

Supplementary Information for

Epigenetic Switch Reshapes Epithelial Progenitor Cell Signatures and Drives Inflammatory Pathogenesis in Hidradenitis Suppurativa

Lin Jin^{1,2,3*}, Yunjia Chen⁴, Suhail Muzaffar^{2,3}, Chao Li⁵, Carlos A. Mier-Aguilar², Jasim Khan^{2,3}, Mahendra P. Kashyap^{2,3}, Shanrun Liu⁶, Ritesh Srivastava^{2,3}, Jessy S. Deshane⁷, Tim M. Townes⁵, Boni E. Elewski², Craig A. Elmets², David K. Crossman⁴, Chander Raman², Mohammad Athar^{1,2,3*}

¹Center for Epigenomics and Translational Research in Inflammatory Skin Diseases, University of Alabama at Birmingham, AL 35294, USA

²Department of Dermatology, School of Medicine, University of Alabama at Birmingham, AL 35294, USA

³UAB Research Center of Excellence in Arsenicals, University of Alabama at Birmingham, AL 35294, USA

⁴Department of Genetics, University of Alabama at Birmingham, AL 35294, USA

⁵Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, AL 35294, USA

⁶UAB Institutional Research Core Program, Flow Cytometry and Single Cell Core, University of Alabama at Birmingham, AL 35294, USA

⁷Division of Pulmonary, Allergy and Critical Care Medicine, University of Alabama at Birmingham, AL 35294, USA

*Co-Corresponding Author: Lin Jin. Email: ljin@uabmc.edu; Mohammad Athar (Lead corresponding author). mohammadathar@uabmc.edu

This PDF file includes:

Main Materials and Methods

Figures S1 to S7 and Table S1

Materials and Methods

Isolation of epidermal cells, primary basal keratinocytes culture, and enrichment of CD49f^{high} keratinocytes for next-generation sequencing

Healthy and HS skin was digested with 2.5U/ml Dispase-II (Millipores Sigma, Cat#4942078001) in Hanks' Balanced Salt Solution (HBSS, Corning, Cat#21-022-CM) overnight at 4°C. After digestion, the epidermis was peeled off and minced. The minced epidermis was incubated with 0.05% trypsin-EDTA at 37°C for 40 mins for further digestion. The suspension was then filtered through a 40µm cell strainer and spun down at 300g at 4°C for 5 mins. Collected single cells were either frozen for the application of next-generation sequencing (NGS) or for primary cell culture supplied with KBM-Gold keratinocyte growth medium (Lonza, Cat# 00192060) at 37°C in a humidified chamber with 5% carbon dioxide. We used keratinocytes (within 2 passages) for performing experiments. For sc-RNA seq, cryovials were thawed and subjected to remove apoptotic and dead cells by using the Annexin Dead Cell Removal Kit (StemCell Technologies, 17899) and EasySep™ Magnet (StemCell Technologies, 18000). Cells were then loaded and processed using the 10xGenomics Chromium platform. For low-input ATAC&mRNA and CUT&RUN samples, enriched living cells were stained with APC anti-human CD49f antibody (1:200 Biolegend, Cat# 313616) for 10 minutes at 4°C. After washing with sorter buffer (PBS supplemented with 0.5% BSA and 2mM EDTA), each sample was incubated with 4 µl of Anti-APC MultiSort MicroBeads (Miltenyi Biotec, Cat# 130-091-255) in 96 µl of sorted buffer for 15 minutes at 4°C. Samples were passed over magnetized MS columns. The columns were washed 3 times to remove unlabeled cells. After the last wash, the columns were removed from the magnetic field, and bound CD49f^{high} cells were eluted in sorter buffer.

Flow cytometry

Disaggregated epidermal cells were labeled with antibodies in PBS + 1% FBS for 10 min at 4 °C and unbound antibodies were washed out three times with PBS. Cells were analyzed on a Beckman Coulter CytoFLEX S flow cytometer and all data were analyzed with FlowJo software (version 10.8). The following antibodies were used: APC anti-human

CD49f antibody (Biolegend, Cat# 313616) (1:200), APC isotype control antibody (Biolegend, Cat# 400511) (1:200), PE anti-human CD71 antibody (Biolegend, Cat# 334106) (1:200) and APC isotype control antibody (Biolegend, Cat# 400511) (1:200).

Cell Immunofluorescence (IF) staining

Immunofluorescence staining was performed on cells grown on coverslips in 6-well culture plates. Cells were then fixed with 4% paraformaldehyde followed by permeabilization with 0.1% Triton X-100 for 15 minutes. After blocking with 1% bovine serum albumin for 45 minutes, cells were incubated with Anti-Keratin 14 (1:500, Biolegend, Cat# 906004) overnight at 4 °C. Following three PBS washes, cells were incubated with donkey anti-chicken™ 488 (1:500, Invitrogen, Cat# A78948). After three times washing with PBS, coverslips were mounted with DAPI and examined under a fluorescence microscope (KEYENCE BZ-X710).

Tissue Hematoxylin & Eosin (H&E) and Immunofluorescence (IF) staining

Excised tissues were fixed in 10% neutral buffered formalin for 24 h at RT, then stored in 70% ethanol. Tissues were embedded in paraffin and then cut into 5 µm sections using a microtome (HM 325, Thermo Fisher Scientific). This was followed by staining with H&E. For IF, sections were placed in blocking buffer containing 5% normal goat serum in PBST (PBS+ 0.4%Triton X100) for 1h at room temperature. The primary antibodies were then added to slides and incubated overnight at 4°C. Sequential staining was done for visualization of more than one target in single specimen. On the following day, sections were re-incubated with the indicated secondary antibodies and further were fixed in DAPI containing Vectashield antifade mounting medium (H-1200, Vectorlabs). Sections were visualized under FLUOVIEW FV3000 confocal microscope (Olympus, USA) equipped with FV3000 Galvo scan unit and FV3IS-SW version 2.3.2.169 software. The following primary antibodies were used: anti-PTTG1 (1:50, Sigma, Cat# HPA008890), anti-HELLS (1:50, Thermal Fisher Scientific, Cat# PA5-64099), anti-COL17A1 (1: 100, Invitrogen, Cat# MA5-31984), Anti-Keratin 14 (1:400, Biolegend, Cat# 906004), anti-Keratin1/10 (Santa Cruz, Cat# sc-53251, 1:400), anti-S100A7 (1:100, Novus, Cat#NB100-56669s),

anti-S100A8 (1:100, Abcam, Cat# Ab92331). Secondary antibodies include goat anti-mouse IgG Dylight 488 (1:200, Thermo Scientific, Cat#35502), goat anti-rabbit IgG Dylight 488 (1:200, Invitrogen, Cat#35552), goat anti-rabbit Alexa FluorTM594 (1:400, Invitrogen, Cat# A11012), donkey anti-chickenTM 488 (1:400, Invitrogen, Cat# A78948).

Quantitative RT-PCR (qRT-PCR)

Total RNA was extracted with TRIzol reagent (Invitrogen, Cat# 15596026) and reverse-transcribed using the SuperScript III First-Strand Synthesis System (Invitrogen, Cat#18080051) with Oligo(dT)20 primers. For gene expression, qPCR was performed with TaqManTM Fast Advanced Master Mix (Applied Biosystems, Cat#4444557) on QuantStudio 12K Flex (Applied Biosystems) machine. The cycling acquisition program was as follows: 95 °C 20 sec, 40 cycles of 95 °C for 3 sec, 60 °C for 30 sec. A set of TaqMan probes from Thermal Fisher used are listed (*SI Appendix*, Dataset S6).

CRISPR/Cas9-mediated enhancer editing

Primary normal human neonatal keratinocytes (NHEK, ATCC, Cat# PCS-200-010) were cultured in KBM-Gold keratinocyte growth medium. Paired dual gRNAs specifically targeting two or scramble control dual gRNAs (*SI Appendix*, Dataset S6) were designed with Genetic Perturbation Platform (<https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design>) and were chemically synthesized (Synthego). 350 pmol of HiFi Cas9 (Integrated DNA Technologies) and 400 pmol sgRNA were mixed and incubated at room temperature for 10 minutes. Each pair of nucleofection of ribonucleoprotein (RNP) was then mixed. 1x10⁶ cells were washed with KGM and resuspended in 100 µl Nucleofector Solution (Lonza, V4XP-3024), gently mixed with pre-assembled RNPs, and electroporated with the program DS-138 using Nucleofector 4D (Lonza). To confirm the target deletion in the bulk transduced cells, PCR primers were designed to amplify the WT or deleted allele (data file S6). The sizes of PCR products were checked by agarose gel electrophoresis. PCR products were subjected to pGEM-T easy vector (Promega) for cloning. Plasmid DNA was extracted from individual colonies and deletions were confirmed by Sanger sequencing.

scRNA-seq library preparation and data analysis

Single epidermal cells were loaded into 10xcontroller and individual cells were captured into oil droplets to form Gel Bead-In-EMulsions (GEMs). The single-cell cDNA libraries were constructed according to the manufacturer's protocol using Chromium NextGEM single cell 3' reagent kit (10XGenomics, version 3.1), and then sequenced by Novaseq 6000 machine with targets minimum of 20,000 reads/cell.

Read alignment was performed by using the default parameters of Cell Ranger single-cell software suite (version 6.0.2) (10x Genomics). The quality of the sample-specific FASTQ file was evaluated by the counts of Cell Ranger which were aligned to the human reference genome (GRCh38-2020-A) using STAR to generate the gene expression matrix. The filtered gene expression matrices were then used for downstream analyses by Seurat R package (version 4.0.5). For quality control, cells with fewer than 50 genes or more than 10,000 genes detected, and more than 15% mitochondrial genes were excluded from the analysis. Sequencing reads for each gene were normalized to total UMIs in each cell to obtain normalized UMIvalues by "NormalizeData" function. Dimensionality reduction was performed with "RunUMAP" function with parameters 'dims=1:30'. To avoid inter-individual differences, we combined all samples using the standard integration function in Seurat. The integrated data were then used for cell clustering by performing 'FindClusters' function at a resolution of 0.5 to define cell identity and visualized by UMAP. Wilcoxon rank sum test was used for heatmap of top 10 differentially expressed genes in each cluster. distinct gene signature. For generating cell cluster-specific heatmaps, we consider genes in individual cell with a fold change > 1.1 or < 0.9 as DEG. To rule out the bias of clustering algorithm on segregating subpopulations, we also re-run all the libraries using the latest Seurat tool (version 4.3.0) with the same parameters abovementioned except that the groups of samples were merged right after using "CreateSeuratObject" and before filtering for quality control. Marker genes for each cluster were determined with the Wilcoxon rank-sum test by "FindAllMarkers" function. Only those with $|avg_logFC| > 0.25$ and 'p_val_adj' < 0.05 were considered as marker genes. To identify differentially expressed genes between clusters, we used the "FindMarkers" Seurat function with logfc. threshold = 0.5 and

`test.use = "wilcox"`. We obtained identical epidermal clusters annotated by the same marker genes and similar expression profiles for top different genes.

Pseudotime analysis

Pseudotime analysis of the keratinocyte lineage was performed using SeuratWrappers (version 0.3.0) and Monocle3 (version 1.0.0). Monocle3 was run on our previously normalized counts matrix created by Seurat 4.0.5. The data were subject to UAMP dimensional reduction and cell clustering using the “cluster_cells” function. A principal graph was plotted through the UMAP coordinates using the “learn_graph” function that represents the path across the keratinocyte differentiation. This graph was further used to order cells in pseudotime using the “ordercells” function. We then re-ran “ordercells” with Basal I cells as the root cell state.

Cell-Cell Communication Analysis

To infer cell-cell communication established by the distinct cell types generated from Seurat (version 4.0.5), we used CellChat (version 1.4.0) tool with default parameters (1). To access the major signaling inputs and outputs among subsets, the “CellChatDB.human” database, as well as “netVisual_aggregate” and “netVisual_heatmap” functions were used to show the strength or weakness of cell-cell communication networks.

Low-input ATAC&mRNA library preparation and data analysis

Low-input ATAC&mRNA-seq was performed on CD49f^{high} keratinocytes following previously-reported protocol with minor modification (2). Briefly, 1X10⁴ single cells from each sample were resuspended in transposition mix (25 µl 2×TD buffer, including 2.5 µl Tn5 (Illumina, Cat#15027865), 16.5 µl PBS, 0.5 µl 1% digitonin, 0.5 µl 10% Tween-20, 1 µl RNase inhibitor, and 4 µl nuclease-free water). Permeabilization/transposition reactions were incubated at 37 °C for 30 min. EDTA and LiCl were added to a final concentration of 10 mM and 0.5 M, respectively, followed by adding 100 µl of Lysis/Binding Buffer in Dynabeads® mRNA DIRECT™ Micro Kit (Invitrogen, Cat# 61021). 20 µl pre-washed Dynabeads® Oligo (dT)25 was added into the cell lysate to

enrich mRNA. The samples were then placed on a Dynal magnet to separate the mRNA and genomic DNA (gDNA). The supernatant containing fragmented gDNA was purified with MinElute PCR purification kit (Qiagen, Cat # 28004) and further amplified using NEBNext® High-Fidelity 2X PCR Master Mix (New England Biolabs, Cat# M0541s). The mRNA captured on beads was washed using Dynabeads® mRNA DIRECTTM Micro Kit. The Dynabeads-mRNA complex was then resuspended in 20 µl of reverse transcription reaction mix without any primer prepared using SuperScriptTM IV First-Strand Synthesis System (Invitrogen, Cat #18091050). Direct fragmentation of the mRNA/cDNA hybrids and PCR amplification were performed on beads using Nextera XT DNA Library Prep Kit (Illumina, Cat# FC1311096).

For ATAC-seq data analysis, reads were trimmed using Trim Galore (version 0.6.7) with the following flags: '--paired --trim-n --trim1'. Trimmed reads were aligned to the human reference genome (GRCh38.p13.Release36) using Bowtie2 (version 2.4.2) with the following flags: '-S -p 4 -X 2000'. Picard tools (version 1.9.1) SortSam then converted the SAM format to BAM format with the following flags: 'SORT_ORDER=coordinate CREATE_INDEX=true' followed by marking duplicates with Picard tools. MACS2 (version 2.1.2) 'callpeak' command was used to call peaks on each sample with flags '-f BAMPE -g hs -q 0.01 -B'. DESeq2 was run for significant differences between healthy and HS samples on the peak matrix and differential peaks with by absolute log2 fold change >0.5 and false discovery rate (FDR) value <0.1 for 'gain' or 'loss'.

For RNA-seq data analysis, sequencing reads were trimmed with Trim Galore and then mapped to the human reference genome (GRCh38.p13.Release36) using STAR (version 2.7.10a_alpha_220601) (options:--outReadsUnmapped Fastx --outSAMtype BAM SortedByCoordinate --outSAMAttributes All). Transcript abundances were calculated using HTSeq-count (version 2.0.2) with options '-r pos -t exon -i gene_id -a 10 -s no -bam'. DESeq2-calculated differentially expressed genes with a *P* value <0.05, as well as a log2 fold change >1, were further analyzed.

CUT&RUN library preparation and data analysis

CUT&RUN experiments were carried out following the protocol of Epicypher CUTANA CUT&RUN kit (version 1.8). Briefly, nuclei from epithelial cells or CD49^{high} keratinocytes (1×10^4) were isolated with Wash Buffer (20 mM HEPES, 150 mM NaCl, 0.5 mM Spermidine, 1X Roche cOmplete, protease inhibitor). The nuclei were captured with Concanavalin A (ConA) beads and incubated with 0.5 µg primary antibodies, including anti-S100A8 (Abcam, Cat#ab92331), anti-H3K4me1(Epicypher, Cat#N013-0040), H3K27me3 (Invitrogen, Cat# MA5-11198), H3K27ac (ActiveMotif, Cat#91193) in Antibody Buffer overnight at 4 °C. Then CUTANA pAG-MNase was added and incubated for 10 min at RT. CaCl₂ was next added at a final concentration of 100 mM to activate the enzyme. The reaction was carried out for 2 hours at 4 °C and stopped by STOP buffer. Samples were incubated for 10 min at 37 °C in a thermocycler. DNA was purified using the CUANA DNA Purification Kit as per the provided instructions. Purified CUT&RUN-enriched DNA was prepared sequencing libraries using the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs, Cat# E7645s). For CUT&RUN sequence data analysis, sequence reads were run through CUT- RUNTools-2.0 (<https://github.com/fl-yu/CUT-RUNTools-2.0>) following their default protocol except for setting “frag_120” to “FALSE”. The histone peaks were scaled by the 5-kb region around the peak summit and were used to profile the heatmap using ‘deepTools’.

Gene Ontology analysis

For the gene ontology (GO) analyses, DEGs were queried into the Gene Functional Annotation Tool from the DAVID Bioinformatics Database (version 2021). GO option GOTERM_Biological Process_ALL and Fisher’s Exact test were selected and the top GO terms with a significant *P*-value were chosen as categories.

Motif analysis

Motif analysis on peak regions was performed using HOMER software (<http://homer.ucsd.edu/homer/ngs/peakMotifs.html>), with a combined list of peaks as the input. The parameters for the main script “findMotifsGenome. pl” in HOMER software were: findMotifsGenome.pl <peak/BED file > <genome > <output directory > -size given – rna –mask. For TOMTOM motif analysis, a database of selected putative enhancer

sequence was run through Tomtom Motif tool-5.5.1 (<https://meme-suite.org/meme/tools/tomtom>) following their default parameters.

Activity-By-Contact (ABC) model analysis

Activity-by-Contact (ABC) model (<https://github.com/broadinstitute/ABC-Enhancer-Gene-Prediction>) (3) was used to predict enhancer-promoter connections, based on datasets of chromatin accessibility (ATAC-seq) and histone modifications (H3K27ac CUT&RUN-seq). Briefly, to identify candidate enhancer regions, peaks were called on previously generated ATAC-Seq bam files with MACS2 using the parameters -p 0.1 –call-summits. The peak file was used with the script makeCandidateRegions.py with the parameters –peakExtendFromSummit 250 –nStrongestPeaks 150,000. Interactions from the output file EnhancerPredictionsFull.txt were used for downstream analysis. For ABC loop Plot creation in IGV, the aligned and duplicate removed BAM files were loaded.

MTT cell viability assay

Cell growth was assessed by a standard MTT assay. Briefly, NHEK cells (passage of 2) were seeded into 96-well culture plates (0.7×10^4 cells/well) in KBM-Gold keratinocyte growth medium (Lonza, Cat# 00192060). The cells were treated with BAY-985 (Selleckchem, Cat#S8935) or vehicle (DMSO). After incubation for 24 h, 10 μ l of MTT (Sigma) solution (5mg/ml) was added to each well, and the plates were then incubated for 3 h at 37 °C. The growth medium was then replaced with 150 μ l of DMSO per well, and the absorbance at 540 nm was measured using a BioTek Cytation 5-imaging multimode reader (Agilent, USA).

Cell treatment and multiplex luminex assay

To determine the optimal concentration of BAY-985 effectively inhibiting the expression of *IL6*, NHEK cells (passage of 2) were pretreated with LPS (40 ng/ml) for 1 h and continually treated with varied concentrations of BAY-985 for an additional 7 h. Cells were collected and quantitative qPCR was performed to analyze the mRNA expression of *IL6*. Healthy or HS primary basal keratinocytes (passage of 2) were treated with the above-

defined concentration of BAY-985 for 24 h. Cells were lysed in RIPA lysis buffer and subjected to multiplex luminex assay using the Milliplex human cytokine/chemokine/growth factor panel A magnetic bead kit (Millipore, Cat# HCYTA-60K-PXBK38) according to the manufacturer's instructions. Standard controls for each protein were prepared by serial dilution and run in duplication. 25 ul of cell lysates for each indicated group were used for analysis on Luminex Magpix Multiplex System (Luminex Corporation, USA) equipped with xPONENT (v. 4.2) software was used to perform the multiplexed assays.

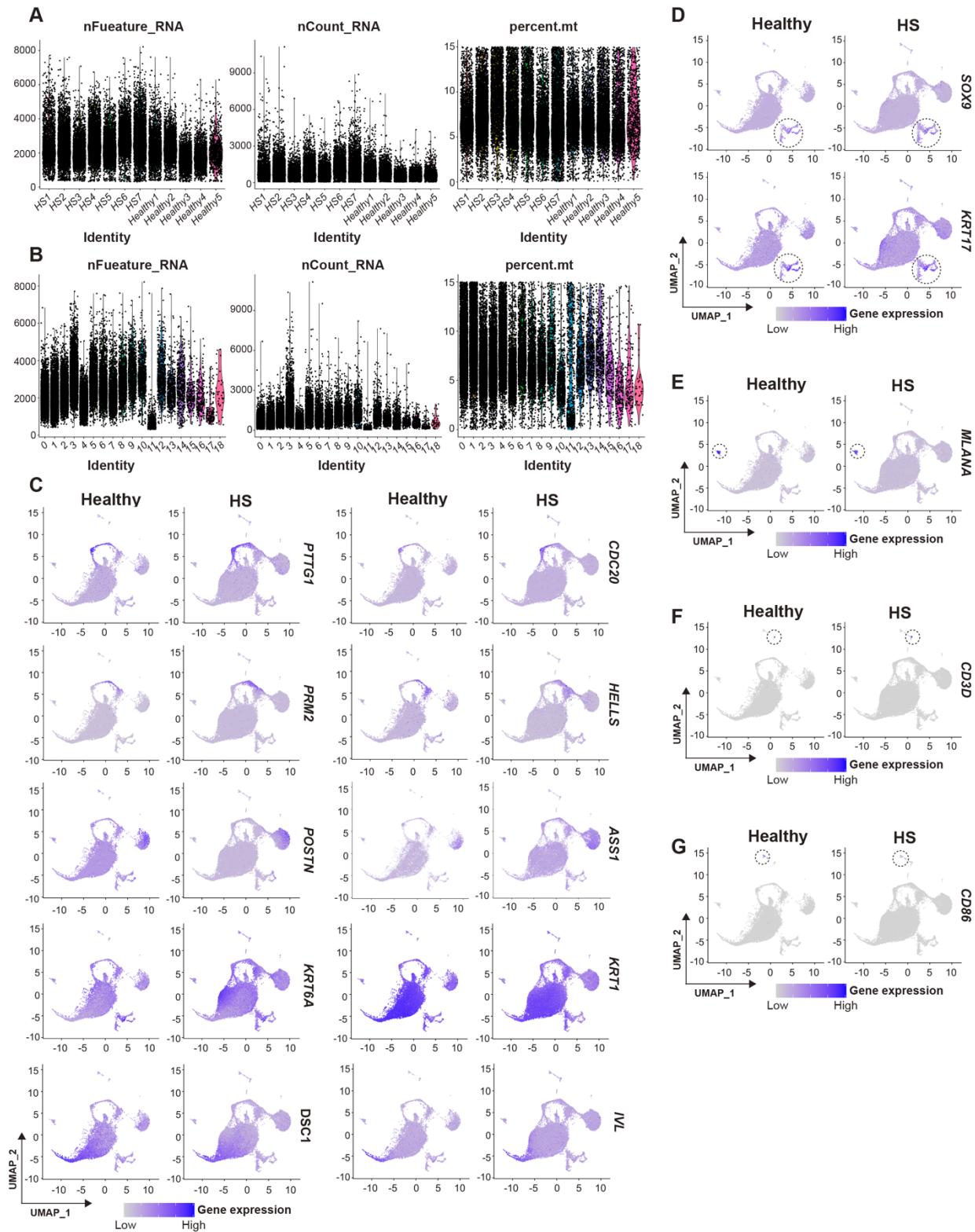
Statistical analysis

All statistical analyses were performed using GraphPad Prism 9 (GraphPad Software, La Jolla, California) or determined by the default algorithm of R pipelines. Unless otherwise specified, all values are expressed as mean \pm standard deviation (SD). Statistical significance between groups was determined using an unpaired Student t-test assuming two-tailed distribution or Chi-squared test. Multiple independent experiments were performed to verify the reproducibility of all experimental findings. A *P* value <0.05 was considered significant unless specified.

References

1. Jin S, et al. Inference and analysis of cell-cell communication using CellChat. *Nature Communications*. 2021; 12(1):1088.
2. Li R, Grimm SA, Wade PA. Low-input ATAC&mRNA-seq protocol for simultaneous profiling of chromatin accessibility and gene expression. *STAR Protoc.* 2021; 2(3):100764.
3. Fulco CP, et al. Activity-by-Contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat Genet.* 2019; 51(12):1664-1669.

Supplementary Figure1



Continued Supplementary Figure 1

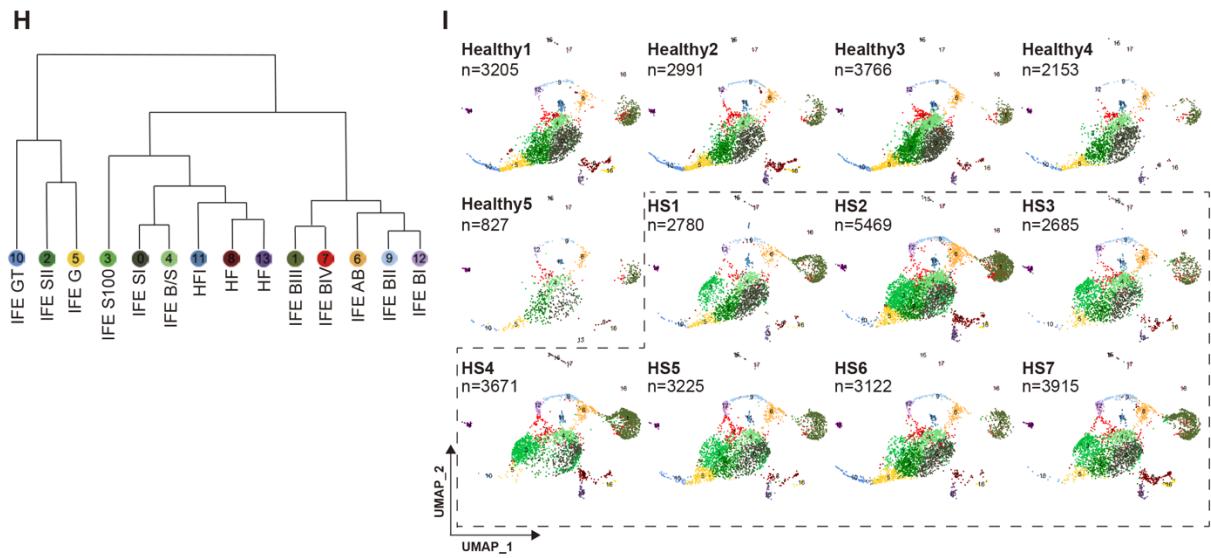


Figure S1. Quality control of individual cell capture and cell subtype identification in healthy and HS scRNA-seq cohorts. (A and B) The number of genes detected in each cell (nFeature-RNA), total molecules detected within a cell (nCount-RNA), and fraction of reads mapping to mitochondria (percent.mt) across recovered samples (A) or cell identity (B). (C) Log-normalized expression of markers for IFE BI (*PTTG1* and *CDC20*), IFE BII (*RRM2* and *HELLS*), IFE BIII (*POSTON* and *ASS1*), IFE BIV (*KRT6A*), IFE S (*KRT1*), and IFE G (*DSC1* and *IVL*). (D) Log-normalized expression of markers for HF (*SOX9* and *KRT17*) as shown in circled clusters. (E, F, and G) Log-normalized expression of markers for melanocyte (*MLANA*) (E), T cell (*CD3D*) (F), and B cell (*CD86*) (G), as shown in the circled cluster, respectively. (H) Hierarchical clustering of gene expression for interfollicular epidermis (IFE) and hair follicle (HF) cell clusters based on data from (B). (I) Original UMAP split by individual libraries for healthy and HS samples, respectively. Cell numbers were also shown in each library.

Supplementary Figure 2

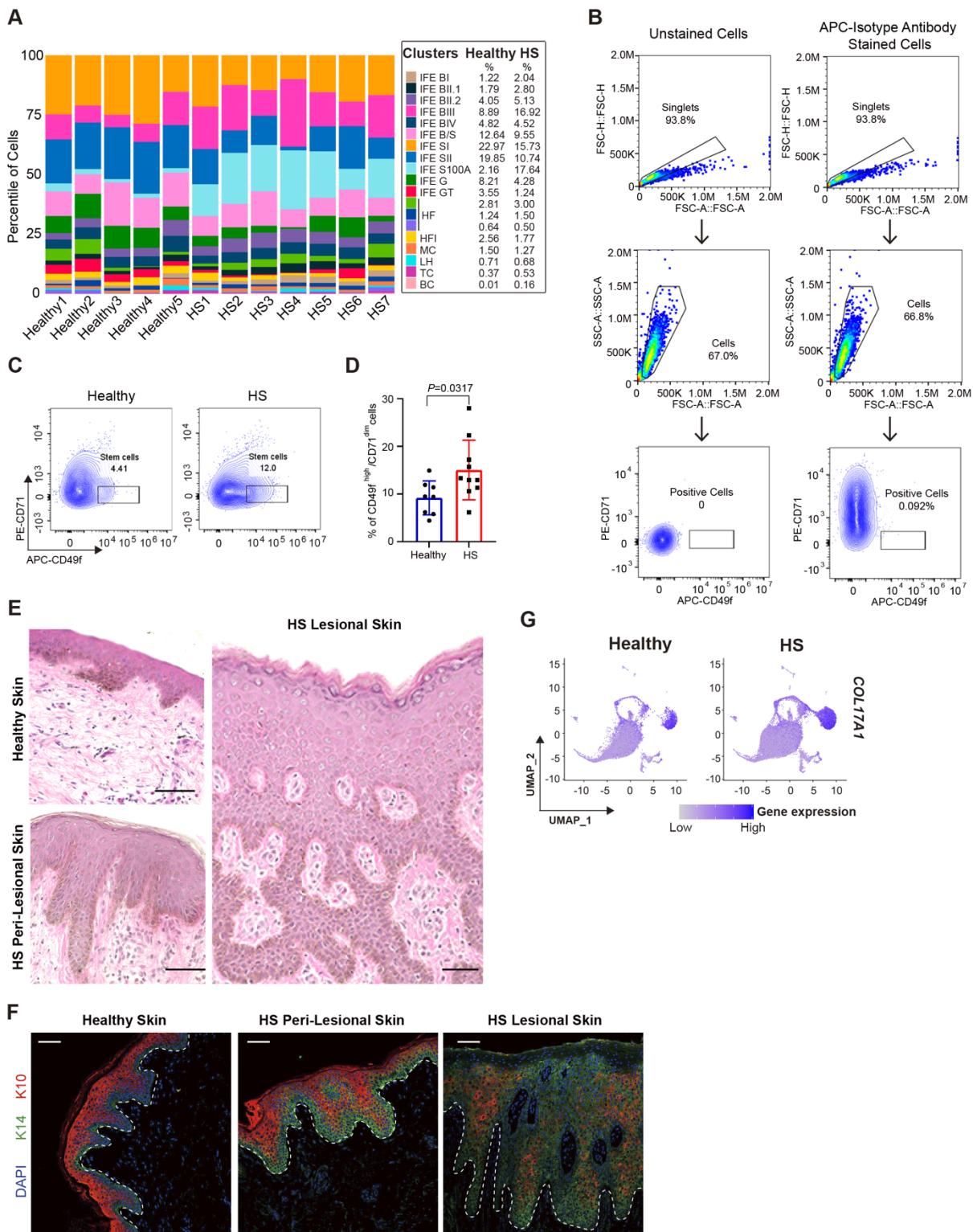


Figure S2. The proliferation of basal stem/progenitor cells in HS lesional skin. (A) Histogram depicting individual cluster fractions of total with healthy epidermis (n=5) or HS epidermis (n=7). (B) FACS strategy for analyzing CD49f^{high}/CD71^{dim} keratinocytes. (C and D) Representative graphs for flow cytometric analyses of CD49f and CD71 protein expression on epidermal cells within healthy and HS samples (C). Quantification of the number of CD49f^{high}/CD71^{dim} cells (n=8 for healthy, n = 10 for HS) (D). (E) Representative H&E staining of healthy, HS peri-lesional, and lesional skin. (F) High-resolution imaging of healthy and HS skin biopsies immunostained for K14, K10, and DAPI. The dashed lines demarcate epidermal–dermal boundaries. (G) Log-normalized expression of COL17A on UMAPs derived from sc-RNA seq datasets in healthy and HS conditions. Magnification, x20 (E and F). Scale bars, 50 µm (E and F).

Supplementary Figure 3

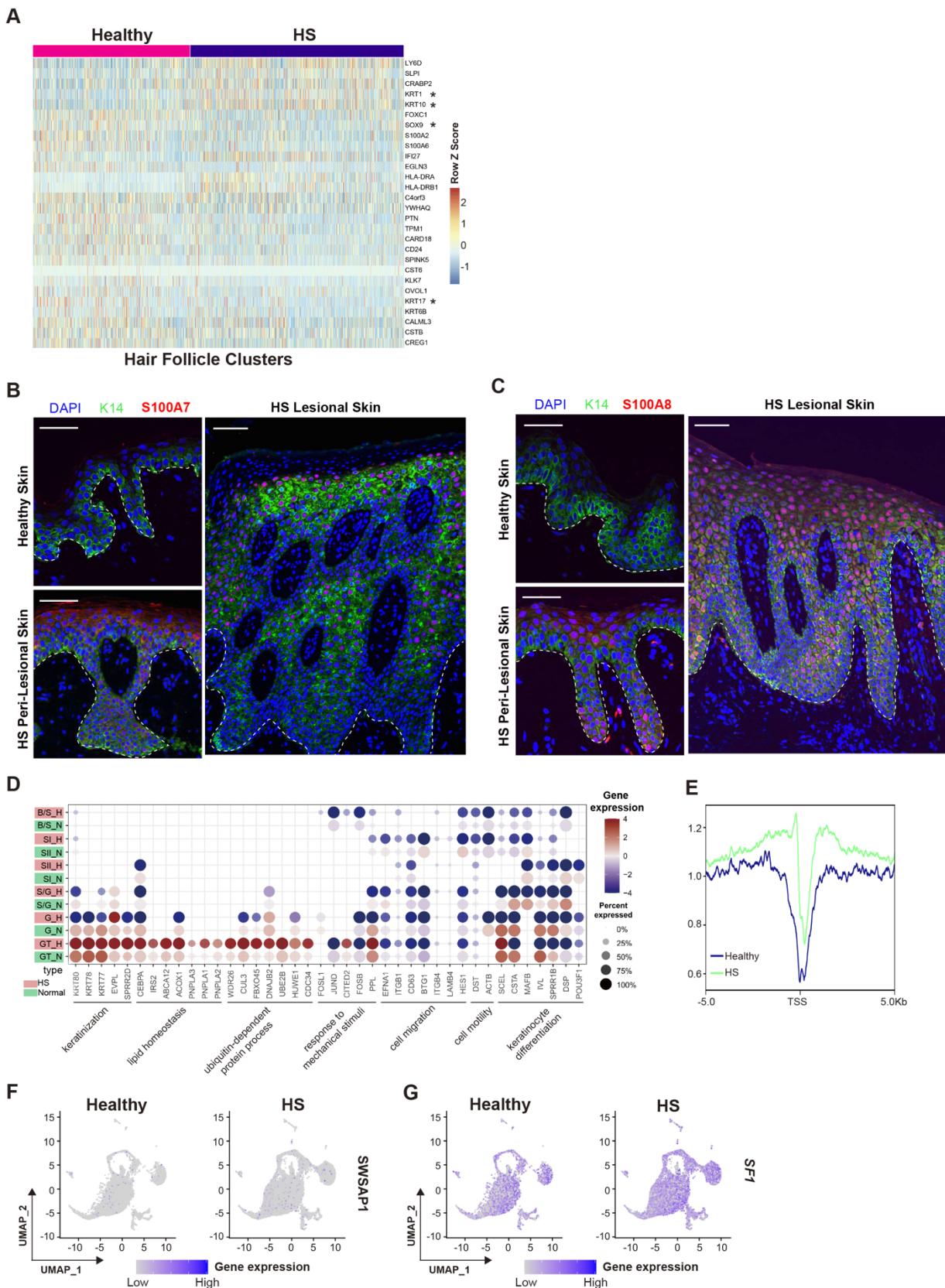


Figure S3. S100A cluster specifically presenting in HS lesional epithelium. (A) The heatmap displaying scaled expression of the top DEGs in individual cell in healthy and HS HF clusters, respectively. (B and C) High-resolution images of healthy and HS skin biopsies immunostained for S100A7 (B), S100A8 (C), as well as K14 and DAPI. Magnification, x20. Scale bars, 50 µm. (D) Dot plot showing representative genes from indicated GO functions across suprabasal clusters in healthy and HS conditions. (E) Distribution of S100A8 CUT&RUN peaks on its target loci (from-5 kb of TSS to +5kb) in healthy and HS epithelial cells. Normalized CUT&RUN signals are shown on the Y-axis. (F and G) The UMAP displaying the scaled expression of *SWSAP1* and *SF1* in healthy and HS clusters.

Supplementary Figure 4

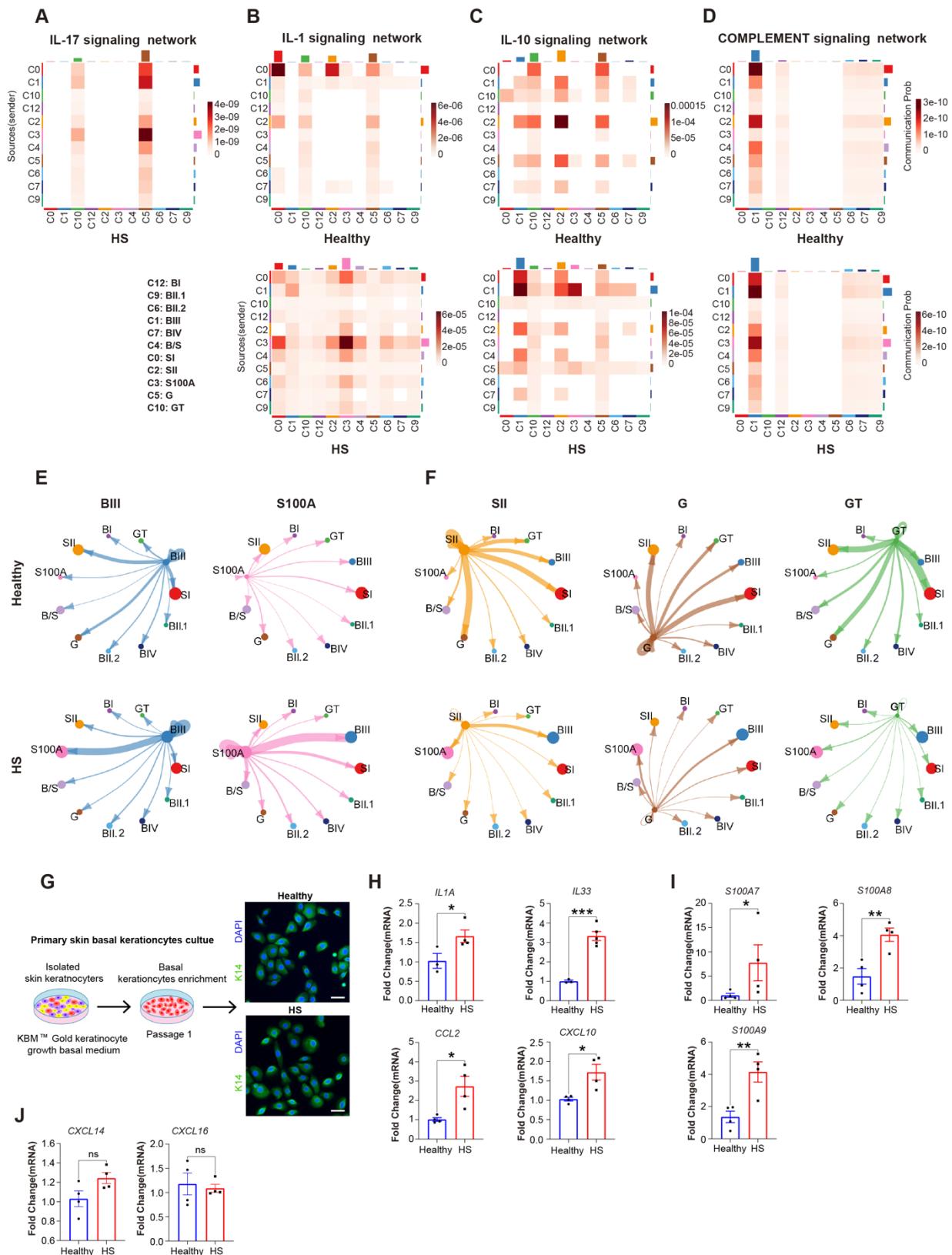


Figure S4. Altered inflammatory signaling strength of putative cell communications in HS disease. (A, B, C and D) Heatmap showing the differential value of possible interactions of the indicated inflammatory signaling between any two cell populations. The X-axis represents the receiver cluster and the Y-axis represents the sender cluster. (E and F) Circle plots displaying the inferred overall signaling communication from the indicated sender cluster to other receiver clusters. (G) *Left*, schematic illustration of the strategy for culturing primary basal keratinocytes. *Right*, immunofluorescence staining of K14 and DAPI in healthy and HS basal keratinocytes. (H to J) qRT-PCR was used to quantify the relative mRNA levels of indicated genes between healthy and HS basal keratinocytes. *P*-values based on unpaired Student t-test. Error bars indicate the mean ± SD from $n = 4$ biological replicates. The experiment was repeated three times independently with similar results. ns, nonsignificant.

Supplementary Figure 5

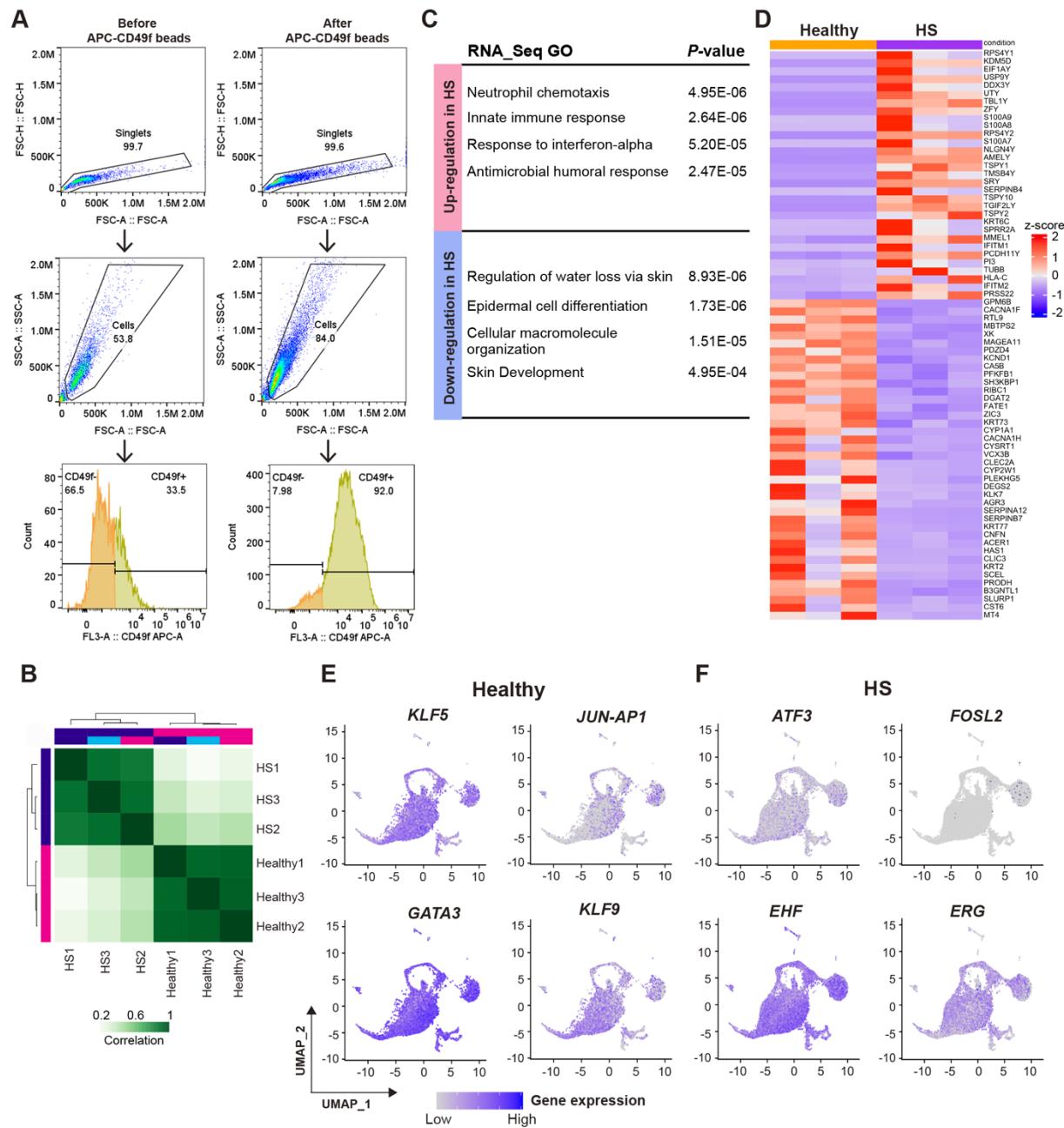


Figure S5. Alterations of the transcriptomic profile in HS CD49f^{high} basal cells. (A) FACS strategy for analyzing CD49f+ keratinocytes before and after APC-CD49f microbeads enrichment. (B) Heatmap clustering of correlation coefficients across all indicated ATAC-seq profiles. (C) Top-scoring GO analysis of DEGs in HS CD49f^{high} basal cells. (D) Heatmap of the top DEGs that were dysregulated in HS CD49f^{high} basal cells. (E and F) Log-normalized expression of markers for representative TFs enriched in healthy and HS conditions, respectively.

Supplementary Figure 6

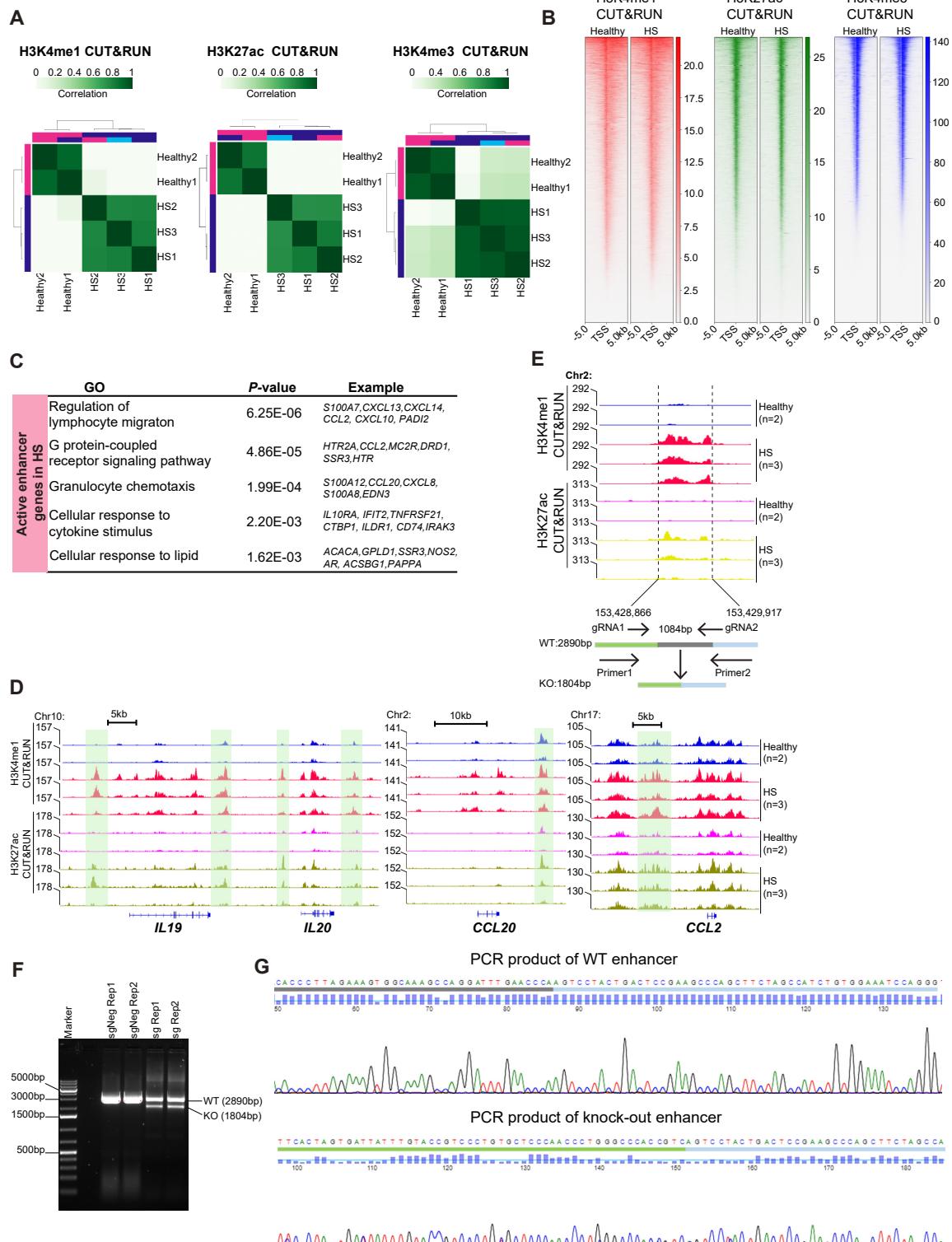


Figure S6. Genetic perturbation abrogating the enhancer on *S100A* family genes loci. (A) Heatmap clustering of correlation coefficients across all indicated CUT&RUN profiles. (B) Heatmaps of merged CUT &RUN H3K4me1, H3K27ac, and H3K4me3-seq peak profiles within 5 kb of the TSS across healthy and HS CD49f^{high} basal cells. CUT&RUN-Seq was performed in biological replicates (n= 2 for healthy, n=3 for HS). (C) Top-scoring GO analyses showing active enhancer genes derived from CUT&RUN peaks. (D) Genome browser views confirming specific histone modification markers enrichment on *IL19*, *IL20*, *CCL20*, and *CCL2* loci, in healthy vs. HS CD49f^{high} basal cells. Green boxes highlighting the greater accumulated signaling over the putative enhancers in HS samples than healthy counterparts. (E) An experimental strategy targeting the enhancer of *S100A7A/8/9* loci using CRISPR/Cas9 plus dual gRNAs flanking ~1 kb central region. Primers 1 and 2 are oligonucleotides for PCR confirmation of depletion. (F) Gel electrophoresis of PCR products on genomic DNA of NHEK cells expressing negative control sgRNAs and E3 sgRNAs. The experiment was repeated three times independently with similar results. (G) Sanger Sequencing to confirm the depletion of E3 enhancer.

Supplementary Figure 7

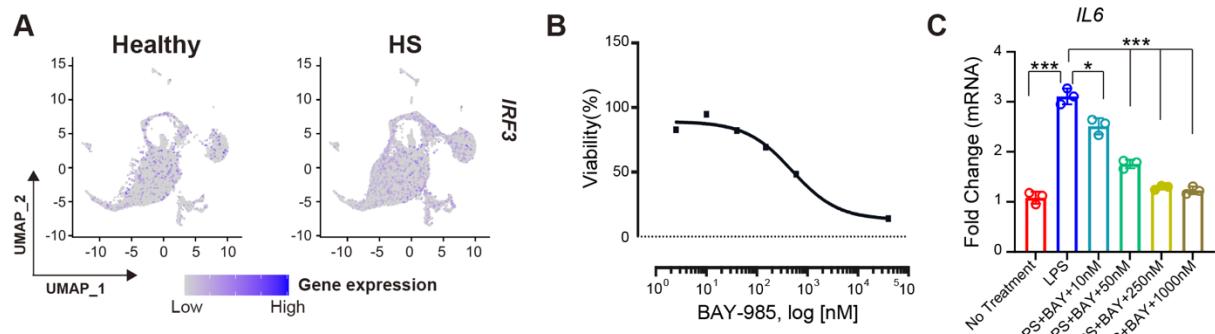


Figure S7. Functional inhibition of IRF3 has an effect on the production of HS inflammatory genes. (A) Log-normalized expression of markers for *IRF3* as shown in healthy and HS, respectively. (B) Cytotoxicity of BAY-985 on NHEK cells for 24 h as determined by the MTT assay. X-axis representing the concentrations in a logarithmic scale. (C) Evaluation of the relative *IL6* mRNA levels by qRT-PCR within the BAY-985 treated groups. *P*-values based on unpaired Student t-test. Error bars indicate the mean \pm SD from $n = 3$ replicates. The experiment was repeated three times independently showing similar results. * $P < 0.05$, *** $P < 0.001$.

Table S1. Description of the patient information and their respective sample for various experiments included in this study.

Condition	Gender	Race	Body site	Experiments
Healthy 1	F	Caucasian	Breast	scRNA,IF,FC,ATAC,CUT&RUN
Healthy 2	F	Black	Breast	scRNA,IF,FC,ATAC,CUT&RUN
Healthy 3	F	Caucasian	Breast	scRNA,IF,FC
Healthy 4	F	Caucasian	Breast	scRNA,IF,FC
Healthy 5	M	Black	Abdomen	scRNA,IF,FC,ATAC
Healthy 6	F	Black	Breast	IF,FC
Healthy 7	F	Black	Breast	FC
Healthy 8	F	Black	Breast	FC
HS 1	M	Black	Axilla	scRNA,IF,FC
HS 2	F	Black	Inframammary	scRNA,IF,FC,ATAC
HS 3	F	Black	Axilla	scRNA,IF,FC
HS 4	M	Caucasian	Axilla	scRNA,IF,FC,ATAC,CUT&RUN
HS 5	F	Black	Axilla	scRNA,FC,CUT&RUN
HS 6	M	Black	Axilla	scRNA,FC
HS 7	M	Black	Groin	scRNA,IF,FC,ATAC,CUT&RUN
HS 8	F	Black	Axilla	IF,FC
HS 9	F	Black	Groin	FC
HS 10	M	Caucasian	Axilla	FC

Abbreviation: scRNA: scRNA-sequence; IF:immunofluorescence; FC: flow cytometry;
ATAC: ATAC sequence;CUT&RUN: CUT&RUN sequence

Note: HS patients above mentioned were Hurley late stage II/III and failure to drug treatment before surgery.