each time point, we calculated a keratinocyte score using the marker genes, *Krt5* and *Lgals7*, which were inferred from our original integration (Figure 1f). We also calculated a fibroblast score using pan-fibroblast markers, *Dcn* and *Lum*, which we inferred from our integration (Figure 1f), as well as *Pdgfra*. To determine upper and lower fibroblast distribution, we profiled the expression of the most significantly expressed upper marker, *Crabp1*, and the most significantly expressed lower marker, *Ly6a*, which were chosen from our curated lists from the literature (Supplementary Table S1). To analyze upper fibroblast enrichment specifically, we weighted *Crabp1* expression by the previously calculated fibroblast score (which was normalized to be between 0 and 1), so that we primarily visualized *Crabp1* expression in fibroblast-enriched spots. Similarly, to analyze lower fibroblast enrichment, we weighted *Ly6a* expression by the previously calculated fibroblast score (which was normalized to be

between 0 and 1), so that we primarily visualized *Ly6a* expression in fibroblast-enriched spots.

Analysis of fibroblasts from Mascharak et al (2022)

To partially validate our integrated analysis of fibroblast dynamics during wound healing, we separately analyzed scRNA-seq data generated from unwounded and SW skin samples reported in Mascharak et al (2022). In this study, scRNA-seq was generated from full-thickness 6-mm SWs sampled at PWD2, 7, 14, and 30. From personal correspondence, we received the Seurat object (Stuart et al, 2019) used for analysis by the authors. Following the original study, we annotated cell types using SingleR and the mouse RNA-sequencing reference dataset (Aran et al, 2019). After automated annotation, we isolated all cells annotated as fibroblasts from the PBS-injected control condition for further analysis, yielding a dataset of 364 cells and 14,236 genes. DEGs for each wound healing time point were identified using a Mann–Whitney U test (method = Wilcoxon in Scanpy). To identify shared DEGs between PWD2 fibroblasts and the PWD4-specific FIB-IV subcluster, we selected only DEGs from each subcluster where their log fold change was greater than 1.0 and their adjusted p-value after correcting for multiple hypothesis testing was $<.05$.

Biased clustering

Single-cell states are typically inferred by discriminatory marker genes. Because scRNA-seq contain upward of 10,000 genes—for this considered data, we analyzed 21,485 genes—it can be difficult to determine whether the genes driving the observed cell states are meaningful—in general, one challenge of single-cell analysis is the lack of interpretability for many of the unsupervised methods. Because many cell types and states can be distinguished by differences in their function, it is worth dissecting how identified cell states can be discriminated by key functions rather than by just a select few marker genes. In particular, we attempt to discriminate the identified fibroblast states (subclusters) by key fibroblast functions, which have been disseminated over time (Plikus et al, 2021). Key fibroblast functions include cell signaling, both sending through secretion of ligands and receiving through receptor expression; synthesis of various ECM components; expression of molecules that degrade and remodel the local ECM; and more. Furthermore, these functions are associated with gene sets that have been recorded in publicly available databases (Gillespie et al, 2022; Jin et al, 2021; Shao et al, 2020). How to analyse scRNA-seq data in terms of these imposed gene categories, which improve interpretability, is of great research interest.

To understand whether the identified fibroblast states differ by key fibroblast functionality, we construct 4 gene sets corresponding to key fibroblast functions: signal ligands, signal receptors, ECM synthesis, and ECM modifying. The signal ligand and receptor gene sets were created by extracting ligands and receptors, respectively, from interactions annotated by either cell–cell contact or secreted signaling from CellChatDB (Jin et al, 2021); the ECM synthesis gene set was created by extracting ligands from interactions annotated by ECM receptor from CellChatDB (Jin et al, 2021), extracting genes from the categories ECM glycoproteins, collagen, and proteoglycans from MatrisomeDB (Shao et al, 2020) and manually adding genes relevant to ECM glycosaminoglycans; the gene set for ECM remodeling was created by extracting genes from the category ECM regulators from MatrisomeDB (Shao et al, 2020). After curation, we obtained 435 signal ligands, 329 signal receptors, 416 ECM synthesis genes, and 228 ECM-modifying genes. We note that one or several of these gene sets overlap partially. Most notably, there are 47 cell–cell contact and surface marker genes that are classified as both signal ligands and signal receptors.

We performed a series of biased integration and clustering analyses, where we considered the full fibroblast data only with respect to the functional gene sets and filtered out all other genes. We integrated the data on all functional genes and integrated the data by isolating the genes for each functional category alone. Each time, we generated a new kNN graph and reclustered the population. We validated the inferred clusters for each considered functional gene subset by examining for meaningful discriminatory marker genes. The functional subclusters were then compared with the original subcluster labels to determine the degree to which the considered gene set captured the full fibroblast transcriptomic heterogeneity. We calculated the similarity between the functional subclusters and the original subclusters using the Adjusted Rand Index and the normalized mutual information, which have also been used to assess data integration and batch effect correction methods for scRNA-seq data (Luecken et al, 2022).

To complement the functional clustering analysis, we also trained the fibroblast data using an interpretable autoencoder Intercode (Rybakov et al, 2020[3]). Intercode constructs an interpretable latent space for the considered scRNA-seq population using gene set annotations—in this case, the fibroblast functions—such that the dimensions of the latent space, which measure the degree of biological variation in the population, are constrained to correspond to variations with respect to the gene sets. The model consists of a nonlinear encoder and a linear decoder, where interpretability is enforced by introducing additional regularization terms in the loss function. Using Intercode, we can determine how well key fibroblast functions distinguish the original fibroblast subclusters, providing information into the differences in functions between the single-cell states. We followed the workflow described in the original study by Rybakov et al (2020)[3]. A binary annotation matrix was constructed from the curated functional gene sets. Per the advice of the study, to improve accuracy of the model during training, we only considered genes comprising the intersection between the functional gene set and the top 2000 highly variable genes. We trained the autoencoder, setting the loss rate of 0.001, the batch size to be 64, and the number of epochs to be 50. For regularization, we set $\lambda_0 = 0.1, \lambda_1 = 0.93, \lambda_3 = 0.57$ and did not consider sparse unannotated factors (we refer to Rybakov et al [2020][3] for more details).

To validate the functional clusters, we performed differential expression analysis with respect to each functional set. For each functional gene subset of the fibroblast population, we built a zero-inflated negative binomial hurdle model using MAST (Finak et al, 2015), where wound healing time was

included as a covariate. DEGs for each cluster and each wound healing time were inferred using a likelihood ratio test. We then verified the specificity of each DEG by manually examining the resultant feature plot.

Regulon inference

To infer possible epigenetic regulators of fibroblast states, we used pySCENIC (Van de Sande et al, 2020) to infer significant regulons containing TFs and corresponding downstream targets. Following the suggested pySCENIC analysis pipeline, we used the raw, untransformed fibroblast gene expression counts as input. Following the analysis pipeline, we inferred significant gene regulatory networks using GRNBoost2 and a curated list of provided TFs relevant to *Mus musculus* provided by pySCENIC developers (Van de Sande et al, 2020). The resultant output was a list of adjacency matrices encoding links between each TF and its targets. Regulons were then inferred using the adjacencies list, using the database of MGI motif annotations provided by pySCENIC and rankings databases of the mouse genome from cisTargetDBs. We used the ranking databases corresponding to the mm10 version of the genome (refseq_r80). Finally, we used regulon inference output to infer cellular enrichment of each significant TF.

To generate regulon-based clusters, we generated a kNN graph ($k = 30$) using the regulon enrichment scores (X_AUCell), on which we Leiden clustered the fibroblast population, setting resolution = 0.6. Then, we considered the dominant one or 2 regulon clusters per wound healing time and calculated the intersection between the downstream targets of the top 20 TFs inferred by pySCENIC and the top 300 DEGs of fibroblasts within each regulon cluster. The intersections were then categorized as corresponding to the previously curated fibroblast functions (signal ligand, signal receptor, ECM synthesis, ECM remodeling) or other functionality.

Trajectory analysis

To generate RNA velocity counts, we first converted BAM files obtained from the repositories mentioned earlier into the raw, unaligned FASTQ files using the 10X Genomics tool bamtofastq. RNA velocity counts were generated using kallisto|bustools (Melsted et al, 2019[4]). We first generated a velocity pseudoalignment index from the *Mus musculus* mm10 genome (GRCm38, Ensembl release 98) by running the command kb ref with the –workflow lamanno option, which identifies the spliced and unspliced transcripts to capture. To properly align the data obtained from Abbasi et al (2020), we appended the tdTomato sequence to the mm10 genome because tdTomato was used to label fibroblasts in this study. Then, we ran the command kb count with the –workflow lamanno to generate spliced and unspliced mRNA count matrices, stored as h5ad files. We did not realign the data from Phan et al (2020); rather, we used the published Loom files provided by the authors at the above Gene Expression Omnibus Datasets repository because the Loom files already contained RNA velocity counts.

To infer possible trajectories in the fibroblast population, we used the package scVelo (Bergen et al, 2020), which extends the solves of the fully time-dependent, dynamical formulation of the single-cell RNA splicing differential equation model put forth by La Manno et al (2018). Splicing rates for each gene are then inferred using a likelihood-based model. For each dataset except those from unwounded postnatal day 21 and SW PWD7 skin, spliced and unspliced mRNA counts for each gene were generated using kallisto|bustools (Melsted et al, 2019[4]). RNA velocity counts from unwounded postnatal day 21 and SW PWD7 skin were generated using velocyto (La Manno et al, 2018). Before downstream analysis, we filtered out genes that did not have at least 20 shared counts across spliced and unspliced mRNA. The kNN graph was calculated using the integrated Scanorama embedding to account for possible batch effects across the multiple datasets. The first and second moments of velocity were calculated for each cell across its 30 neighbors. Velocities and splicing rates were then estimated using the dynamical option of scVelo. For visualization purposes, we projected the estimated velocities and directional flows onto the original Uniform Manifold Approximation and Projection produced during subclustering, labeling both wound healing times and fibroblast subclusters. To obtain a more interpretable, coarse-grained picture of possible transitions between fibroblast subclusters, we also calculated the partition-based graph abstraction, that is, the PAGA graph (Wolf et al, 2019), which estimates connectivity between graph partitions on the basis of the generated kNN graph used for clustering and, in turn, leads to more robust inference of lineage trajectories in scRNA-seq data. The edge weights of the PAGA graph then give a measure of the likeliest transitions between graph partitions, that is, the graph nodes. To obtain estimates of edge directionality, we use the pseudotime values inferred from RNA velocity, which are calculated using a random-walk–based distance measure between cells on the velocity graph obtained from scVelo. We then overlay the directed PAGA graph onto the previously calculated Uniform Manifold Approximation and Projection.

Cell–cell communication analysis

We inferred significant cell–cell communication activity using CellChat (Jin et al, 2023[5], 2021). For the time points—unwounded postnatal day 21 skin, unwounded postnatal day 49 skin, and wounded tissues (SW PWD4, SW PWD7, LW PWD12, and LW REG PWD18), we isolated the immune subclusters and the dominant fibroblast subclusters present at those times. At each considered time point, we implemented the standard CellChat pipeline, but to avoid biasing cell–cell communication inference due to the presence of small cell populations, we removed cell types that represented <0.1% of the population using the filterCommunication command. To summarize cell–cell communication activity, we summed up individual interaction scores for each inferred signaling pathway between immune cells and fibroblasts. To increase interpretability of results, we removed weaker signaling pathways where the maximum aggregate score fell below a specified threshold, threshold = 0.05.

[4] Melsted P, Sina Booeshaghi A, Gao F, Beltrame E, Lu L, Hjorleifsson KE, et al. Modular and efficient pre-processing of single-cell RNA-seq. bioRxiv 2019:1–18.

[5] Jin S, Plikus MV, Nie Q. CellChat for systematic analysis of cell-cell communication from single-cell and spatially resolved transcriptomics. bioRxiv 2023:1–65.

To better summarize signaling patterns, we focused on the pathways that were exclusively driven by signaling ligands secreted from immune cells to fibroblasts (Figure 6a), exclusively driven by signaling ligands secreted by fibroblasts to other fibroblasts (Figure 6b), and exclusively driven by fibroblasts to immune cells (Figure 6c). We identified 5 unique immune-to-fibroblast signaling pathways in wound healing, which exhibited 3 types of signaling patterns (Figure 6a). Three paracrine pathways, IL-1, IL-4, and oncostatin M, are dynamically unimodal and peak in early PWD4 wound. One pathway, TNF, is dynamically bimodal, emerging in early wound and reactivating in late wound. The remaining pathway, TGFβ, emerges initially in early wound and is active for a prolonged period of biological time. In contrast, we identified only one pathway that was exclusively driven from fibroblasts to other fibroblasts, PROS, which is dynamically bimodal, activating in early wounds and reactivating in late wounds (Figure 6b). We identified 4 signaling pathways that are exclusively secreted by fibroblasts to immune cells. Of the 4, 3 pathways, which are annexin, colony-stimulating factor, and complement, exhibit a unimodal pattern, whereas galectin is bimodal in signaling behavior (Figure 6c).

SUPPLEMENTARY REFERENCES

Abbasi S, Sinha S, Labit E, Rosin NL, Yoon G, Rahmani W, et al. Distinct regulatory programs control the latent regenerative potential of dermal fibroblasts during wound healing [published correction appears in Cell Stem Cell 2021;28:581–3]. Cell Stem Cell 2020;3:396–412.e6.

Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol 2019;20:163–72.

Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol 2020;38:1408–14.

Birbrair A, Zhang T, Files DC, Mannava S, Smith T, Wang ZM, et al. Type-1 pericytes accumulate after tissue injury and produce collagen in an organ-dependent manner. Stem Cell Res Ther 2014;5:122.

Bodnar RJ, Satish L, Yates CC, Wells A. Pericytes: a newly recognized player in wound healing. Wound Repair Regen 2016;24:204–14.

Cui HS, Joo SY, Cho YS, Kim JB, Seo CH. CPEB1 or CPEB4 knockdown suppresses the TAK1 and Smad signalings in THP-1 macrophage-like cells and dermal fibroblasts. Arch Biochem Biophys 2020;683:108322.

Doğan A. Apelin receptor (Aplnr) signaling promotes fibroblast migration. Tissue Cell 2019;56:98–106.

Domínguez Conde C, Xu C, Jarvis LB, Rainbow DB, Wells SB, Gomes T, et al. Cross-tissue immune cell analysis reveals tissue-specific features in humans. Science 2022;376:eabl5197.

Driskell RR, Lichtenberger BM, Hoste E, Kretzschmar K, Simons BD, Charalambous M, et al. Distinct fibroblast lineages determine dermal architecture in skin development and repair. Nature 2013;504:277–81.

Elmentaite R, Kumasaka N, Roberts K, Fleming A, Dann E, King HW, et al. Cells of the human intestinal tract mapped across space and time. Nature 2021;597:250–5.

Eming SA, Wynn TA, Martin P. Inflammation and metabolism in tissue repair and regeneration. Science 2017;356:1026–30.

Fang F, Shangguan AJ, Kelly K, Wei J, Gruner K, Ye B, et al. Early growth response 3 (Egr-3) is induced by transforming growth factor-β and regulates fibrogenic responses. Am J Pathol 2013;183:1197–208.

Fantauzzo KA, Christiano AM. Trps1 activates a network of secreted Wnt inhibitors and transcription factors crucial to vibrissa follicle morphogenesis. Development 2012;139:203–14.

Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol 2015;16:278.

Gay D, Ghinatti G, Guerrero-Juarez CF, Ferrer RA, Ferri F, Lim CH, et al. Phagocytosis of Wnt inhibitor SFRP4 by late wound macrophages drives chronic Wnt activity for fibrotic skin healing. Sci Adv 2020;6:eaay3704.

Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, et al. The reactome pathway knowledgebase 2022. Nucleic Acids Res 2022;50:D687–92.

Goss G, Rognoni E, Salameti V, Watt FM. Distinct fibroblast lineages give rise to NG2+ pericyte populations in mouse skin development and repair. Front Cell Dev Biol 2021;9:675080.

Guerrero-Juarez CF, Dedhia PH, Jin S, Ruiz-Vega R, Ma D, Liu Y, et al. Single-cell analysis reveals fibroblast heterogeneity and myeloid-derived adipocyte progenitors in murine skin wounds. Nat Commun 2019;10:650.

Haensel D, Jin S, Sun P, Cinco R, Dragan M, Nguyen Q, et al. Defining epidermal basal cell states during skin homeostasis and wound healing using single-cell transcriptomics. Cell Rep 2020;30:3932–47.e6.

Hie B, Bryson B, Berger B. Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. Nat Biotechnol 2019;37:685–91.

Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Myung P, Plikus MV, et al. Inference and analysis of cell-cell communication using CellChat. Nat Commun 2021;2021:1–20. https://doi.org/10.1101/2020.07.21.214387.

Korosec A, Frech S, Gesslbauer B, Vierhapper M, Radtke C, Petzelbauer P, et al. Lineage identity and location within the dermis determine the function of papillary and reticular fibroblasts in human skin. J Invest Dermatol 2019;139:342–51.

Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 2019;16:1289–96.

La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, et al. RNA velocity of single cells. Nature 2018;560:494–8.

Leppäranta O, Pulkkinen V, Koli K, Vähätalo R, Salmenkivi K, Kinnula VL, et al. Transcription factor GATA-6 is expressed in quiescent myofibroblasts in idiopathic pulmonary fibrosis. Am J Respir Cell Mol Biol 2010;42:626–32.

Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. Deep generative modeling for single-cell transcriptomics. Nat Methods 2018;15:1053–8.

Luecken MD, Büttner M, Chaichoompu K, Danese A, Interlandi M, Mueller MF, et al. Benchmarking atlas-level data integration in single-cell genomics. Nat Methods 2022;19:41–50.

Mascharak S, Talbott HE, Januszyk M, Griffin M, Chen K, Davitt MF, et al. Multi-omic analysis reveals divergent molecular events in scarring and regenerative wound healing. Cell Stem Cell 2022;29:315–27.e6.

McKellar DW, Walter LD, Song LT, Mantri M, Wang MFZ, De Vlaminck I, et al. Large-scale integration of single-cell transcriptomic data captures transitional progenitor states in mouse skeletal muscle regeneration. Commun Biol 2021;4:1280.

Melms JC, Biermann J, Huang H, Wang Y, Nair A, Tagore S, et al. A molecular single-cell lung atlas of lethal COVID-19. Nature 2021;595:114–9.

Noda S, Asano Y, Nishimura S, Taniguchi T, Fujiu K, Manabe I, et al. Simultaneous downregulation of KLF5 and Fli1 is a key feature underlying systemic sclerosis. Nat Commun 2014;5:5797.

Olson LE, Zhang J, Taylor H, Rose DW, Rosenfeld MG. Barx2 functions through distinct corepressor classes to regulate hair follicle remodeling. Proc Natl Acad Sci U S A 2005;102:3708–13.

Petitpré C, Faure L, Uhl P, Fontanet P, Filova I, Pavlinkova G, et al. Single-cell RNA-sequencing analysis of the developing mouse inner ear identifies molecular logic of auditory neuron diversification. Nat Commun 2022;13:3878.

Phan QM, Fine GM, Salz L, Herrera GG, Wildman B, Driskell IM, et al. Lef1 expression in fibroblasts maintains developmental potential in adult skin to regenerate wounds. eLife 2020;9:e60066.

Phan QM, Sinha S, Biernaskie J, Driskell RR. Single-cell transcriptomic analysis of small and large wounds reveals the distinct spatial organization of regenerative fibroblasts. Exp Dermatol 2021;30:92–101.

Plikus MV, Wang X, Sinha S, Forte E, Thompson SM, Herzog EL, et al. Fibroblasts origins, definitions, and functions in health and disease. Cell 2021;184:3852–72.

Polański K, Young MD, Miao Z, Meyer KB, Teichmann SA, Park JE. BBKNN: fast batch alignment of single cell transcriptomes. Bioinformatics 2020;36:964–5.

Qin Z, Worthen CA, Quan T. Cell-size-dependent upregulation of HGF expression in dermal fibroblasts: impact on human skin connective tissue aging. J Dermatol Sci 2017;88:289–97.

Qiu C, Cao J, Martin BK, Li T, Welsh IC, Srivatsan S, et al. Systematic reconstruction of cellular trajectories across mouse embryogenesis. Nat Genet 2022;54:328–41.

Sarkar A, Stephens M. Separating measurement and expression models clarifies confusion in single-cell RNA sequencing analysis. Nat Genet 2021;53:770–7.

Shao X, Taha IN, Clauser KR, Gao YT, Naba A. MatrisomeDB: the ECM-protein knowledge database. Nucleic Acids Res 2020;48:D1136–44.

Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, et al. Comprehensive integration of single-cell data. Cell 2019;177:1888–902.e21.

Stumm J, Vallecillo-García P, Vom Hofe-Schneider S, Ollitrault D, Schrewe H, Economides AN, et al. Odd skipped-related 1 (Osr1) identifies muscle-interstitial fibro-adipogenic progenitors (FAPs) activated by acute injury. Stem Cell Res 2018;32:8–16.

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 2018;138:802–10.

Telerman SB, Rognoni E, Sequeira I, Pisco AO, Lichtenberger BM, Culley OJ, et al. Dermal Blimp1 acts downstream of epidermal TGFβ and Wnt/β-catenin to regulate hair follicle formation and growth. J Invest Dermatol 2017;137:2270–81.

Vallecillo-García P, Orgeur M, vom Hofe-Schneider S, Stumm J, Kappert V, Ibrahim DM, et al. Odd skipped-related 1 identifies a population of embryonic fibro-adipogenic progenitors regulating myogenesis during limb development. Nat Commun 2017;8:1218.

Van de Sande B, Flerin C, Davie K, De Waegeneer M, Hulselmans G, Aibar S, et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nat Protoc 2020;15:2247–76.

Van den Berge K, Roux de Bézieux H, Street K, Saelens W, Cannoodt R, Saeys Y, et al. Trajectory-based differential expression analysis for single-cell sequencing data. Nat Commun 2020;11:1201.

Wei J, Ghosh AK, Sargent JL, Komura K, Wu M, Huang QQ, et al. PPARγ downregulation by TGFß in fibroblast and impaired expression and function in systemic sclerosis: a novel mechanism for progressive fibrogenesis. PLoS One 2010;5:e13778.

Wendisch D, Dietrich O, Mari T, von Stillfried S, Ibarra IL, Mittermaier M, et al. SARS-CoV-2 infection triggers profibrotic macrophage responses and lung fibrosis. Cell 2021;184:6243–61.e27.

Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol 2018;19:15.

Wolf FA, Hamey FK, Plass M, Solana J, Dahlin JS, Göttgens B, et al. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. Genome Biol 2019;20:59.

Wolock SL, Lopez R, Klein AM. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. Cell Syst 2019;8:281–91.e9.

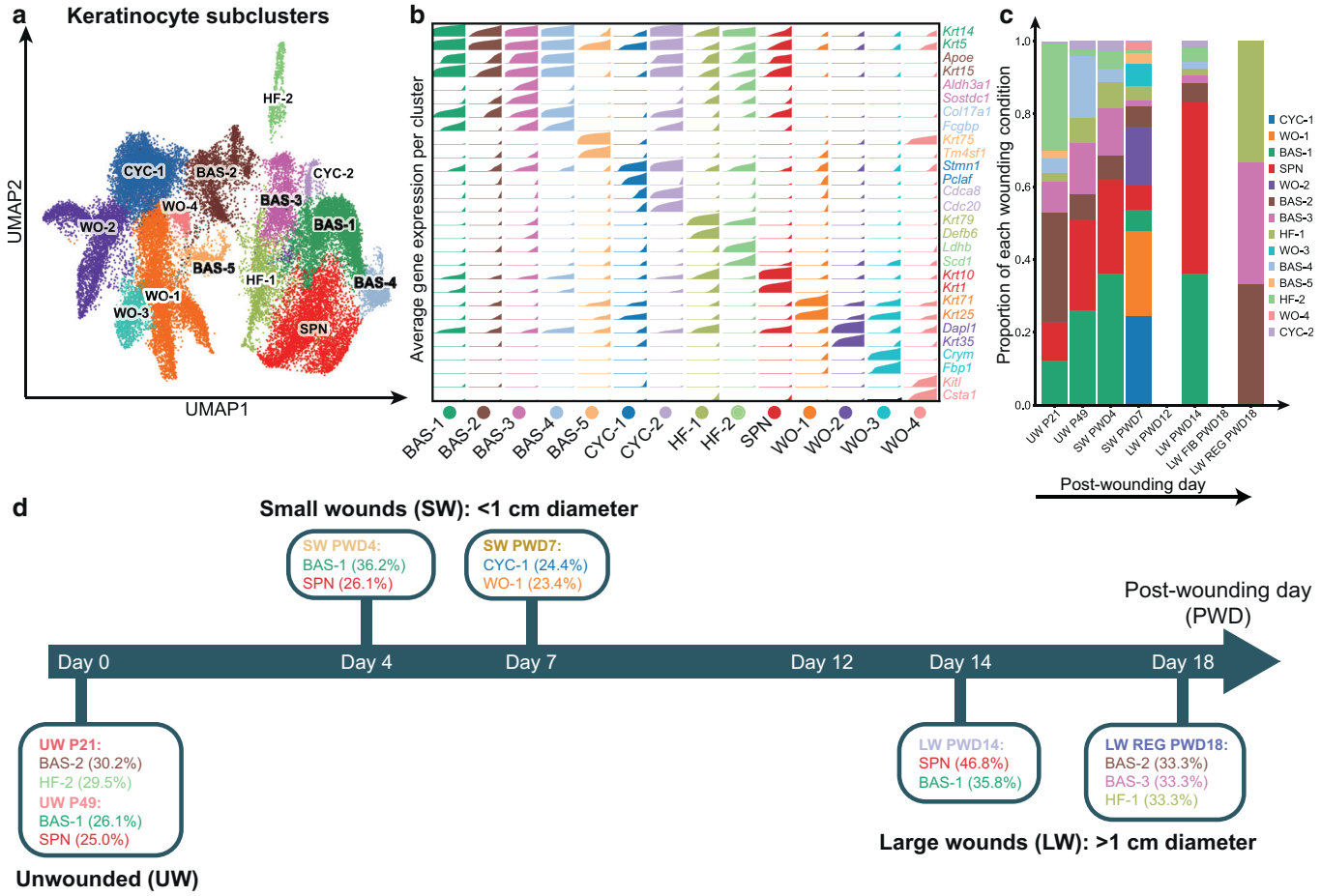
Xiao Z, Xi C. Hepatocyte growth factor reduces hypertrophy of skin scar: in vivo study. Adv Skin Wound Care 2013;26:266–70.

Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, et al. Gene set knowledge discovery with Enrichr. Curr Protoc 2021;1:e90.

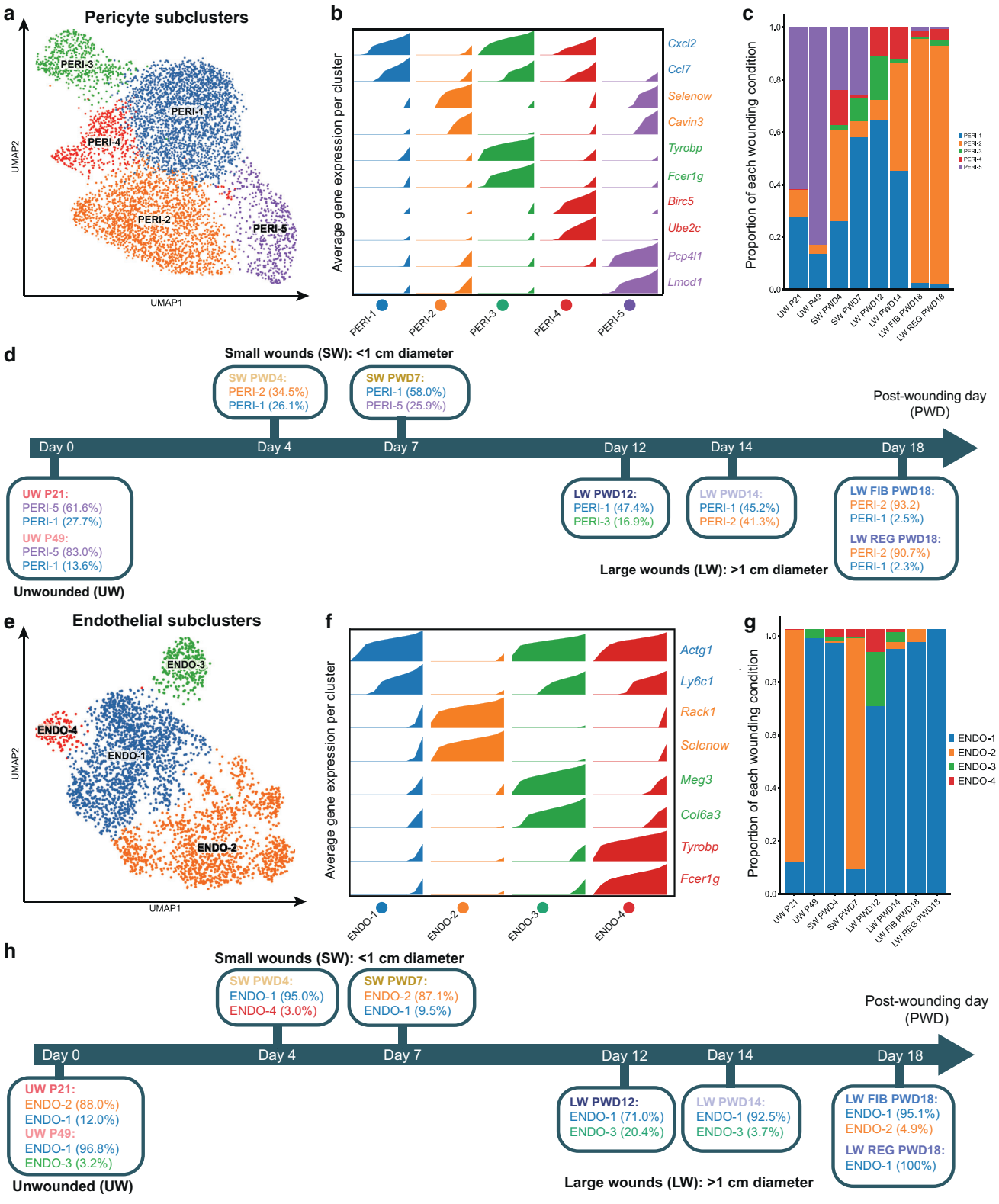
Yamazaki T, Nalbandian A, Uchida Y, Li W, Arnold TD, Kubota Y, et al. Tissue myeloid progenitors differentiate into pericytes through TGF-β signaling in developing skin vasculature. Cell Rep 2017;18:2991–3004.

Yokoyama Y, Sekiguchi A, Fujiwara C, Uchiyama A, Uehara A, Ogino S, et al. Inhibitory regulation of skin fibrosis in systemic sclerosis by apelin/APJ signaling. Arthritis Rheumatol 2018;70:1661–72.

Zhang T, He X, Caldwell L, Goru SK, Ulloa Severino LU, Tolosa MF, et al. NUAK1 promotes organ fibrosis via YAP and TGF-β/SMAD signaling. Sci Transl Med 2022;14:eaaz4028.

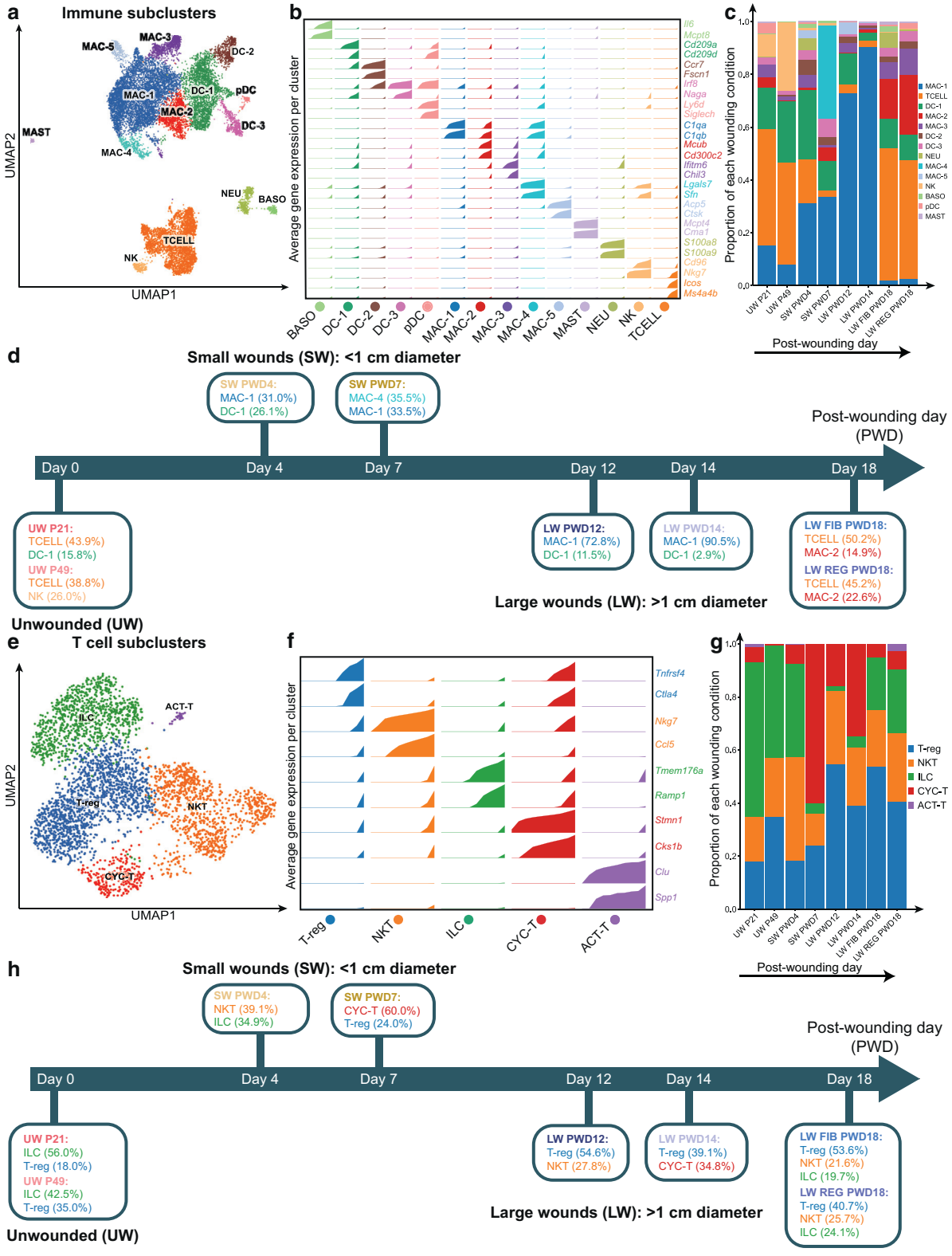


Supplementary Figure S1. Identification of keratinocyte subclusters in integrated dataset. (**a**) UMAP plot of the identified epidermal cell subclusters in integrated dataset shown in main Figure 1c. Cells are color coded and annotated as shown. (**b**) Tracks plot of the top 2 differentially expressed genes for each subcluster and their expressions across all keratinocytes. Color coding and annotation are according to **a**. (**c**) Relative epidermal cell composition of individual scRNA-seq datasets arranged according to study timeline from Figure 1a and annotated along the *x*-axis. Stacked bars show relative abundance of distinct epidermal cell subclusters. Bar segments are colored and annotated according to **a**. **(d**) Summary of epidermal cell subclusters at each time point along the biological timescale of skin wound healing, as originally defined in Figure 1a. UW denotes unwounded skin. LW, large wound; P21, postnatal day 21; P49, postnatal day 49; PWD, postwounding day; scRNA-seq, single-cell RNA-sequencing; SW, small wound; UMAP, Uniform Manifold Approximation and Projection.

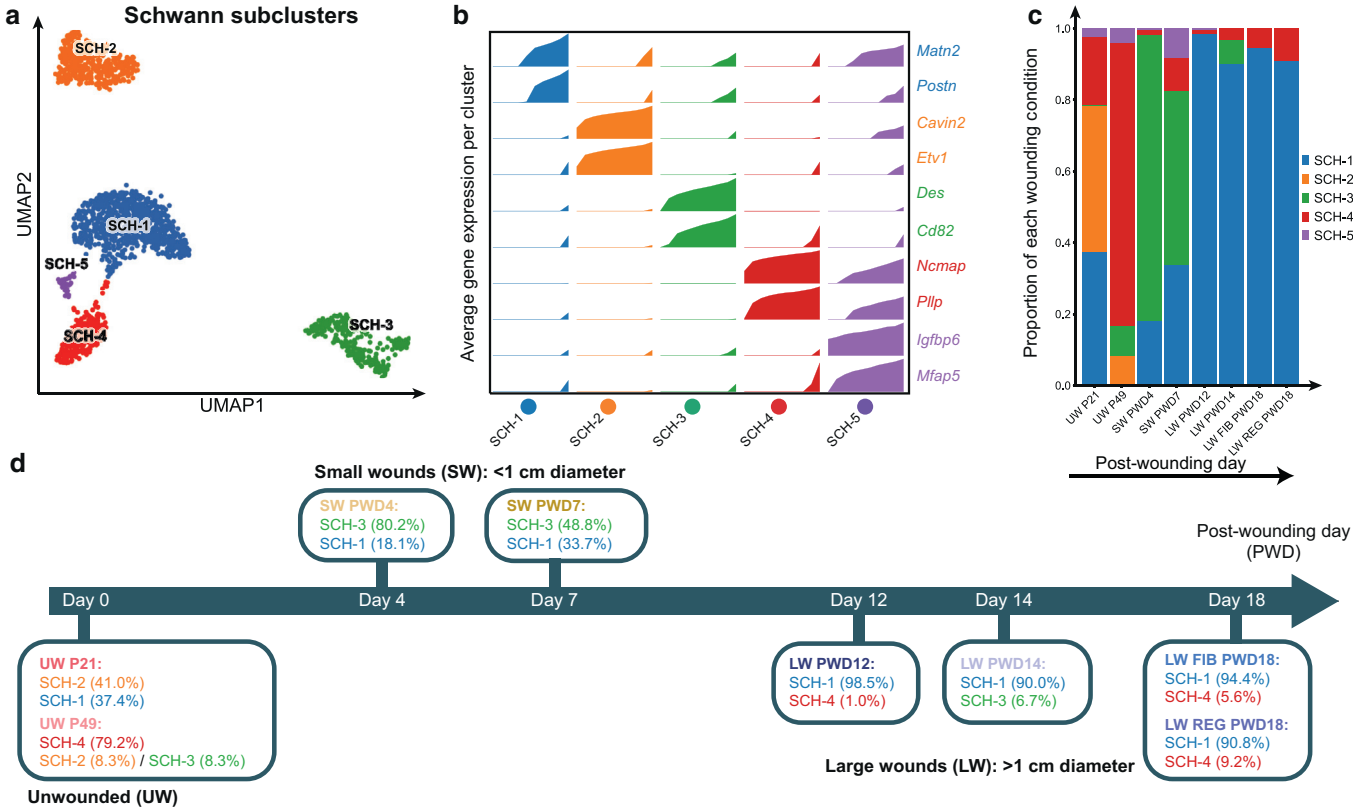


Supplementary Figure S2. Identification of dominant vascular cell lineage subclusters in integrated dataset. (**a**) UMAP plot of the identified pericyte subclusters in integrated dataset shown in main Figure 1c. Cells are color coded and annotated as shown. (**b**) Tracks plot of the top two differentially expressed genes for each subcluster and their expressions across all pericytes. Color coding and annotation are according to **a**. (**c**) Relative pericyte composition of individual scRNA-seq datasets arranged according to study timeline from Figure 1a and annotated along the *x*-axis. Stacked bars show relative abundance of distinct pericyte subclusters. Bar segments are colored and annotated according to **a**. (**d**) Summary of pericyte subclusters at each time point along the biological timescale of skin wound healing, as originally defined in Figure 1a. (**e**) UMAP plot of the identified endothelial cell subclusters in integrated dataset shown in main Figure 1c. Cells are color coded and annotated as shown. (**f**) Tracks plot of the top two differentially expressed genes for each subcluster and their expressions across all endothelial cells. Color coding and annotation are according to **e**. (**g**) Relative endothelial cell composition of individual scRNA-seq datasets arranged according to study timeline from Figure 1a and annotated along the *x*-axis. Stacked bars show relative abundance of distinct endothelial

subclusters. Bar segments are colored and annotated according to **e**. (**h**) Summary of endothelial subclusters at each time point along the biological timescale of skin wound healing, as originally defined in Figure 1a. UW denotes unwounded skin. LW, large wound; P21, postnatal day; P49, postnatal day 49; PWD, postwounding day; scRNA-seq, single-cell RNA-sequencing; SW, small wound; UMAP, Uniform Manifold Approximation and Projection.

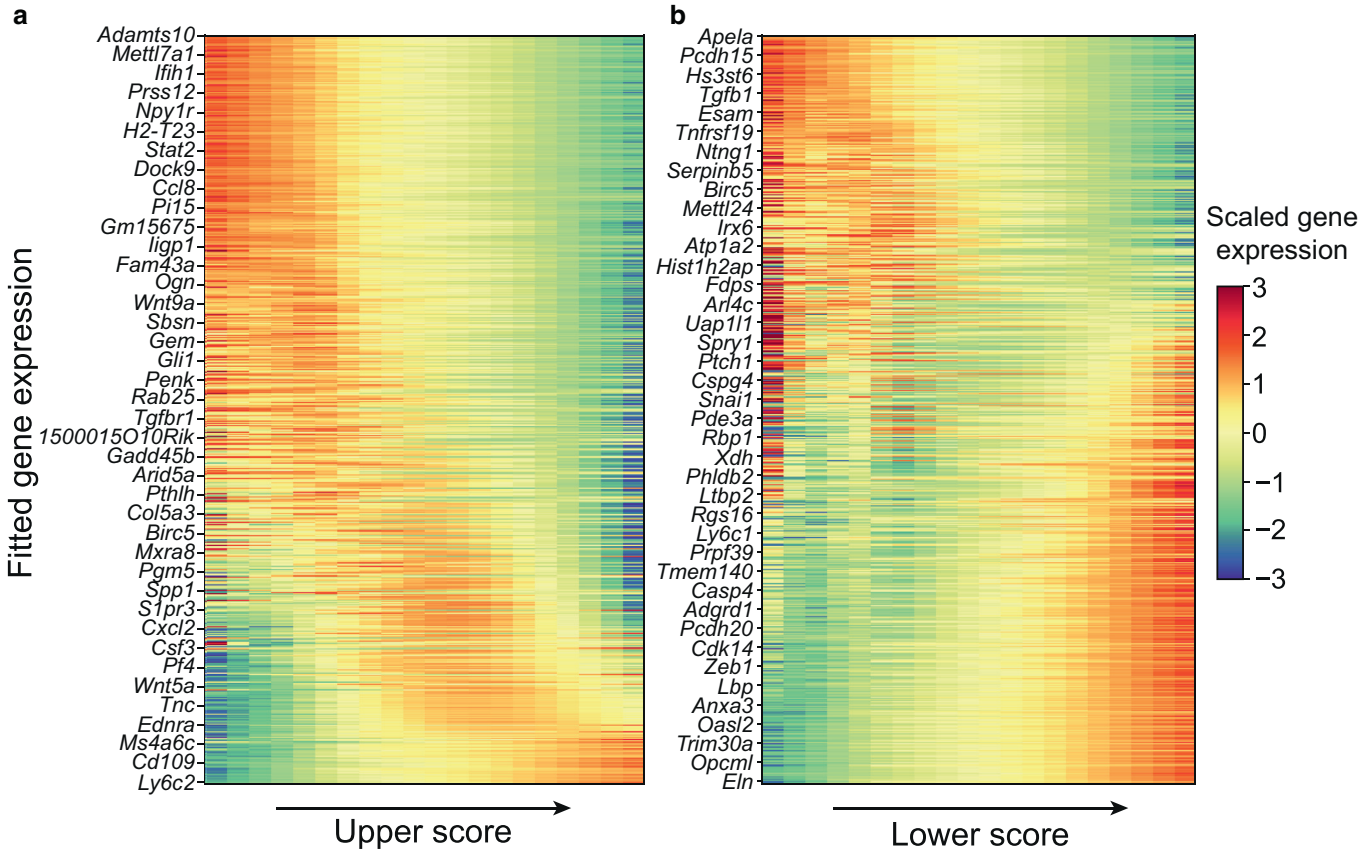


Supplementary Figure S3. Identification of dominant immune cell types in integrated dataset. (**a**) UMAP plot of the identified immune cell subclusters in integrated dataset shown in main Figure 1c. Cells are color coded and annotated as shown. (**b**) Tracks plot of the top 2 differentially expressed genes for each subcluster and their expressions across all dermis-resident immune cell types. Color coding and annotation are according to **a**. (**c**) Relative immune cell composition of individual scRNA-seq datasets arranged according to study timeline from Figure 1a and annotated along the *x*-axis. Stacked bars show relative abundance of distinct immune cell subclusters. Bar segments are colored and annotated according to **a**. (**d**) Summary of dominant immune cell subclusters at each time point along the biological timescale of skin wound healing, as originally defined in Figure 1a. (**e**) UMAP plot of the identified T-cell subclusters in integrated dataset shown in main Figure 1c. Cells are color coded and annotated as shown. (**f**) Tracks plot of the top two differentially expressed genes for each subcluster and their expressions across all T cells. Color coding and annotation are according to **e**. (**g**) Relative T-cell composition of individual scRNA-seq datasets arranged according to study timeline from Figure 1a and annotated along the *x*-axis. Stacked bars show relative abundance of distinct T-cell subclusters.

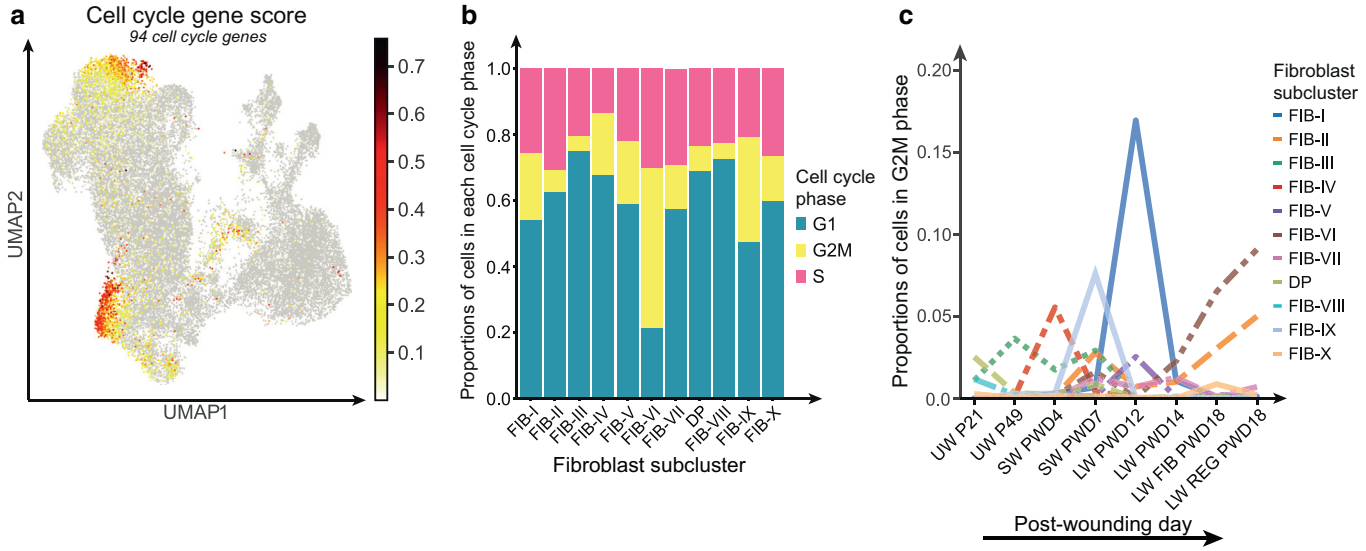


Supplementary Figure S4. Identification of Schwann cells present in integrated dataset. (**a**) UMAP plot of the identified Schwann cells subclusters in integrated dataset shown in main Figure 1c. Cells are color coded and annotated as shown. (**b**) Tracks plot of the top two differentially expressed genes for each subcluster and their expressions across all Schwann cell subclusters. Color coding and annotation are according to **a**. (**c**) Relative Schwann cell composition of individual scRNA-seq datasets arranged according to study timeline from Figure 1a and annotated along the *x*-axis. Stacked bars show relative abundance of distinct Schwann cell subclusters. Bar segments are colored and annotated according to **a**. (**d**) Summary of dominant Schwann cell subclusters at each time point along the biological timescale of skin wound healing, as originally defined in Figure 1a. UW denotes unwounded skin. LW, large wound; P21, postnatal day 21; P49, postnatal day 49; PWD, postwounding day; scRNA-seq, single-cell RNA-sequencing; UMAP, Uniform Manifold Approximation and Projection.

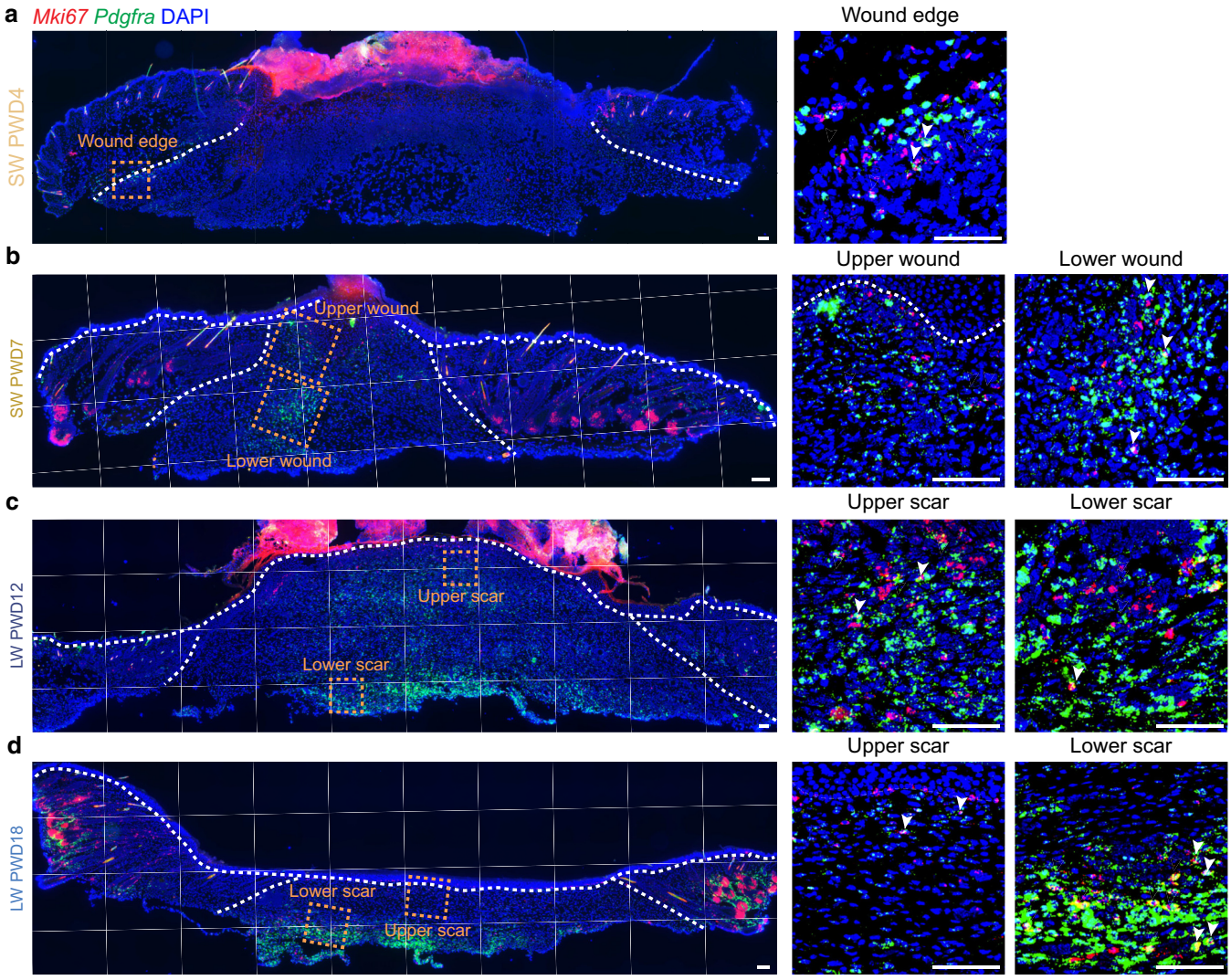
Bar segments are colored and annotated according to **e**. (**h**) Summary of T-cell subclusters at each time point along the biological timescale of skin wound healing, as originally defined in Figure 1a. UW denotes unwounded skin. ILC, innate lymphoid cell; LW, large wound; P21, postnatal day 21; P49, postnatal day 49; PWD, postwounding day; scRNA-seq, single-cell RNA-sequencing; T-reg, regulatory T cell; UMAP, Uniform Manifold Approximation and Projection.



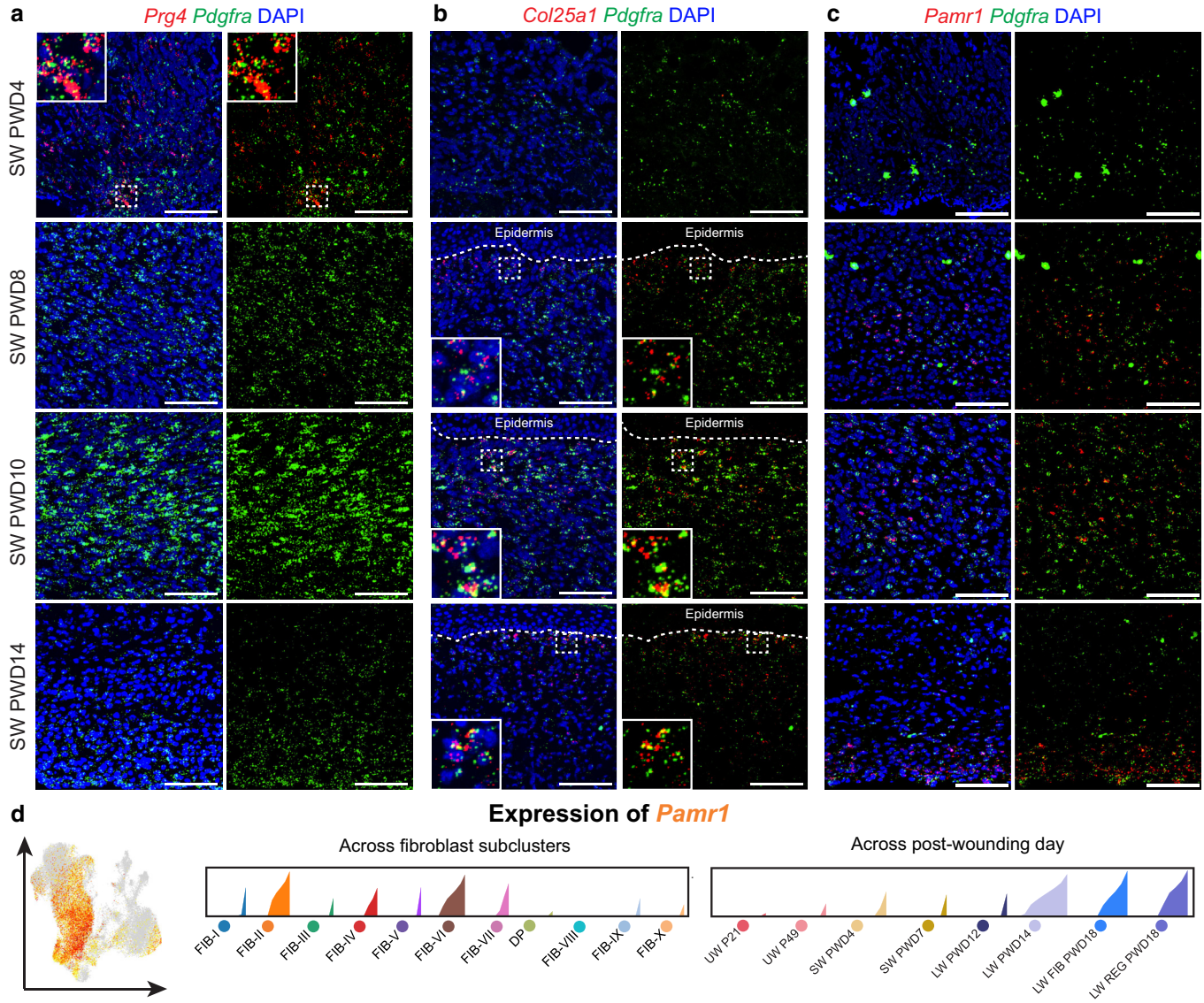
Supplementary Figure S5. Differential expression analysis of upper and lower score–enriched fibroblasts. We fit gene expression as continuous functions of the upper and lower fibroblast score separately. The fitted expressions were aggregated into 10 bins determined by the spatial identity scores. (**a**) Heatmap of the fitted expression of genes (from the top 2,000 highly variable genes) inferred to be significantly associated with the calculated upper fibroblast score. Genes are ordered by their Spearman correlation with the upper fibroblast score values (lowest correlation at the top, highest correlation at the bottom). For visualization purposes, we only labeled every 40 genes. We extracted the top five genes that correlated least with the upper score (top) and the top five genes that correlated most with the upper score (bottom). (**b**) Heatmap of the fitted expression of the highly variable genes inferred to be significantly associated with the calculated lower fibroblast score. Genes are ordered by their Spearman correlation with the lower fibroblast score values (lowest correlation at the top, highest correlation at the bottom). For visualization purposes, we only labeled every 40 genes. We extracted the top 5fivegenes that correlated least with the lower score (top) and the top genes that correlated most with the upper score (bottom), including those that were unwounded FIB-III–specific genes (labeled green).



Supplementary Figure S6. Fibroblast proliferative competence is maintained throughout wound healing time. (a) Feature plot of cell cycle gene score, as calculated from 94 genes, across all fibroblasts. Orange indicates gene score–expressing fibroblasts. (**b**) Relative proportion of fibroblasts in each major cell cycle phase (G1, G2M, S) across each fibroblast subcluster, annotated along the *x*-axis. Stacked bars are colored according to cell cycle phase, as defined on the right. (**c**) Proportions of fibroblasts in G2M phase across study timeline from Figure 1a and annotated along the *x*-axis. Graph lines are colored according to information in the right. UW denotes unwounded skin. LW, large wound; PWD, postwounding day; SW, small wound; UMAP, Uniform Manifold Approximation and Projection.

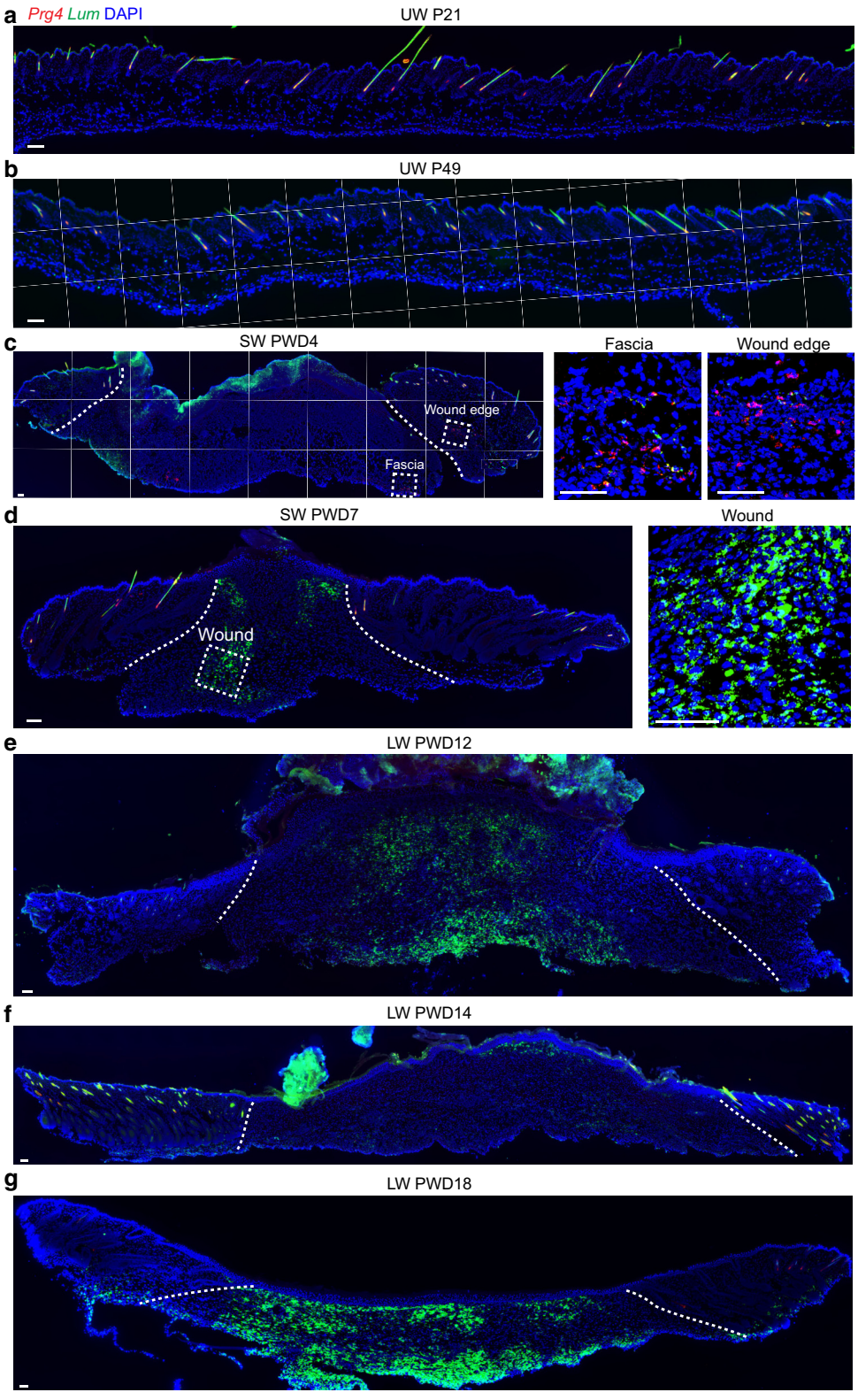


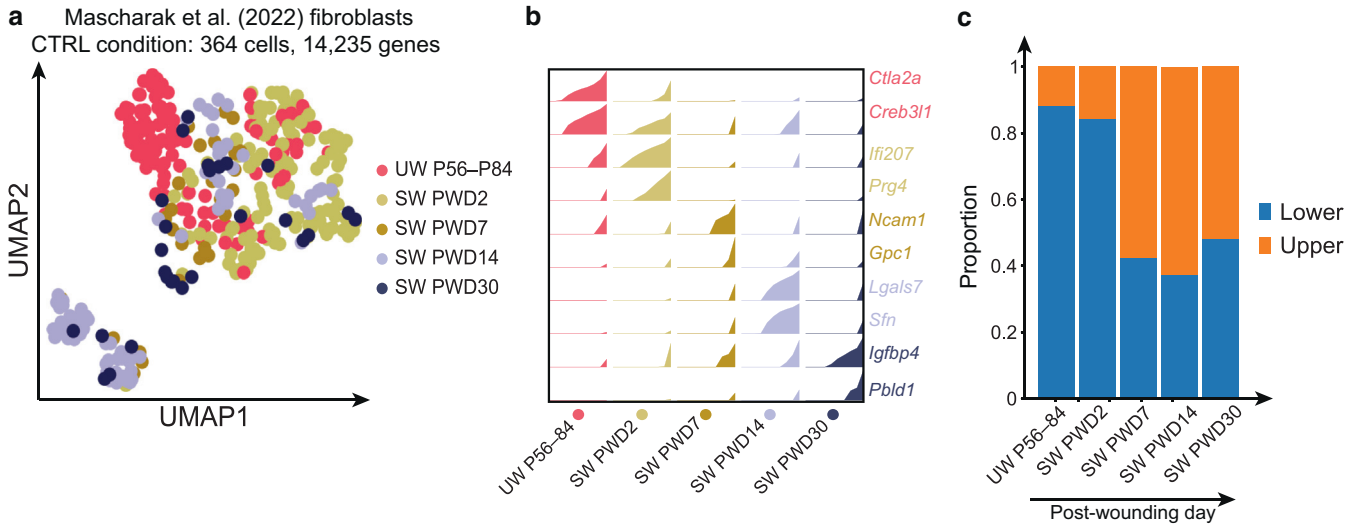
Supplementary Figure S7. Coexpression of proliferation marker *Mki67* and fibroblast marker *Pdgfra* across wound healing times. (**a**–**d**) In situ expression of *Pdgfra* (green) and proliferation marker *Mki67* (red) in the following wound samples: (**a**) SW PWD4, (**b**) SW PWD7, (**c**) LW PWD12, and (**d**) LW PWD18. Tissues were counterstained with DAPI (blue). White dashed lines mark original wound edges and/or epidermis. Several marked and annotated regions are magnified and shown on the right. Selected *Mki67*$^{+}$/*Pdgfra*$^{+}$ cells are marked with white arrowheads. Bars in **a**–**d** = 100 μm. LW, large wound; PWD, postwounding day; SW, small wound.

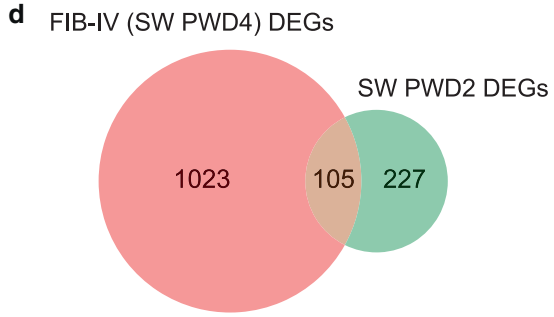


Supplementary Figure S8. Marker gene staining in 6-mm wound model confirms conserved fibroblast dynamic predicted by scRNA-seq analysis. (**a**–**c**) In situ expression of *Pdgfra* (green) and (**a**) early wound fibroblast marker *Prg4* (red), (**b**) mid-late upper scar fibroblast marker *Col25a1* (red), and (**c**) late lower scar fibroblast marker *Pamr1* (red) is shown in 6-mm wound tissues sampled at PWD4, 8, 10, and 14. All tissues were counterstained with DAPI (blue). In **a**–**b,** the white solid boxes show the magnified version of the regions marked by the white dashed boxes. White dashed lines in **b** mark the epidermis. Bars in **a**–**c** = 100 μm. (**d**) Left panel shows feature plot of *Pamr1*, a DEG for late wound-specific fibroblasts. Orange indicates expressing fibroblasts. Center panel shows track plots of *Pamr1* expression across fibroblast subclusters, annotated and color coded according to Figure 2a, whereas the right panel shows track plots of *Pamr1* expression across biological time points, annotated and color coded according to Figure 1b. UW denotes unwounded skin. DEG, differentially expressed gene; LW, large wound; P21, postnatal day 21; P49, postnatal day 49; PWD, postwounding day; scRNA-seq, single-cell RNA-sequencing; SW, small wound.

Supplementary Figure S9. Replication study shows transient *Prg4* expression in early wounds and lack thereof in unwounded skin and wounds at other time points. In situ expression of *Lum* (green) and *Prg4* (red) is shown in (**a, b**) unwounded skin samples and wound samples: (**c**) SW PWD4, (**d**) SW PWD7, (**e**) LW PWD12, (**f**) LW PWD14, and (**g**) LW PWD18. Tissues were counterstained with DAPI (blue). White dashed lines mark original wound edges. In **c**–**d,** magnified regions are shown on the right. No significant *Prg4* expression was detected in all samples shown. Bars in **a**–**g** = 100 μm. LW, large wound; PWD, postwounding day; SW, small wound.







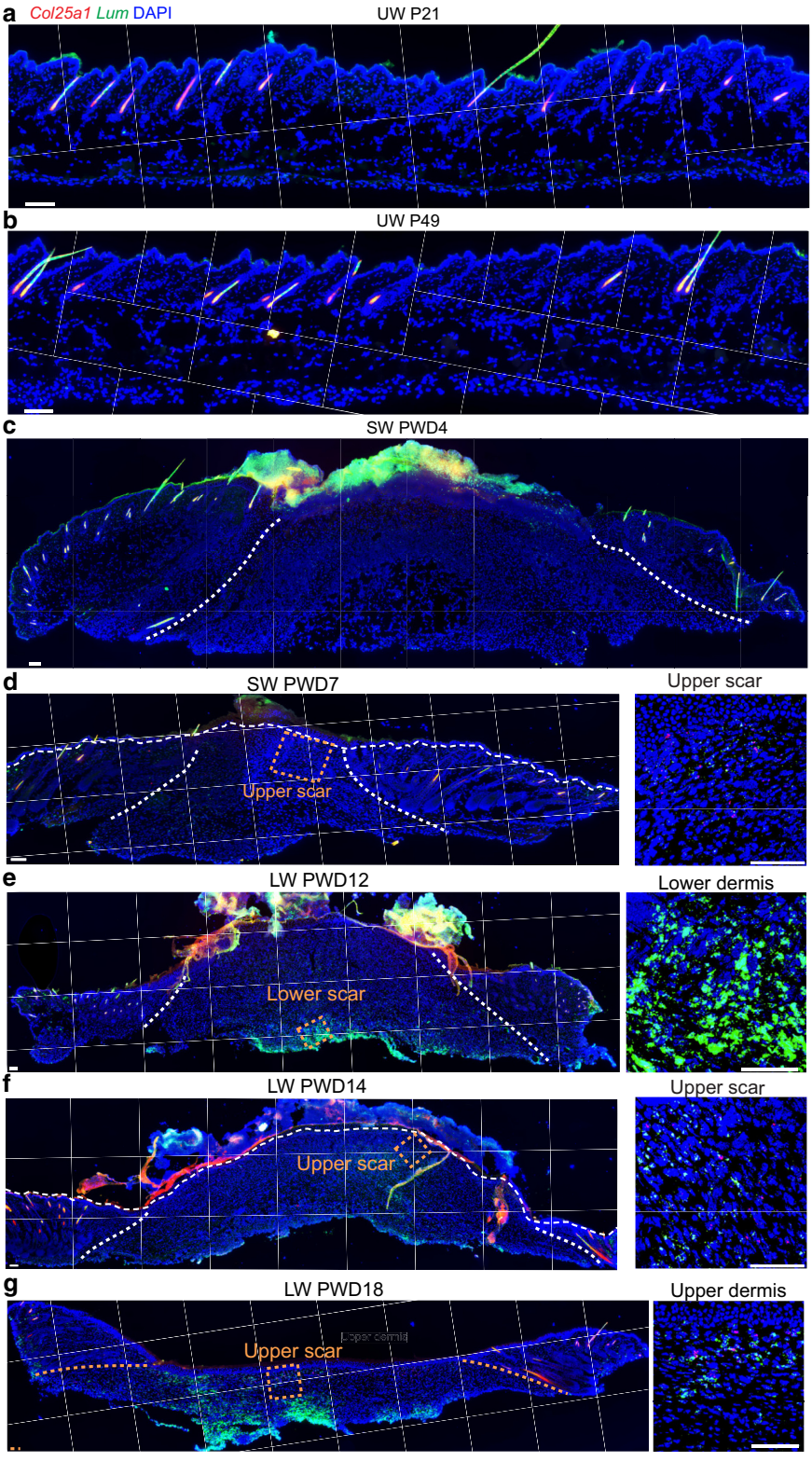
e

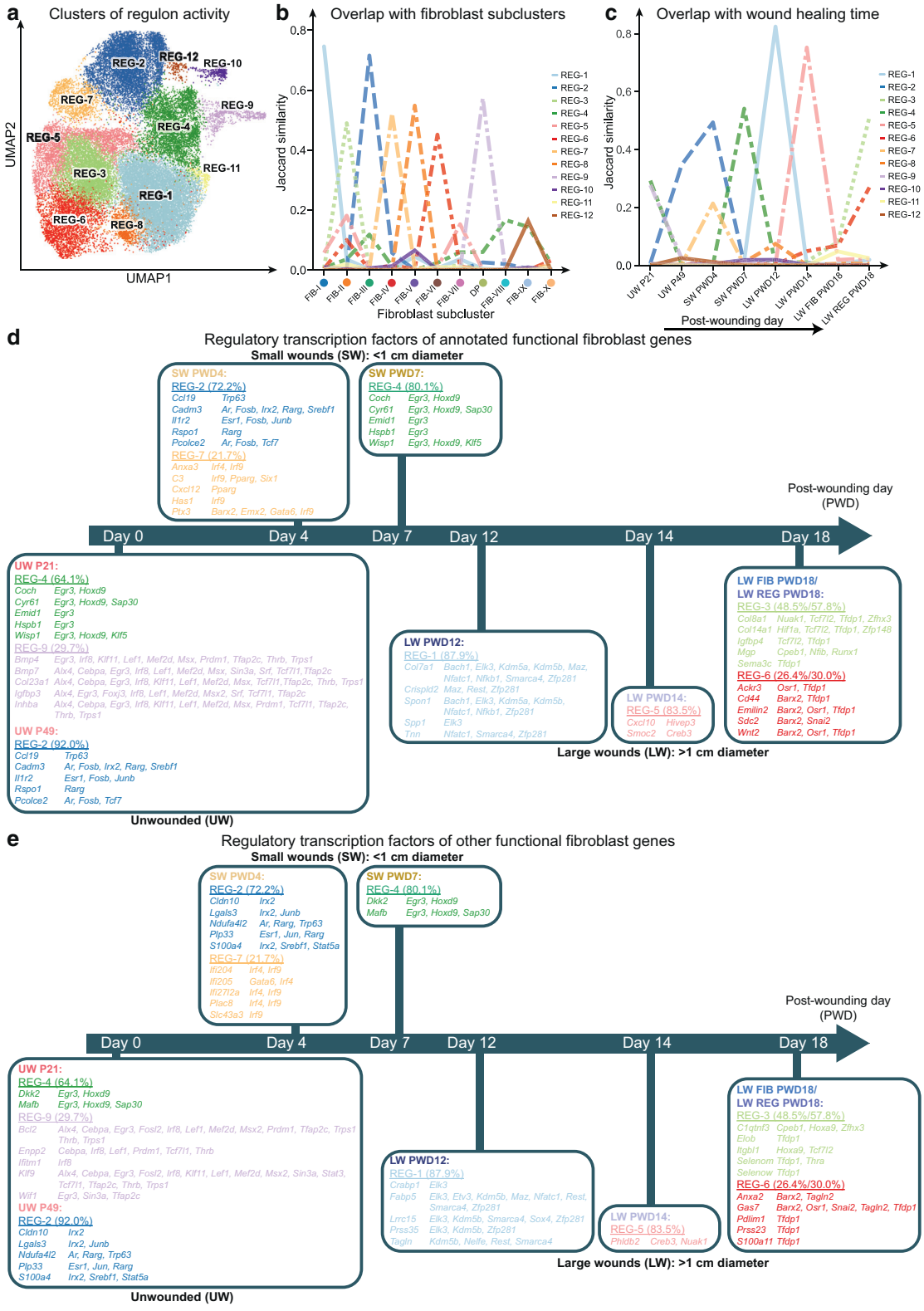
Top shared DEGs between FIB-IV (SW PWD4) and SW PWD2 fibroblasts

Cell growth	Chemokines	Inflammation	ECM synthesis
Cygb	*Ccl2*	*Ifi204*	*Gfpt2*
Plac8	*Ccl7*	*Ifi205*	*Has1*
	Cxcl1	*Il6*	*Prg4*
	Cxcl12	*Ptx3*	*Sod3*
		Tnfaip6	*Ugdh*

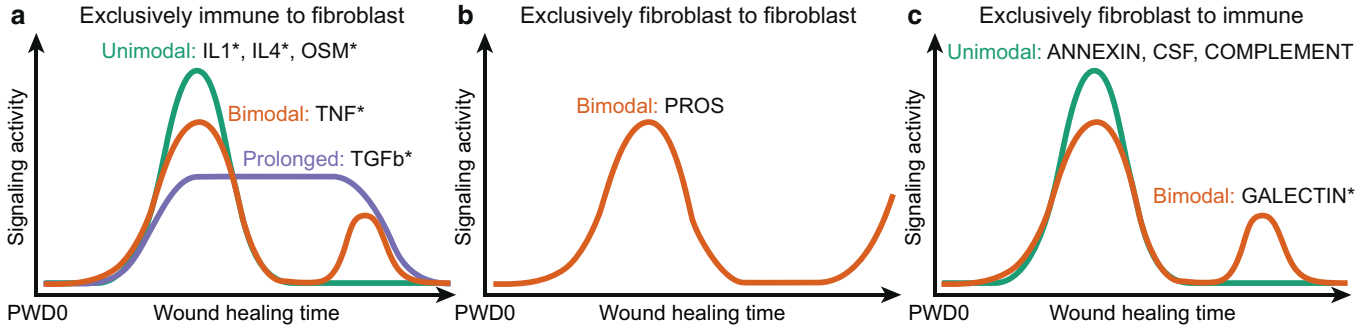
Supplementary Figure S10. Reanalysis of scRNA-seq data on fibroblasts sampled from small wounds from Mascharak et al (2022). (**a**) UMAP showing the neighborhood embedding of the fibroblast population labeled by wound healing time. UW denotes unwounded skin. (**b**) Tracks plots of the top two DEGs in fibroblasts for each sample from **a**. *Prg4* is enriched in wound fibroblasts derived from SW PWD2. (**c**) Composition of upper-enriched (orange) versus lower-enriched fibroblasts (blue) for each sample from **a**. (**d**) Overlap between DEGs of early wound fibroblast subcluster FIB-IV from main Figure 2a and DEGs of *Prg4*$^{+}$ fibroblasts from SW PWD2 from **a**. (**e**) Summary of the top DEGs in the intersection between original FIB-IV wound fibroblasts from main Figure 2a and *Prg4*$^{+}$ SW PWD2 wound fibroblasts from **a**. UW denotes unwounded skin. DEG, differentially expressed gene; P56, postnatal day 56; P84, postnatal day 84; PWD, postwounding day; scRNA-seq, single-cell RNA-sequencing; SW, small wound; UMAP, Uniform Manifold Approximation and Projection.

Supplementary Figure S11. Replication study shows the expression of mid-late wound fibroblast marker gene *Col25a1 in* upper scar and lack thereof in unwounded skin and wounds at other time points. In situ expressions of *Lum* (green) and *Col25a1* (red) are shown in (**a, b**) unwounded skin samples and wound samples: (**c**) SW PWD4, (**d**) SW PWD7, LW PWD12 (**e**), (**f**) LW PWD14, (**g**) LW PWD18. Tissues were counterstained with DAPI (blue). White dashed lines mark original wound edges. In **d**–**g**, magnified regions are shown on the right, displaying significant *Col25a1* expression. Bars in **a**–**e** = 100 μm. LW, large wound; P21, postnatal day 21; P49, postnatal day 49; PWD, postwounding day; SW, small wound.





Supplementary Figure S12. pySCENIC-inferred epigenetic regulation of fibroblast states. (**a**) UMAP plot shows fibroblasts clustered on the basis of regulon activity, which group cells with similar inferred epigenetic states. All fibroblast clusters are color coded and annotated as shown. (**b, c**) Jaccard similarities between the (**b**) regulon-based clusters and the original fibroblast subclusters as well as (**c**) regulon-based clusters and wound healing times. These data indicate significant overlap between fibroblast states and additional epigenetic states due to wound healing time. (**d, e**) Summary of the dominant fibroblast regulon states and (**d**) the inferred transcriptional factors that regulate fibroblast-specific functional genes and (**e**) nonfibroblast-specific functional genes over wound healing time. UW denotes unwounded skin. LW, large wound; PWD, postwounding day; SW, small wound; UMAP, Uniform Manifold Approximation and Projection.



Supplementary Figure S13. Summary of signaling interactions between dominant immune cells and fibroblast states over wound healing time. (**a**) Summary of wound healing time–dependent signaling patterns for pathways that operate from wound-resident immune cells to fibroblasts. (**b**) Summary of wound healing time–dependent signaling patterns for pathways that operate between fibroblasts only. (**c**) Summary of wound healing time–dependent signaling patterns for pathways that operate from fibroblasts to wound-resident immune cells. In **a**–**c**, signaling pathways are arranged and color coded according to their time-dependent pattern: unimodal (ie, used once and time limited; green), prolonged (ie, used for extended time; purple), and bimodal (ie, use twice at distinct times of wound healing; red). Pathways marked with an asterisk are predicted to be mitogenic. PWD, postwounding day.

Supplementary Table S1. List of Upper and Lower Fibroblast Marker Genes and Associated Literature Evidence

Fibroblast Marker	Position (Upper/Lower)	Sample Type	Reference
Cnn1	Lower	Unwounded	Korosec et al (2019)
Crabp1	Upper	Wounded	Abbasi et al (2020)
Dlk1	Lower	Wounded	Abbasi et al (2020)
Dpp4	Upper	Unwounded	Driskell et al(2013) and Tabib et al (2018)
Fabp5	Upper	Wounded	Abbasi et al (2020)
Fmo1	Lower	Unwounded	Tabib et al (2018)
Lef1	Upper	Unwounded	Phan et al (2020)
Ly6a	Lower	Wounded	Abbasi et al (2020) and Driskell et al (2013)
Mest	Lower	Wounded	Abbasi et al (2020)
Prdm1	Upper	Unwounded	Driskell et al (2013)
Prss35	Upper	Wounded	Abbasi et al (2020)
Runx1	Upper	Wounded	Abbasi et al (2020)