Defining Transcriptional Signatures of Human Hair Follicle Cell States

Rie Takahashi[1,8], Adrienne Grzenda[2,8], Thomas F. Allison[3,8], Jeffrey Rawnsley[4], Samuel J. Balin[1], Shan Sabri[5], Kathrin Plath[3,5,6] and William E. Lowry[1,5,6,7]

The epidermis and its appendage, the hair follicle, represent an elegant developmental system in which cells are replenished with regularity because of controlled proliferation, lineage specification, and terminal differentiation. Although transcriptome data exists for human epidermal and dermal cells, the hair follicle remains poorly characterized. Through single-cell resolution profiling of the epidermis and anagen hair follicle, we characterized the anatomical, transcriptional, functional, and pathological profiles of distinct epidermal, hair follicle, and hair follicle–associated cell subpopulations including melanocytes, endothelial cells, and immune cells. We additionally traced the differentiation trajectory of interfollicular and matrix cell progenitors and explored the association of specific cell subpopulations to known molecular signatures of common skin conditions. These data simultaneously corroborate prior murine and human studies while offering new insights into epidermal and hair follicle differentiation and pathogenesis.

Journal of Investigative Dermatology (2020) **140,** 764–773; doi:10.1016/j.jid.2019.07.726

INTRODUCTION

Skin has been the subject of extensive characterization at the pathological and molecular level in humans and mice for decades (Fuchs, 1998-1999). During that time, numerous cell types within the epidermis and its appendages, the hair follicle and sweat glands, have been identified via immunohistochemical and array-based bulk gene expression analyses in each species. The characterization of epidermal, dermal, and hair follicle cells at a single-cell resolution has lagged compared with other tissues, largely owing to limited sample sizes and the limitations of early single-cell technologies. Previous single-cell transcriptional methods (e.g., Fluidigm) required labor-intensive manual sorting of cells for array hybridization or sequencing, which necessitated significant sample sizes to overcome processing losses for reliable downstream application. Newer methods (e.g., Drop-seq and 10X Genomics), however, employ multiplex barcoding and microfluidics that enable the rapid parallel processing of thousands of cells at a lower cost with significantly improved sensitivity (Klein and Macosko, 2017; Picelli, 2017; Ziegenhain et al., 2017).

The first single-cell mapping of murine epidermal and hair follicle cells yielded 1,422 unique transcriptome profiles, clustering into 25 unique subpopulations (Acosta et al., 2017; Joost et al., 2016). These human studies focused primarily on elucidation of interfollicular epidermal and dermal cell types and differential transcriptional programs activated by pathological processes such as psoriasis and wound healing (Cheng et al., 2018; Philippeos et al., 2018). Comparable characterization of human hair follicle cell types and differentiation states is currently lacking, as profiling of human hair follicles is difficult because of their relative scarcity within the skin compared with mice.

In this study, we gained access to follicule-enriched fractions of human skin that were discarded from hair transplant procedures and subjected them to high-throughput single-cell RNA (scRNA) procedures to gain single-cell transcriptomes for many of the cells associated with human follicles. This allowed for the generation of single-cell transcriptomes of numerous cell states within the follicle as well as epidermal keratinocytes, endothelial cells, mesenchymal populations, immune cells, and melanocytes. With these data, cell fate trajectories can be generated, and patterns of gene expression in skin diseases can be probed to find individual cell types that are targets of disease.

RESULTS

Single-cell profiling of follicle-enriched human skin grafts

Prior human single-cell investigations used fractions from total epidermis, resulting in a high ratio of epidermal to follicular cells (Cheng et al., 2018). To improve single-cell resolution of follicular cell types, anagen hair follicles were obtained from discarded human scalp micrografts collected for transplantation (Figure 1a). Micrografts are composed of

[1]*Division of Dermatology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California;* [2]*Department of Psychiatry, Semel Institute for Neuroscience and Behavior, University of California, Los Angeles, Los Angeles, California;* [3]*Department of Biological Chemistry, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California;* [4]*Department of Plastic Surgery, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California;* [5]*Molecular Biology Institute, University of California, Los Angeles, Los Angeles, California;* [6]*Eli and Edythe Broad Center for Regenerative Medicine, University of California, Los Angeles, Los Angeles, California; and* [7]*Department of Molecular Cell and Developmental Biology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California*

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

Correspondence: William E. Lowry, Department of Molecular Cell and Developmental Biology, David Geffen School of Medicine, University of California Los Angles, 621 Charles Young Drive, Los Angeles, California 90095. E-mail: blowry@ucla.edu

Abbreviations: DEG, differentially expressed gene; IFE, interfollicular epidermis; IRS, inner root sheath; scRNA, single-cell RNA

Received 29 May 2018; revised 17 July 2019; accepted 26 July 2019; accepted manuscript published online 31 October 2019; corrected proof published online 19 December 2019

several hair follicles, and variable amounts of surrounding tissues, including interfollicular epidermis, dermis, and sebaceous and/or apocrine glands, among others. Micrografts were collected from five patients for single-cell analysis and immunohistochemistry. Anagen phase was confirmed by hematoxylin and eosin–labeled sections of adjacent graft samples. Following enzymatic and physical dissociation, single-cell suspensions underwent fluorescence-activated cell sorting to remove dead cells and debris. Libraries were generated and sequenced using the Drop-seq ($n = 2$) and commercial 10X Genomics ($n = 2$) platforms (Klein and Macosko, 2017; Weisenfeld et al., 2017). The 10X method identified more transcripts and genes than Drop-seq (Supplementary Figure S1). Dataset integration was accomplishing using the Seurat integration method (Butler et al., 2018).

Unsupervised clustering and cell-type identification

Unsupervised, graph-based clustering revealed 23 primary clusters of cells, visualized by t-distributed stochastic neighbor embedding (TSNE) (Figure 1b). Overlaid t-distributed stochastic neighbor embedding projections of the Drop-seq and 10X samples and calculated cell counts from each technique were compared (Supplementary Figure S1c and d). Differentially expressed genes (DEGs) between clusters, expression values of keratin isoforms, and *MKI67* expression were employed to identify the individual subpopulations. The Gene Expression Deconvolution Interactive Tool additionally was employed for cell-type identification using the Skin Signatures database to assess cluster DEGs (Supplementary Figure S2a) (Swindell et al., 2013).

Hierarchical clustering of significant DEGs (log fold change > 1, adjusted P-value < 0.01) showed that distinct transcriptional profiles define each cell type (Figure 1c). Most of the dataset (49.9%, Figure 1d) consisted of hair follicle subpopulations present during anagen (e.g., bulge, lower bulge, outer root sheath, inner root sheath [IRS], matrix, medulla, cortex, isthmus, and infundibulum), indicative of a successful enrichment procedure. Note that lower bulge, matrix, medulla, and cortex are not present in telogen follicles, confirming that the follicles analyzed were indeed in the anagen stage. Interfollicular epidermal cells constituted the next most prevalent cell type (39%), and T cells, Langerhans cells, endothelial cells, apocrine and eccrine gland cells, melanocytes, and sebocytes were additionally identified (~11.1%). One cluster appeared to represent mesenchymal cell types, potentially including dermal papillae, dermal sheath, smooth muscle, and fibroblasts (Supplementary Figure S2b and c), but the cell number and the diversity within this cluster precluded definitive judgement. Highly differentially expressed keratin and nonkeratin gene expression profiles are shown in Figure 2a–c to demonstrate specificity of expression. The epidermal transcriptional profiles largely reproduced the findings of Cheng et al. (2018). The full list of identified cell-type makers are presented in Supplementary Table S1.

Immunostaining to confirm scRNA sequencing–identified expression patterns

Melanocyte and Langerhans cell subpopulations were confirmed by langerin and CD74 labeling, respectively, in hair follicle sections (Figure 3a). To confirm the utility of scRNA sequencing analysis in identifying previously undescribed subpopulation markers, we looked for enrichment of immunostaining with markers such as *CXCL14* through immunostaining of hair follicle sections alongside known bulge marker *CD200*, demonstrating colocalization (Figure 3b). Gene ontology analysis indicated enrichment of extracellular matrix protein synthesis associated with the bulge (Supplementary Figure S3), confirmed by colocalization of extracellular matrix markers *TNC* and *EFEMP1* (Figure 3b and lower magnification images in Supplementary Figure S4a). These three secreted extracellular matrix components were indeed found at the protein level in the bulge compartment of the follicle in a pattern suggestive of their secretion. Although these markers were not completely exclusive to the bulge in either the RNA sequencing or immunostaining analyses, the presence of these proteins in the bulge was confirmatory for the scRNA sequencing procedures.

We additionally analyzed the expression patterns of a subset of other putative markers identified by cluster differential gene expression analysis using an independently derived resource, the Human Tissue Atlas (Uhlén et al., 2015), demonstrating the particular fidelity of this subset for hair follicle layers (Figure 3c, Supplementary Figure S4b). Specifically, *DAPL1* was predicted to be expressed in the cortex and medulla, *DCD* in the sweat glands, *DSC1* in the IRS, *DSG4* in the cortex and medulla, *CD74* in Langerhans cells, *ELOVL5* in sebaceous glands, *FABP9* in the IRS, *S100A3* in the cuticle, and *CD59* in melanocytes.

Lineage trajectory of epidermal and follicular cell progenitors

To examine the differentiation of the interfollicular epidermis (IFE), IFE cells were placed in pseudotemporal order (Figure 4a). Cells deriving from the basal IFE were predicted from this analysis to differentiate into the spinous and later granular layers, as would be predicted by decades of research on both human and murine skin (Figure 4b). Consistent with what was reported in Cheng et al. (2018), we detected a mitotically active subset of basal IFE cells that do not follow the traditional IFE differentiation pattern. As a result, we are confident in both the data and the methods employed to generate lineage trajectories. Also, as predicted, progressive loss of basal markers *KRT5*, *KRT14*, and *COL17A1* is observed coincident with a gain of spinous marker *KRT10* and granular marker *CALML5* expression as differentiation progresses (Figure 4c). The top genes driving differentiation are shown in Figure 4d.

Probing for lineage trajectories in the hair follicle, we ordered data from the cortex/medulla/matrix, IRS Huxley's and Henle's layers, and outer root sheath companion layer (Figure 5a–d). When placed in pseudotemporal order, the cortex/medulla/matrix population branches off, terminating in outer root sheath derived companion layer cells as well as a further differentiated subset of cells, likely representing hair shaft medulla and cortex components. Additionally, the cortex/medulla/matrix subpopulation additionally gives rise to the IRS Huxley's and Henle's layers. These data are consistent with a prior report by Mesler et al. (2017), in which early matrix progenitors give rise to the companion layer and later

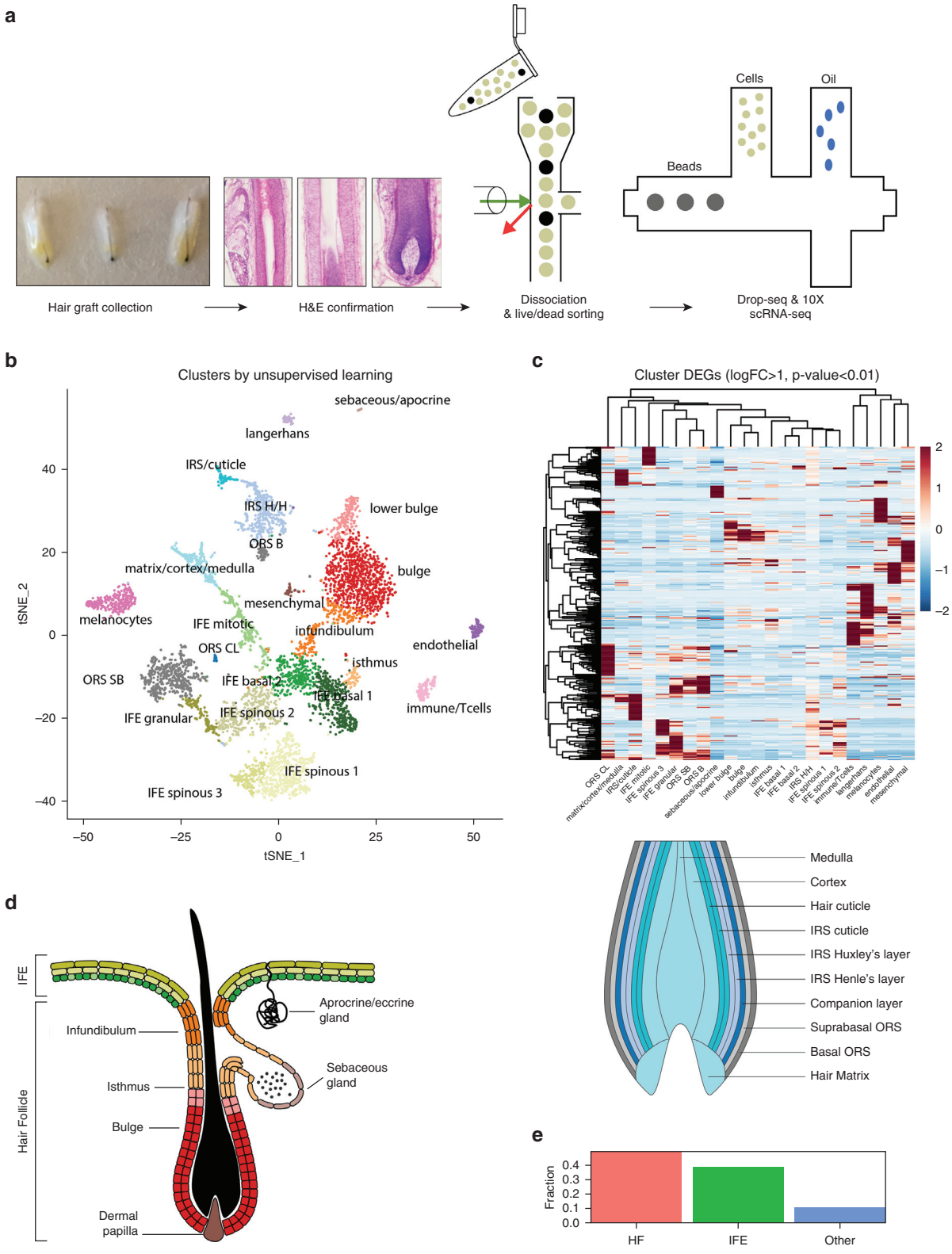


Figure 1. Identification of unique cell types from follicular-enriched scalp grafts by scRNA-seq. (**a**) Experimental workflow schematic. Grafts were collected from human scalp, examined microscopically to confirm anagen phase, dissociated, then sorted to remove dead cells and other debris before Drop-seq or 10X scRNA-seq. (**b**) t-SNE visualization of clusters generated by unsupervised, graph-based clustering of the integrated dataset. (**c**) heatmap of DEGs that define individual subtypes of cells from the analysis (**d**) (left) Illustration of the primary epidermal and follicular compartments present during telogen identified; (right) depiction of the layers and cell types found in the bulb of anagen follicle (bulb not present during telogen). (**e**) Percentage of epidermal, follicular, and other cell types identified. DEG, differentially expressed gene; FC, fold change; H&E, hematoxylin and eosin; HF, hair follicle; IFE, interfollicular epidermis; IRS, inner root sheath; IRS H/H, inner root sheath Huxley's/Henley's layers; ORS CL, outer root sheath companion layer; ORS SB/S, outer root sheath suprabasal/basal layer; scRNA-seq, single-cell RNA sequencing; t-SNE, t-distributed stochastic neighbor embedding.

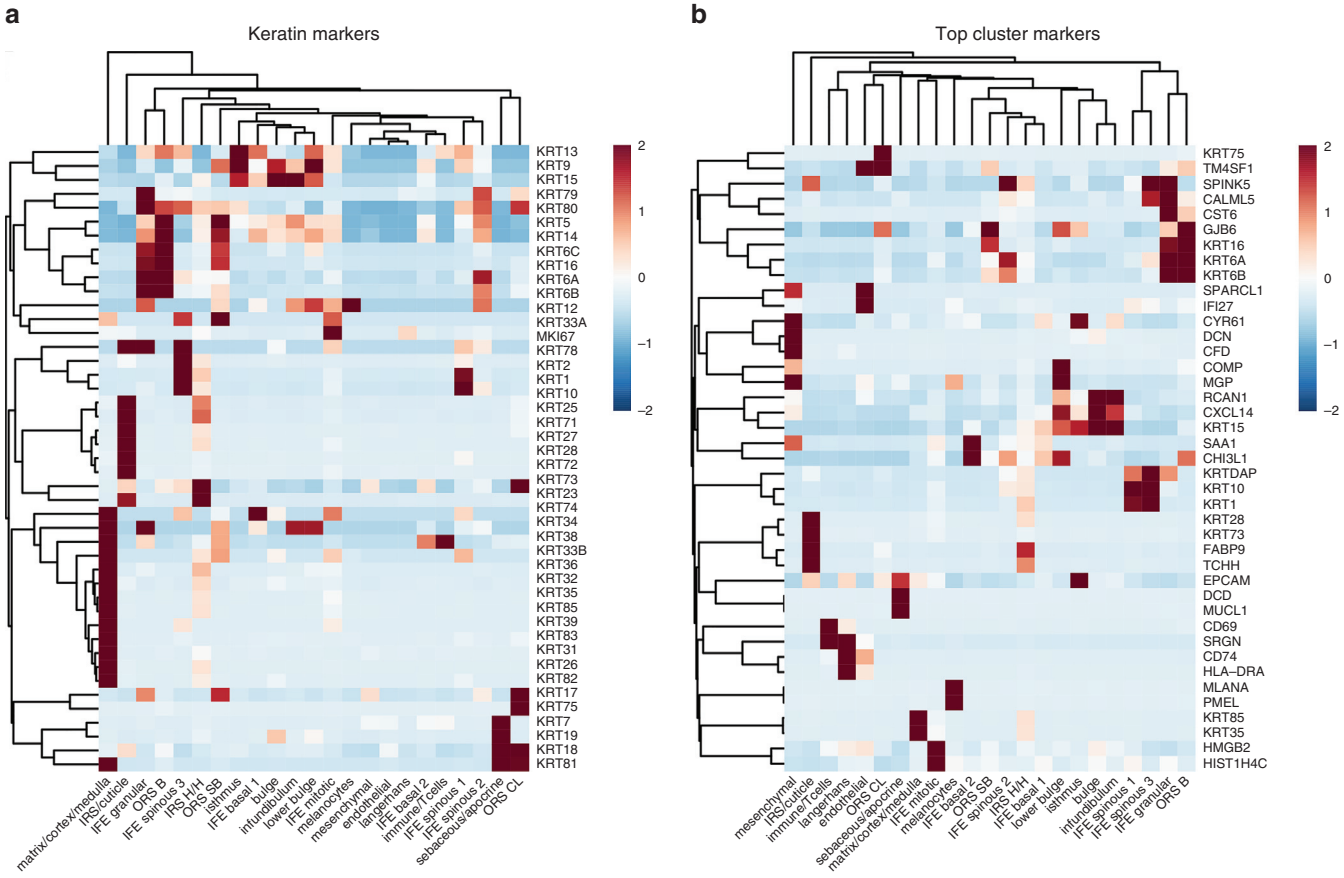


Figure 2. Cell-type identification and gene expression profiles. (**a**) Clustered Heatmap of keratin genes differentially expressed at a logFC > 1, adjusted *P*-value < 0.01, by cluster. (**b**) Clustered heatmap of the top cluster-specific markers of each cluster by FC. (**c**) t-SNE visualization of known and previously undescribed markers of cell type. DP, dermal papillae; DS, dermal sheath; FC, fold change; IFE, interfollicular epidermis; IRS H/H, inner root sheath Huxley's/Henley's layers; ORS CL, outer root sheath companion layer; ORS SB/S, outer root sheath suprabasal/basal layer; t-SNE, t-distributed stochastic neighbor embedding.

matrix progenitors give rise to the IRS and lower hair shaft components (Mesler et al., 2017).

Cell-type enrichment of pathological signatures

Another potential utility of these datasets is to link disease states to particular cell types in human skin. Numerous databases have been generated to describe the gene expression patterns associated with the onset or progression of skin diseases. In particular, DermDB2 (http://chlamy.mcdb.ucla.edu/NewDermDB/NewDermDB.html) provided a deep resource of data on gene expression patterns induced by various injuries, skin infections, and cancers. The limitation of these databases is the fact that the analyses were performed on bulk tissue samples, so information about the effect on particular cell types is obscured. Here, we probed the expression of groups of genes associated with various skin syndromes across our scRNA datasets. As shown in Figure 6, 14 disease or injury states showed distinct patterns of expression in each of the clusters identified within the scRNA datasets. Of note, some of the diseases known to be particularly related to immune responses showed a strong pattern of regulation in Langerhans and immune and/or T cells (leprosy, Stevens–Johnson Syndrome, and mycosis fungoides). As expected, the three cancer expression patterns were most linked to cell types that have been described pathologically to be most similar to the indicated cancer type (IFE granular for squamous cell carcinoma; bulge, outer root sheath, and matrix for basal cell carcinoma; and melanocytes for melanoma). These results demonstrate that gene expression patterns of skin diseases can be linked to particular target cells through correlational analysis with scRNA sequencing.

DISCUSSION

Our use of follicle-enriched samples permitted single-cell deconvolution of the transcriptomes of several layers of the human interfollicular epidermis, infundibulum, and hair follicle, as well as a wide variety of follicle-associated cells (e.g., immune, glandular, and pigment). Furthermore, we demonstrated the feasibility of integrated datasets generated from different technologies while maintaining clear cell-type resolutions. We utilized a multidimensional identification and validation strategy, examining known keratin and proliferative markers, ontological enrichment of DEGs, immunohistochemical labeling, and examination of the literature regarding newly identified but less well-characterized putative markers. Overall, the strategy proved highly effective in discerning most of the known hair follicle cell types and several cell types known to be associated with the follicle such as immune cells and melanocytes. The physical isolation of the hair follicles still allowed for the capture of cells from endothelial lineages and sweat gland cells, allowing for

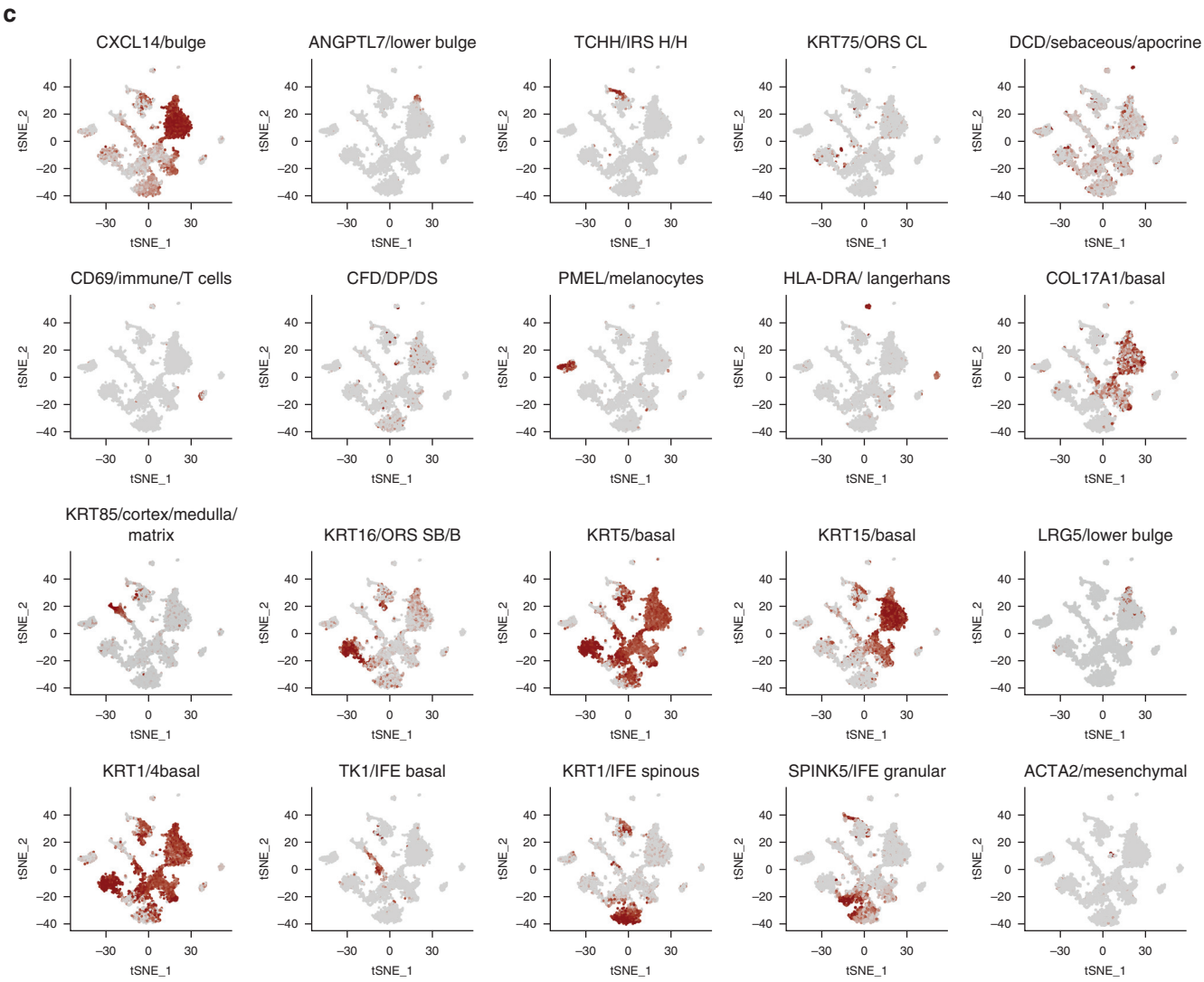


Figure 2. Continued.

the additional characterization of these important cell types in the skin.

We used expression of keratins to first establish cell identities, as this has proven useful in the past. In addition, we also used an existing database to confirm identities based on the entire profile of each cell cluster (Supplementary Figure S2). These identities and cell-type specificity of expression patterns were confirmed at the protein level using the immunostaining of follicles in our own lab (Figure 3 and Supplementary Figure S4) as well as a database of immunostaining patterns freely available (Supplementary Figure S4, Human Protein Atlas).

By performing ontological molecular overrepresentation analysis on each cell type's DEGs, we uncovered some interesting patterns that further confirmed cell identities and revealed previously undescribed biological pathways (Supplementary Figure S3). For instance, the epidermal subpopulations displayed enrichment for genes related to cell-cell adhesion and epidermal development, as would be expected for cells from stratified epithelia for whom barrier formation is a paramount activity; immune cells in the follicle preps showed enrichment for peptide antigen binding; the dermal sheath and papillae population showed enrichment for collagen binding and extracellular matrix organization; and cells of the IRS showed enrichment for genes related to keratinization. Enrichment profiling of identified cell types for known skin condition gene expression signatures permitted identification of previously undescribed epidermal and follicular constituents of skin pathogenesis. As expected, hyperactivation of immune response associated with conditions such as leprosy, Stevens–Johnson Syndrome, and mycosis fungoides revealed enrichment of signatures in Langerhans and T cells. Additionally, the three neoplastic signatures (squamous cell carcinoma, basal cell carcinoma, and melanoma) were linked to known pathological identification (i.e., epidermis and infundibulum for squamous cell carcinoma, bulge and lower bulge for basal cell carcinoma, and melanocytes for melanoma).

The primary limitation of scRNA sequencing is depth of coverage. Although this study permitted expanded coverage through integration of Drop-seq and 10X platforms from five distinct samples, deconvolution of highly related or sparse tissue populations (e.g., cortex vs. medulla vs. matrix) was not possible with the current sample size. However, together

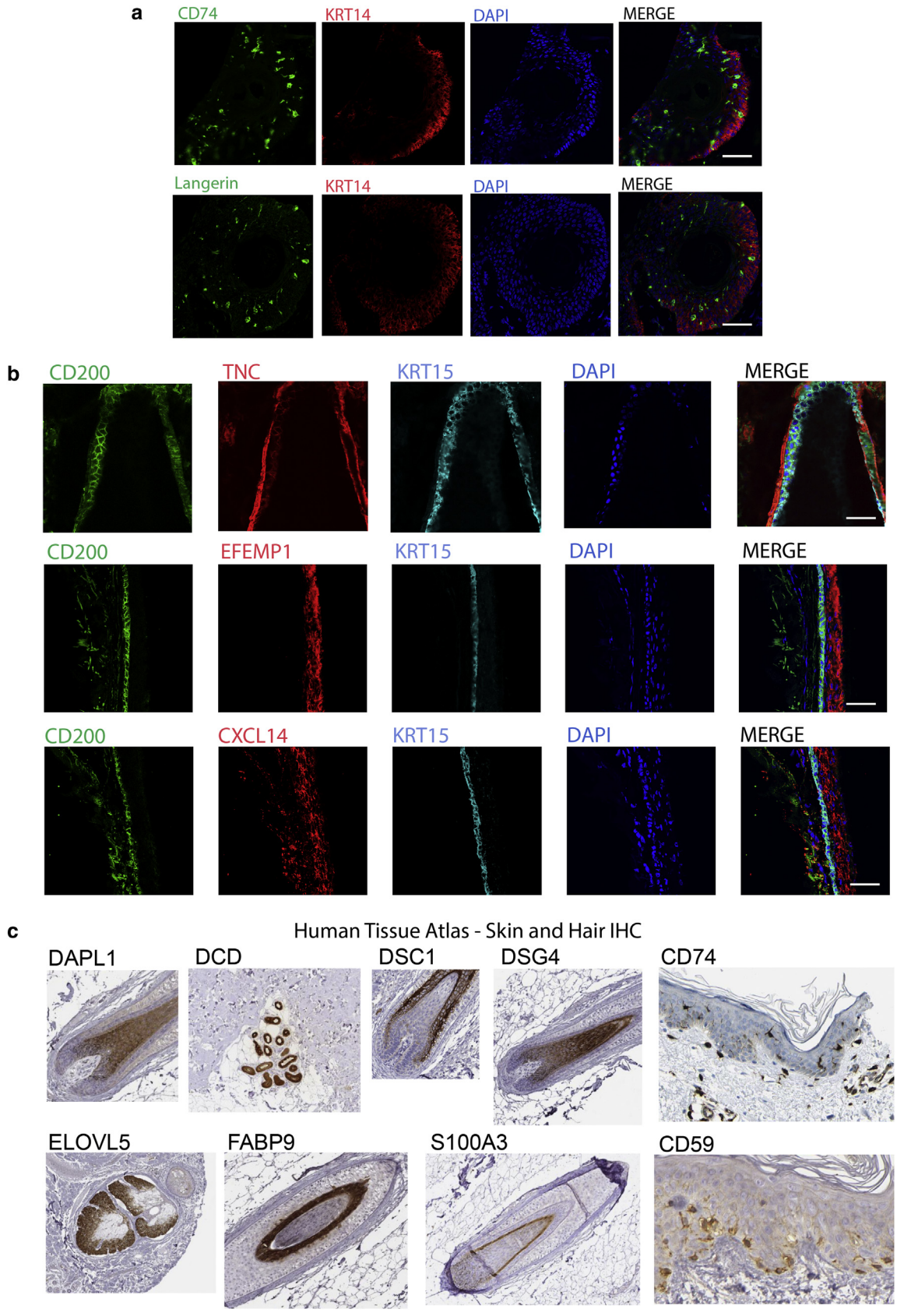


Figure 3. Validation of enrichment of gene expression at the protein level. (**a**) Immunostaining of serial sections of frozen hair follicle samples labeled with CD74 and Langerin (green) highlight Langerhans cells in the basal interfollicular epidermis (KRT14^{+}, red). Nuclei labeled with DAPI. Bar = 50 μm. (**b**) Immunostaining of serial sections of frozen hair follicle samples labeled with antibodies against CXCL14, TNC, and EFEMP1 (red) compared with canonical bulge markers CD200 (green) and KRT15 (blue). Nuclei labeled with DAPI. Bar = 50 μm. (**c**) IHC of the indicated epitopes derived from the Human Tissue Atlas for markers identified by scRNA-seq as specific to particular cell types. IHC, immunohistochemistry; scRNA-seq, single-cell RNA sequencing.

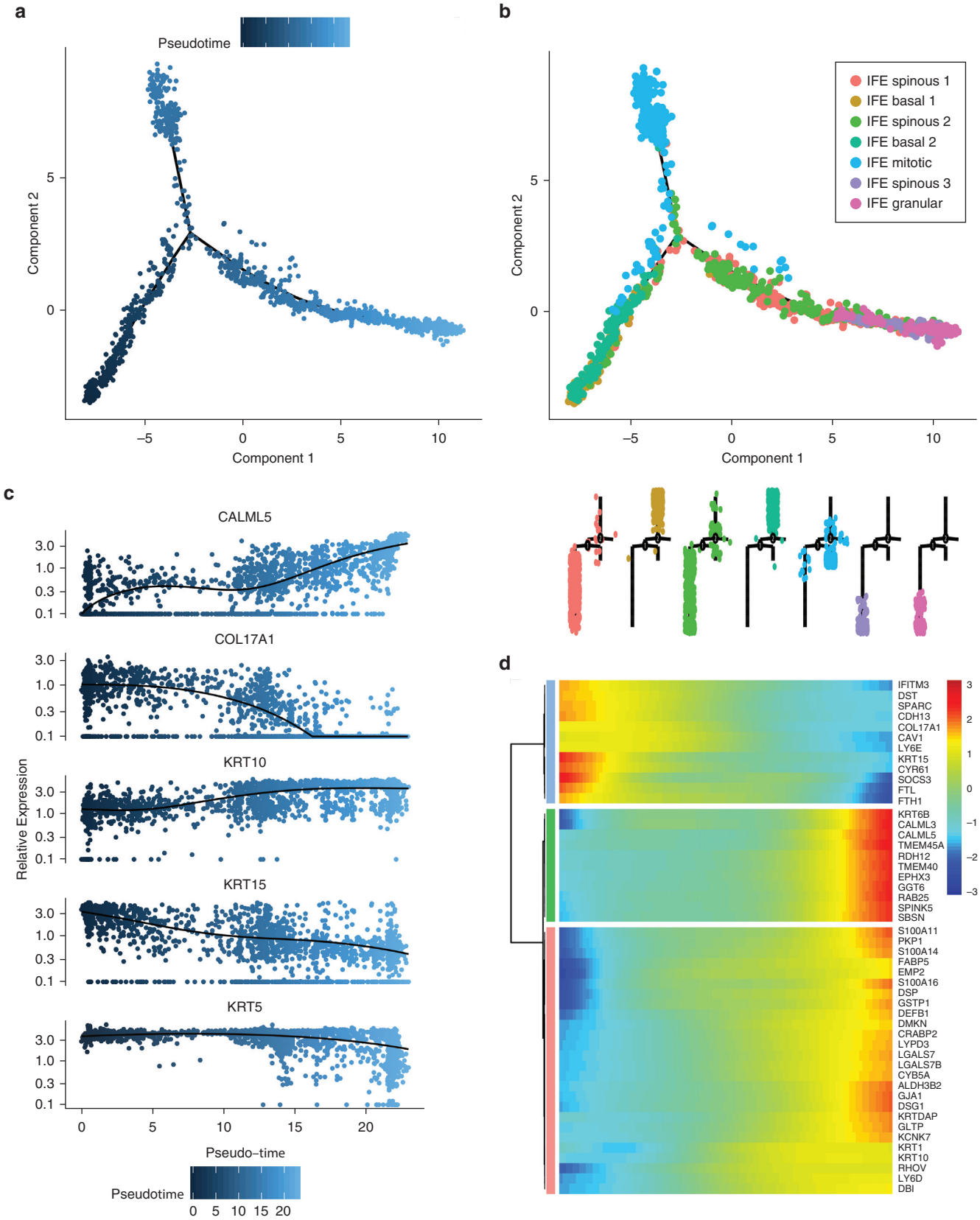


Figure 4. Trajectory analysis of interfollicular progenitor cells. (**a**) Pseudotemporal trajectory of analyzed interfollicular subpopulations. (**b**) Basal, spinous, and granular subpopulations in pseudotemporal order. (**c**) Expression of known critical keratin markers in pseudotime (branchpoint variations in expression not indicated). (**d**) Heatmap of top-scoring targets with differential expression across pseudotime. IFE, interfollicular epidermis.

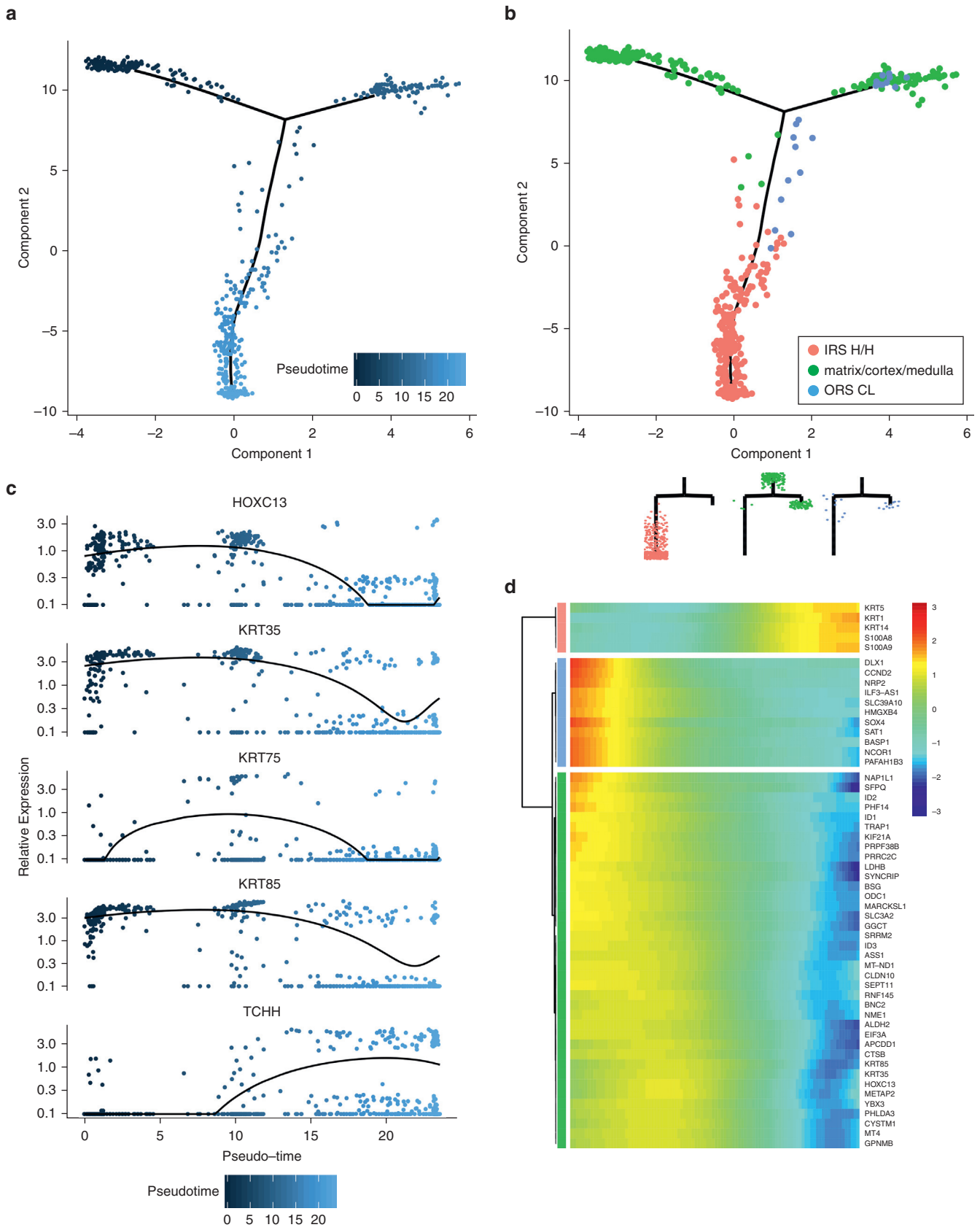


Figure 5. Trajectory analysis of matrix progenitor cells. (**a**) Pseudotemporal trajectory of analyzed interfollicular subpopulations. (**b**) Matrix, IRS H/H, IRS cuticle, cuticle subpopulations in pseudotemporal order. (**c**) Expression of critical differentiation markers in pseudotime (branchpoint variations in expression not indicated). (**d**) Heatmap of top-scoring targets with differential expression across pseudotime. IRS H/H, inner root sheath Huxley's/Henley's layers; ORS CL, outer root sheath companion layer.

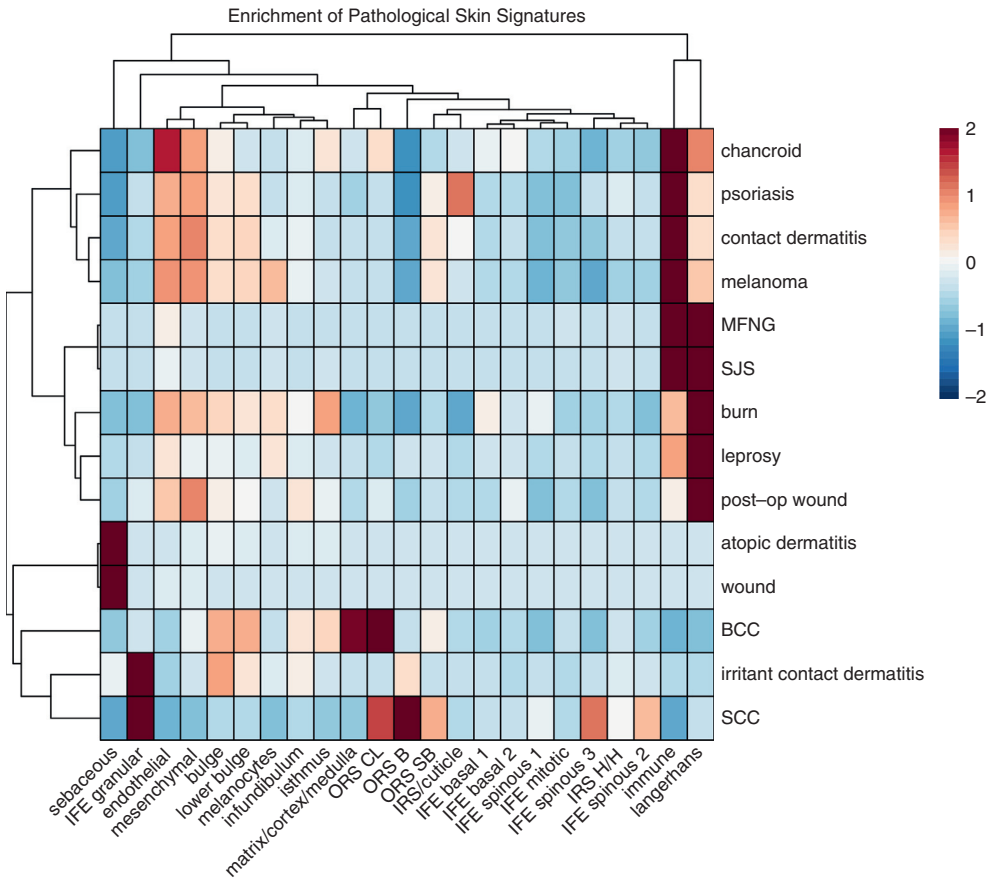


Figure 6. Cell type–specific enrichment of disease signatures. Heatmaps of the average expression of skin pathology signatures across identified cell types. This analysis looks at the pattern of previously established gene expression differences specific to the indicated skin disorders to determine whether individual cell types show enrichment for the patterns found in these syndromes (Developed by Matteo Pellegrini and available at: http://chlamy.mcdb.ucla.edu/NewDermDB/NewDermDB.html). BCC, basal cell carcinoma; IFE, interfollicular epidermis; IRS, inner root sheath; IRS H/H, inner root sheath Huxley's/Henley's layers; MFNG, mycosis fungoides; ORS B, outer root sheath basal layer; ORS CL, outer root sheath companion layer; ORS SB, outer root sheath suprabasal layer; SCC, squamous cell carcinoma; SJS, Stevens–Johnson Syndrome.

these data represent an important advance in the investigation of hair follicle biology. We anticipate these data will be borne out by others to supplement a variety of analyses and comparisons with other datasets to further elucidate critical components and mechanisms in epidermal and follicular cell development and differentiation.

METHODS

Sample harvest and preparation

Hair grafts were obtained from patients undergoing hair transplantation. Written informed consent was obtained from each patient and the study protocol was approved through the Institutional Review Board (IRB #16-000681-AM-00002). Single-cell suspensions were generated from the micrografts as previously described (Ohyama and Vogel, 2006). Briefly, the grafts were flushed with phosphate buffered saline and incubated in dispase overnight at 4 °C. The following day, the grafts were incubated in the same dispase solution at 37 °C for 30 minutes. Hair grafts were gently dissociated with a P1000 pipette and then incubated for 10 minutes in trypsin 0.05% solution diluted with phosphate buffered saline. The grafts were again gently dissociated with a P1000 pipette and placed back in 37 °C for 10 minutes. Trypsin was deactivated with 5% fetal bovine serum. Fibrous tissue and debris were filtered out with a 40-μm strainer. Single cells were visualized and counted with a hemocytometer and washed with phosphate buffered saline. Live cells were preferentially sorted using a FACS Aria III High-speed Cell Sorter and submitted for single-cell sequencing.

Immunostaining

Hair grafts were embedded in OCT compound, frozen, and sectioned for immunostaining. Frozen sections were fixed in acetone and labeled with the following primary antibodies: CD59 (Abcam, Cambridge, United Kingdom; ab69084, 1:50), Melan A (Santa Cruz Biotechnology, Dallas, TX; sc-20032, 1:50), CD200 (Bio-Rad, Hercules, CA; MCA1960, 1:75), CD74 (Santa Cruz Biotechnology; sc-6262, 1:50), Keratin 14 (Covance, Princeton, NJ; PRB-155P, 1:500), Keratin 15 (Covance; PCK-153P, 1:500), CXCL14 (Abcam; ab36622, 1:50), EFEMP1 (Abcam; ab106429, 1:50), and Langerin (Santa Cruz Biotechnology; sc-271272, 1:50). Samples were imaged using a Leica TCS SP8 Digital Light Sheet Microscope at the University of California, Los Angeles CNSI Advanced Light Microscopy/Spectroscopy Shared Resource Facility.

scRNA sequencing

The raw Drop-seq data was processed using the Drop-seq tools v1.12 pipeline from the McCaroll lab, utilizing the standard parameters as shown in the documentation (https://github.com/broadinstitute/Drop-seq/releases/tag/v1.12). In brief, cell and molecular barcodes were extracted from raw sequencing data based on bases 1–12 for cell and 13–20 for molecular barcodes while filtering out reads with poor quality bases (TagBamWithReadSequenceExtended). Subsequently, reads were trimmed to remove SMART adapter sequences as well as PolyA tails (FilterBAM, TrimStartingsequence, and PolyATrimmer). HiSat2 was used to align these filtered reads to the human reference genome, hg38. Aligned reads were then merged with the unaligned reads to recapture

molecular and cell BAM tags, and reads were tagged with GE if they overlapped with gene exons (MergeBamAlignment and TagReadWithGeneExon). Bead synthesis errors were then corrected and uMIs merged (DetectBeadSynthesisErrors). Finally, differential gene expressions were generated using standard parameters (DigitalExpression). We performed all downstream analysis on these differential gene expressions after filtering out cells with fewer than 250 genes. For 10X data, CellRanger2.2 pipelines were used to generate expression matrices with all standard parameters. Raw fastq files were processed and aligned to the human GrCh38 genome, and the CellRanger cell detection algorithm was utilized to determine the numbers of gene expression microarrays per run. No other parameters were changed.

Computational methods

All analyses were performed in R. Gene expression analysis and cell-type identification were performed using the Seurat package (Butler et al., 2018; Satija et al., 2015). Seurat objects were created for each sample. Only those genes that were expressed in more than five cells and cells that expressed more than 200 genes were retained (22,000 cells). Cells with a high proportion (>5%) of mitochondrial expression were filtered out, as these typically represent cells damaged during isolation. Filtering was also performed on the number of detected transcripts within each sample to eliminate partial cells and doublets, respectively. Datasets were normalized and scaled according to default settings with regression against cell number and mitochondrial content, following by variable gene expression calculation (FindVariableGenes, x low cutoff = 0.03, x high cutoff = 3, y cutoff = 1). The union of the top 2,000 variable genes was used to perform canonical correlation analysis across the different samples and align the subspaces (FindIntegrationAnchors and IntegrateData), followed by integrated t-distributed stochastic neighbor embedding visualization of all cells. A total of 5,270 remaining cells were used in the final analysis. Gene expression markers were calculated for each subpopulation (FindAllMarkers, method = *t*-test). Average, log-normalized expression profiles were calculated for each gene using the AverageExpression function. Gene ontology and pathway enrichment analyses were executed in clusterProfiler. Heatmaps were generated with the pheatmap package. Pseudotime calculations were performed in Monocle 2 as previously described using the clusters and DEGs identified in Seurat using default settings (Qiu et al., 2017). Cell subpopulations were identified as described in the primary text.

Data availability statement

The scRNA sequencing data are in NIH-GEO (GSE129611). The averaged data for all cell clusters are available in Supplementary Table S2. The raw data are available through Figshare: https://figshare.com/s/74a818b57f08b0873a8c.

ORCIDs

Rie Takahashi: http://orcid.org/0000-0003-0255-7040
Adrienne Grzenda: http://orcid.org/0000-0001-5952-6953
Thomas F. Allison: http://orcid.org/0000-0001-5881-6991
Jeffrey Rawnsley: http://orcid.org/0000-0002-4987-6754
Samuel J. Balin: http://orcid.org/0000-0002-7646-1319
Shan Sabri: http://orcid.org/0000-0002-0885-6159
Kathrin Plath: http://orcid.org/0000-0001-7796-3372
William E. Lowry: http://orcid.org/0000-0003-2932-2276

CONFLICT OF INTEREST

WEL is a co-founder, shareholder, and president of Pelage Pharmaceuticals. Pelage Pharmaceuticals did not sponsor or support this work in any way. The other authors have no conflicts of interest to disclose.

ACKNOWLEDGMENTS

The authors are indebted to Dr. Rassman for his support of this project and their patients who donated follicles for this effort. In addition, the authors are grateful to the BSCRC Genomics Core Facility at University of California, Los Angeles for RNA sequencing. Furthermore, the authors acknowledge Michael Rendl for his expertise on Dermal populations and Matteo Pellegrini (University of California, Los Angeles) for development and sharing of the DermDB2 database. Confocal laser scanning microscopy was performed at the CNSI Advanced Light Microscopy/Spectroscopy Shared Resource Facility at University of California, Los Angeles. This work was supported by awards to RT (Dermatology Foundation, Dermatologic Research Foundation of California, and the T32 Training Grant Dermatological Scientist Training Program, University of California, Los Angeles) and to WEL (NIH NIAMS 5R01AR070245).

AUTHOR CONTRIBUTIONS

Conceptualization: WEL, RT, TFA; Data Curation: RT SS, SJB, KP, TFA; Resources: JR, SJB; Software: AG, KP; Supervision: KP, WEL; Validation: RT; Writing - Original Draft Preparation: WEL

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at www.jidonline.org, and at https://doi.org/10.1016/j.jid.2019.07.726.

REFERENCES

Acosta JR, Joost S, Karlsson K, Ehrlund A, Li X, Aouadi M, et al. Single cell transcriptomics suggest that human adipocyte progenitor cells constitute a homogeneous cell population. Stem Cell Res Ther 2017;8:250.

Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 2018;36:411–20.

Cheng JB, Sedgewick AJ, Finnegan AI, Harirchian P, Lee J, Kwon S, et al. Transcriptional programming of normal and inflamed human epidermis at single-cell resolution. Cell Rep 2018;25:871–83.

Fuchs E. Beauty is skin deep: the fascinating biology of the epidermis and its appendages. Harvey Lect 1998-1999;94:47–77.

Joost S, Zeisel A, Jacob T, Sun X, La Manno G, Lönnerberg P, et al. Single-cell transcriptomics reveals that differentiation and spatial signatures shape epidermal and hair follicle heterogeneity. Cell Syst 2016;3:221–37.e9.

Klein AM, Macosko E. InDrops and Drop-seq technologies for single-cell sequencing. Lab Chip 2017;17:2540–1.

Mesler AL, Veniaminova NA, Lull MV, Wong SY. Hair follicle terminal differentiation is orchestrated by distinct early and late matrix progenitors. Cell Rep 2017;19:809–21.

Ohyama M, Terunuma A, Tock CL, Radonovich MF, Pise-Masison CA, Hopping SB, et al. Characterization and isolation of stem cell-enriched human hair follicle bulge cells. J Clin Invest 2006;116:249–60.

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.

Picelli S. Single-cell RNA-sequencing: the future of genome biology is now. RNA Biol 2017;14:637–50.

Qiu X, Hill A, Packer J, Lin D, Ma YA, Trapnell C. Single-cell mRNA quantification and differential analysis with Census. Nat Methods 2017;14:309–15.

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

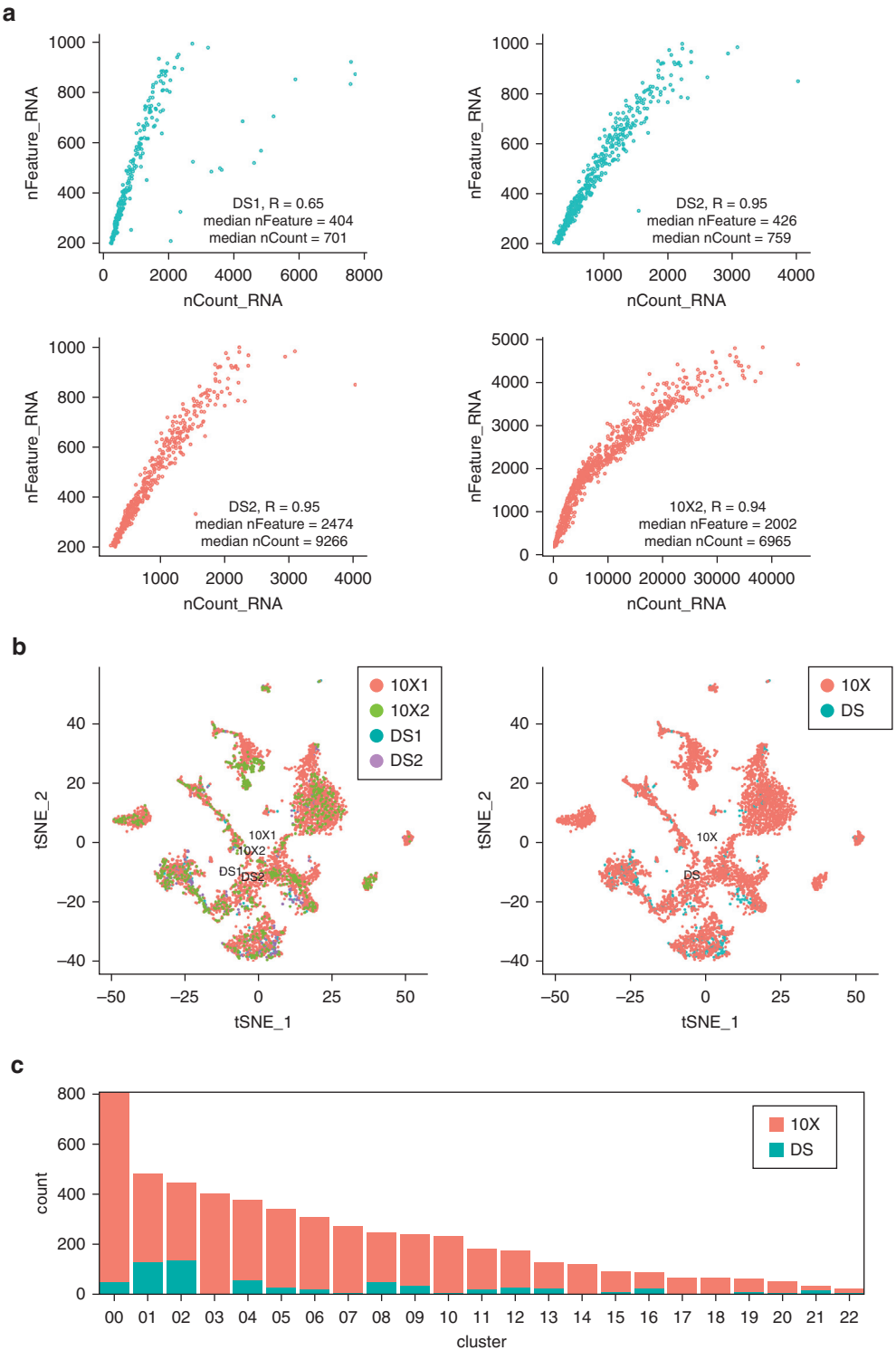
Swindell WR, Johnston A, Voorhees JJ, Elder JT, Gudjonsson JE. Dissecting the psoriasis transcriptome: inflammatory- and cytokine-driven gene expression in lesions from 163 patients. BMC Genomics 2013;14:527.

Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. Science 2015;347:1260419.

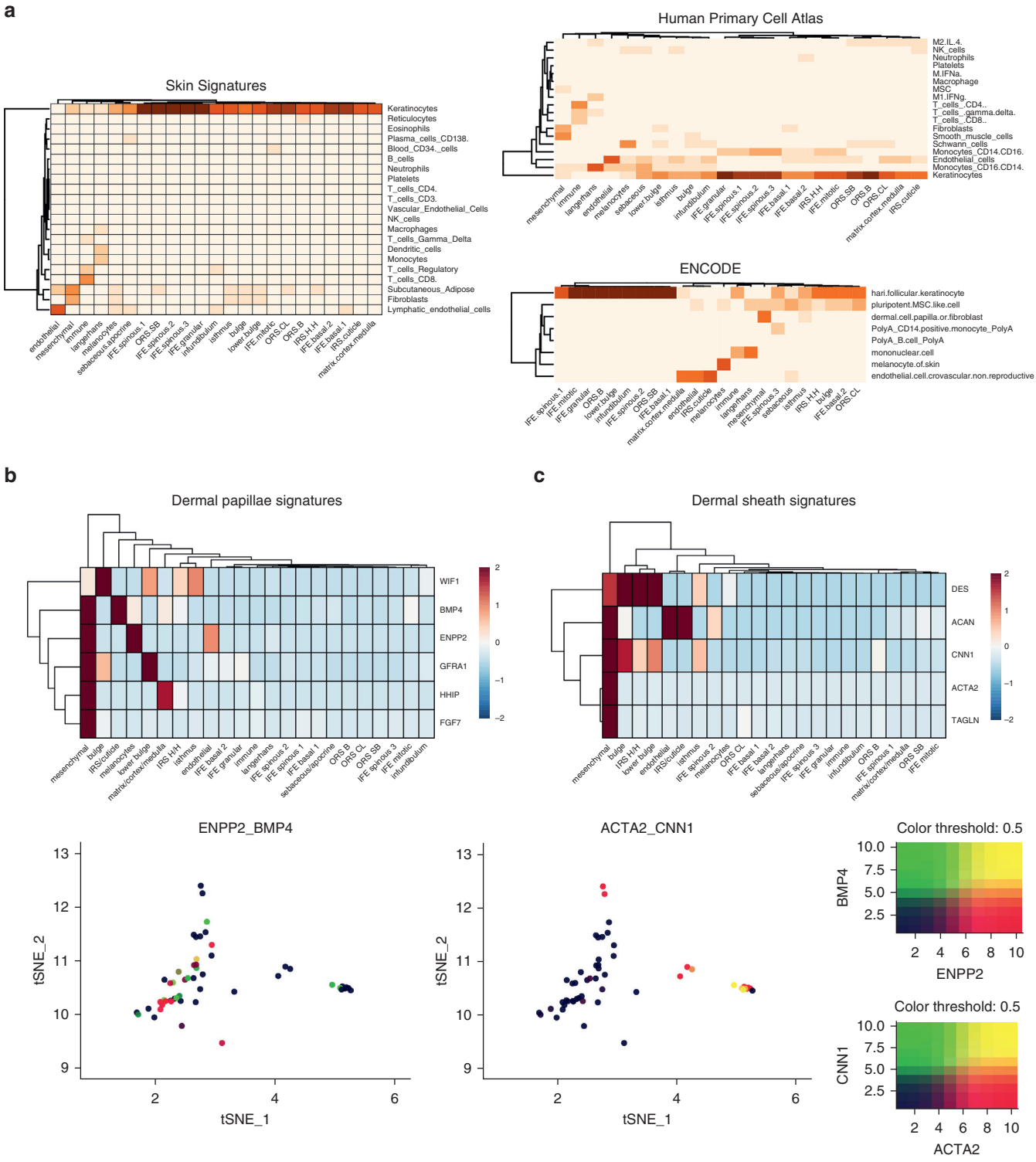
Weisenfeld NI, Kumar V, Shah P, Church DM, Jaffe DB. Direct determination of diploid genome sequences. Genome Res 2017;27:757–67.

Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, et al. Comparative analysis of single-cell RNA sequencing methods. Mol Cell 2017;65:631–43.e4.

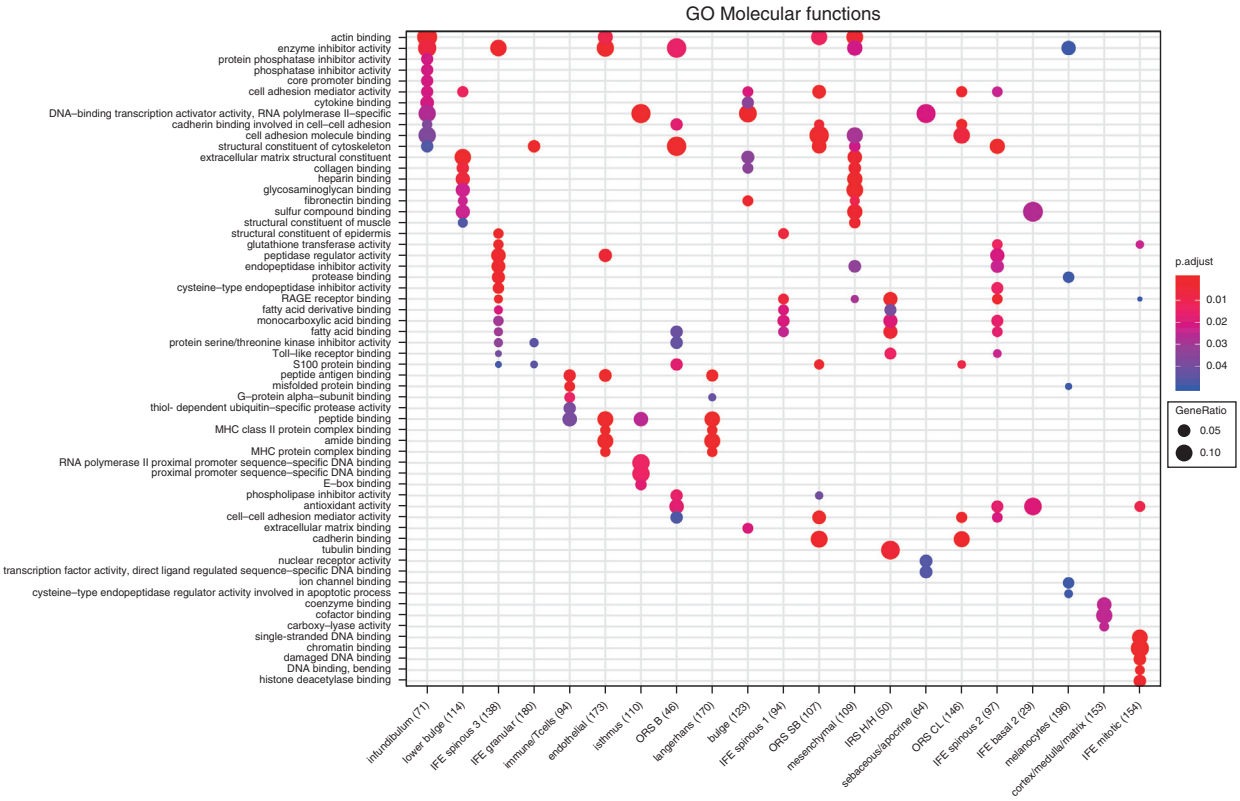
SUPPLEMENTARY MATERIAL



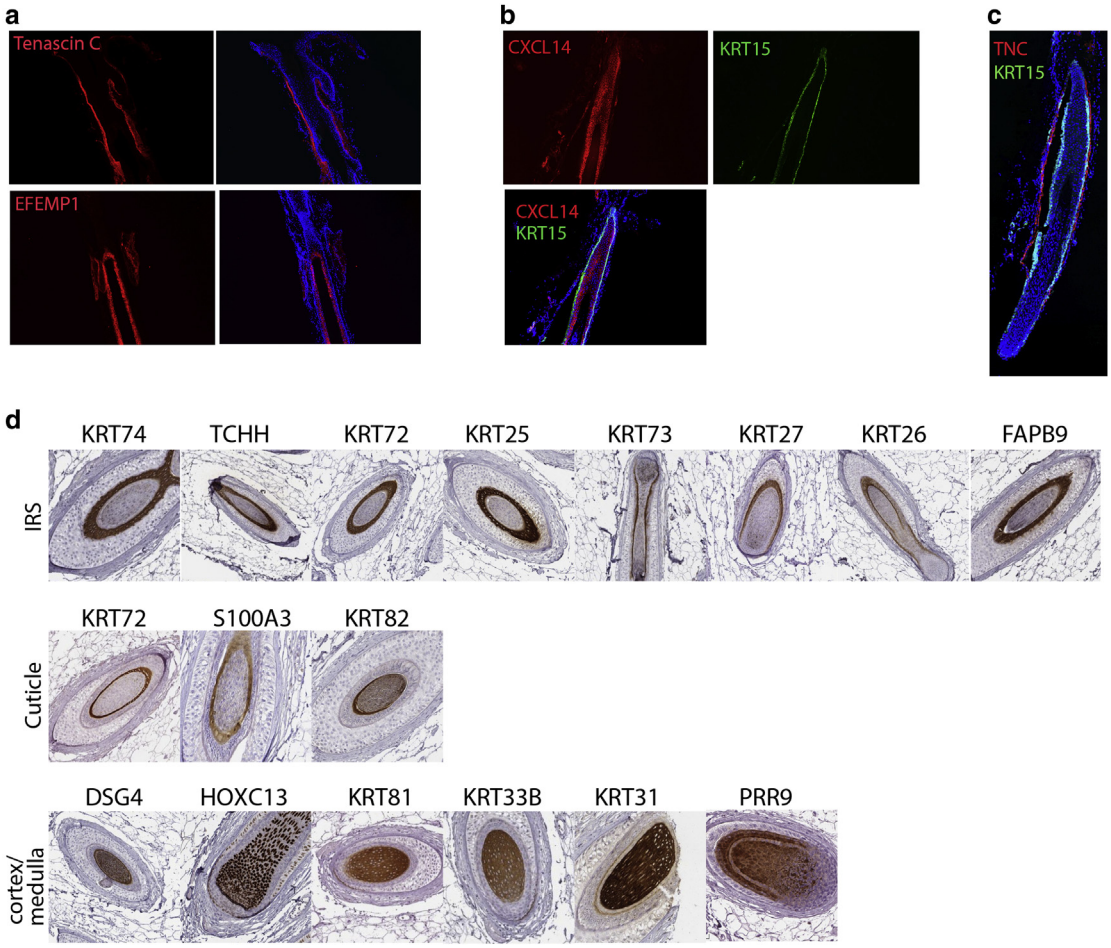
Supplementary Figure S1. Cross-platform data integration of Drop-seq and 10X datasets. (**a**) UMI and gene detection rates of Drop-seq (top, $n = 2$) and 10X (bottom, $n = 2$) platforms, respectively, with median values per cell. (**b**) t-SNE visualization of the results of the unsupervised, graph-based clustering on the integrated dataset by (left) individual dataset or (right) platform). (**c**) Proportion of cells from the 10X and Drop-seq datasets within each identified cluster. DS, Drop-seq; t-SNE, t-distributed stochastic neighbor embedding.



Supplementary Figure S2. Gene Expression Deconvolution Interactive Tool analysis of gene expression. The tool may be found at: http://webtools.mcdb.ucla.edu. (**a**) Enrichment of skin signatures (Swindell et al., 2013), Human Primary Cell Atlas, and ENCODE. (**b** and **c**) Heatmap of dermal papillae and dermal sheath signature genes (provided by Michael Rendl, personal communication) and t-SNE of colocalized expression of two markers for each cell type. IFE, interfollicular epidermis; IRS, inner root sheath; IRS H/H, inner root sheath Huxley's/Henley's layers; MSC, mesenchymal stem cell; NK, natural killer; ORS B, outer root sheath basal layer; ORS CL, outer root sheath companion layer; ORS SB, outer root sheath suprabasal layer; t-SNE, t-distributed stochastic neighbor embedding.



Supplementary Figure S3. Overrepresentation analysis of GO molecular functions by cluster genes. Significantly differentially expressed genes between clusters were defined as logFC $>$ 1 and adjusted *P*-value $<$ 0.05. Only enriched categories with an adjusted *P*-value $<$0.05 are shown. Overrepresentation analysis and visualization performed using ClusterProfiler R package. FC, fold change; GO, Gene Ontology; IFE, interfollicular epidermis; IRS H/H, inner root sheath Huxley's/Henley's layers; MHC, major histocompatibility complex; ORS B, outer root sheath basal layer; ORS CL, outer root sheath companion layer; ORS SB, outer root sheath suprabasal layer.



Supplementary Figure S4. Immunohistochemical confirmation of expression of proteins predicted by scRNA-seq. (**a-c**) Immunofluorescence for the same markers shown in Figure 3, but at lower magnification (5x). (**d**) Immunostains were derived from the Human Protein Atlas (Uhlén et al., 2015). scRNA-seq, single-cell RNA sequencing.