UNIVERSITY OF CALIFORNIA,

IRVINE

Interfollicular Epidermal Differentiation is Gradualistic Rather than Stepwis with GRHL3 Controlling Progression from Stem to Transition Cell States

DISSERTATION

submitted in partial satisfaction of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in Cellular & Molecular Biosciences

by

Ziguang Lin

Dissertation Committee:
Professor Bogi Andersen, M.D., Chair
Assistant Professor Kai Kessenbrock, Ph.D.
Associate Professor Angela Fleischman, M.D., Ph.D.

2020

Table of Contents

List of Figures

Acknowledgements

First and foremost, I would like to thank Dr. Bogi Andersen for his mentorship and guidance. He had tremendous faith and patience in me and supported me in letting me determine the direction of my research. Thanks to him I have become more mature as a scientist and a person. Next, I would like to thank Dr. Zhengquan Yu for guiding me through a productive and successful collaboration with his laboratory. I would also like to thank my committee members for their guidance and suggestions that improved my project.

I would like to thank my mother for her unconditional love and support throughout life. It would not have been possible for me to achieve this without her unwavering encouragement and support.

I am also grateful for my family, friends and lab-mates for their support and company. They have supported me through good and rough times.

Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Number T32CA009054. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health

Vita

Ziguang Lin

Education

University of California, Irvine 2014-2020
Ph.D. Cellular Molecular Biology; GPA 4.0
NIH T32 Cancer Biology Research Fellowship recipient

University of California, Los Angeles 2007**-**2009
Los Angeles, CA
B. S. Molecular Cellular Developmental Biology; graduation with honors; GPA 3.81

Research Experience

University of California, Irvine
Irvine, CA

Ph.D Researcher
2014 **-** present

Bogi Andersen Lab – mouse epidermal differentiation and intestinal regeneration.

Use single cell RNA-seq (scRNA-seq) to study mouse epidermal stem cell differentiation, intestinal regeneration.

Use immunofluorescence, microscopy, single molecule RNA-FISH to visualize specific RNA/protein targets.

Discovered the role of transcription factor GRHL3 in a continuous epidermal differentiation process and in regulating the Wnt pathway.

Discovered that Msi1+ cells in the intestinal crypt to be highly proliferative cells that contribute in a major way to intestinal regeneration following irradiation independent of Lgr5+ cells.

Use ChIP-seq (H3K27Ac, K4Me1) on cultured human keratinocytes induced to proliferate, migrate or differentiate to characterize the epigenetic landscape of each cell state, respectively; use ChIP-qPCR to validate ChIP-seq findings.

Use a variety of tools in R, Linux, python, MATLAB to analyze high-throughput sequencing data (QC, alignment, differential gene expression, peak calling, super enhancer calling, motif enrichment analysis, PCA, tSNE, single cell analysis, clustering, pseudotime, RNA-velocity, etc.

City of Hope National Medical Center

Research Associate II 2010 - 2014

Center for Biomedicine & Genetics – development and optimization of industrial scale stem cell manufacturing, differentiation, and banking process.

Routine manufacturing of large scale hESC, cardiomyocytes, neural stem cells and retino-pigmented epithelium cell banks under cGMP standard.

QC/QA of products via FACS, qPCR, ELISA, MycoAlert etc.

Development of a scalable, xeno-free, enzyme-free, GMP-compliant hESC culture system.

University of California, Los Angeles Los Angeles, CA

Laboratory Assistant
2007 - 2009

Siavash Kurdistani Lab – cellular levels of epigenetic aberrations in cancer

Use FACS and immunofluorescence to establish a correlation between phase of cell cycle and cellular levels of histone modifications.

Publications

Lin Z, Jin S, Chen J, Li Z, Lin ZQ, Nie Q, Andersen B. Interfollicular epidermal differentiation is gradualistic rather than stepwise with GRHL3 controlling progression from stem to transition cell states., Nature Communications, in review

Sheng X#, **Lin Z**#, Lv C#, Shao C, Bi X, Deng M, Xu J, Guerrerro-Juarez C, Li M, Wu X, Zhao R, Liu X, Wang Q, Nie Q, Cui W, Gao S, Zhang H, Liu Z, Plikus M, Lengner C, Andersen B, Ren F*, Yu Z*. Rapidly-cycling stem cells mediate Lgr5-high cell-independent intestinal regeneration., Nature Communications, in review

Herndon-Klein R, **Lin Z**, Hopkin A, Gordon W, Tsoi LC, Liang Y, Gudjonsson J, Andersen B. GRHL3 chromatin binding and super-enhancer landscape are reorganized in different functional states of epidermal keratinocytes. 2017. *PLOS Genetics*.

Herndon-Klein R, Hu W, **Lin Z**, Nguyen T, Andersen B. Identification of chromatin regulatory domains in human corneal epithelial cells links enhancers to epithelial diseases and defines a role for KLF7 in corneal epithelial progenitors. 2017. J Biol Chem.

Tompkins, J. D., Jung, M., Chen, C., **Lin, Z**., Ye, J., Godatha, S., Riggs, A. D. (2016). EBioMedicine Mapping Human Pluripotent-to-Cardiomyocyte Differentiation : Methylomes , Transcriptomes , and Exon DNA Methylation " Memories ." *EBIOM*, *4*, 74–85. https://doi.org/10.1016/j.ebiom.2016.01.021

Chen, V. C., Ye, J., Shukla, P., Hua, G., Chen, D., **Lin, Z**., … Couture, L. A. (2015). Development of a scalable suspension culture for cardiac differentiation from human pluripotent stem cells. *Stem Cell Research*, *15*(2), 365–375.

Chen, V. C., Couture, S. M., Ye, J., **Lin, Z**., Hua, G., Huang, H. I. P., … Couture, L. A. (2012). Scalable GMP compliant suspension culture system for human ES cells. *Stem Cell Research*, *8*(3), 388–402

Abstract of the Dissertation

Interfollicular Epidermal Differentiation is Gradualistic Rather than Stepwis with GRHL3 Controlling Progression from Stem to Transition Cell States

Ziguang Lin

Doctor of Philosophy in Cellular & Molecular Biosciences

University of California, Irvine, 2020

Professor Bogi Andersen, Chair

Although interfollicular epidermal (IFE) differentiation is thought to be stepwise as reflected in sharp boundaries between the basal, spinous, granular and cornified layers, this prediction has not been studied at a single cell resolution. We used single cell RNA-seq to show that IFE differentiation is best described as a single step gradualistic process with a large number of transition cells between the basal and spinous layer. RNA-velocity analysis identifies a commitment point that separates the plastic basal and transition cell state from the unidirectionally differentiating cells. We also show that GRHL3, best known for promoting IFE terminal differentiation, has a major function in suppressing epidermal stem cell expansion and the emergence of an abnormal stem cell state by suppressing Wnt signaling in stem cells.

Chapter 1. Background

Epidermis of the skin

The epithelium is composed of continuous sheets of tightly packed cells that lines the surface (ie. skin) and the internal cavities of the body. It acts to mediate the interaction between the organism and the environment. Some examples of epithelial tissues are epidermis of the skin, intestine epithelium, mammary gland and cornea.

The skin is composed of the upper epidermis and the lower dermis layers separated by a basement membrane. The skin is essential for an animal's survival by providing a water-tight physical barrier, thermo-regulation, defense against radiation and infections, as well as tactile feedback. In the epidermis of the skin, such diverse biological functions are compartmentalized in specialized cell types of distinct embryonic origins. During development, the epidermis emerges from the surface ectoderm as a single layer of unspecified progenitor cells, which later forms the stratified inter-follicular epidermis (IFE), the hair follicles (HF), the sebaceous glands (SG), and Merkel cells that enables tactile sensing. Meanwhile the epidermis is also embedded with mesoderm-derived cells: T-cells, Langerhans cells, as well as melanocytes that renders skin/hair colors and UV protection. During homeostasis, the epidermal barrier is maintained by mitotically active keratinocyte stem cells in the basal layer of the stratified epidermis. The basal cells progressively differentiate and migrate supra-basally as they form the spinous and granular layers. Eventually, the terminally

differentiated cells cease transcriptional activities and die; becoming flattened, de-nucleated and form the water-tight cornified barrier that is shed continuously.

Epidermal Structure

Four morphologically distinct layers of cells are visible through hematoxylin-eosin staining. The basal proliferating cells contact dermis through a specialized layer of extracellular matrix (ECM) called the basement membrane. Hemidesmosomes and integrin-based adhesions link the basal cells and the ECM while desmosome junctions are found throughout basal, spinous and the granualar layers between the cells. Adherens junctions are found in the basal and the spinous layer cells. The keratinocyte stem cells replenish the epidermis during homeostasis and regeneration. The cytoarchitechture changes depending on differentiation state of the cells to support specific functions of each layer. For example, increase in adhesions is observed in stratum spinosum; cornified envelope is assembled in the granulosum layer; while the water-tight cutaneous barrier is formed in the stratum corneum layer. Distinct cytokeratins are expressed in different layers. For instance, cytokeratin 5 (K5) and cytokeratin 14 (K14) are only expressed in the stratum basale while cytokeratin 1 (K1) and cytokeratin 10 (K10) are expressed in the spinosum and granulosum layers. Where as Loricrin, Filagrin are only expressed in the cornified layer. Some junctional proteins are also found to be expressed in a graded manner. DSC1 and DSG4 level is the highest in the granulosum layer and drops gradually until gone in the stratum spinosum layer. While DSG3, DSC3, DSC2 levels are

highest in the stratum basale and decreases gradually toward the stratum granulosum layer.

Epidermal Development

The ectoderm gives rise to the IFE, hair follicle, sebaceous gland, and Merkel cells while the mesoderm gives rise to the connective tissues of the dermis. A single layer of epithelial cells at embryonic day 9.5 (E9.5) starts to stratify into multiple layers by E14.5. The sub-ectodermal mensenchymes are thought to send signals to initiate epidermal stratification. The proliferative intermediate layer and the transient protective endodermis-like periderm forms by E14.5. The intermediate layer divide and mature into the spinous layer E15.5. Both the intermediate layer and the spinous layer express K1 at this time point. K1 expression is to be induced by Notch signaling[1], which is thought to be downstream of Np63[2]. And spinous cells continue to differentiate, mature and forms the granular layer by E16.5. The water-tight cornified cell envelope is intact by E18.5. Increase in extracellular Ca^{2+} concentration and the corresponding intracellular Ca^{2+} sensing machinery induces the formation of granular and the cornified layers. At the end of epidermal development, transcription factors Klf4 and Grhl3 are essential for skin barrier formation[3,4].

The embryonic ectoderm can also give rise to the nervous system. The choice between epidermal and neural fate depends on the balance of Wnt, FGF and BMP signaling. In the absence of Wnt signaling shortly after gastrulation, the ectodermal cells develop toward a neural fate in response to upregulated FGF signaling and down regulated BMP signaling[5]. Whereas when Wnt signling is

present, the ectodermal cells develops an epidermal fate as a result of upregulated BMP signaling[6].

Basal Stem Cell Heterogeneity and Self-Renewal

Studies on stem cells in the basal layer of the adult epidermis have suggested cellular heterogeneity within this layer, although the nature of this heterogeneity remains controversial. The hierarchical model of stem cell differentiation suggests that the basal layer comprises rare slow-cycling long-term stem cells and their fast-cycling committed progenitors[7]. Another related model suggests that slow-cycling and fast-cycling stem cells occupy distinct regions of the basal layer and renew within their respective regions[8,9]. The stochastic model of stem cell differentiation suggests that all basal cells are equivalent in terms of their stemness and that they differentiate in a stochastic manner[10–12]. All these models share the notion that basal cell divisions are asymmetric; one progeny gives rise to self-renewing progenitor while the other is committed for differentiation. Lineage tracing experiments in mice using *Axin2* has shown that epidermal stem cells in the basal layer competes neutrally and require Wnt/β-catenin signaling to proliferate[11]. These stem cells regulate self-renewal in an autocrine manner by producing Wnt ligands and inhibitors.

Transcriptional Regulation of Inter-follicular Epidermal Differentiation

To date, we know that the gene expression programs underpinning epidermal terminal differentiation process is orchestrated by a number of key transcription factors such as *Trp63, Maf/Mafb, Klf4, Znf750, Prdm1 and Grhl3*[13].

GRHL3 (also referred to as GET1) is a member of the highly conserved CP2-like transcription factor family. In *Drosophila,* GRHL3 plays essential roles in cuticle formation and repair. In mice, *Grhl3* is expressed in the surface ectoderm during embryogenesis and in the supra-basal layer of the epidermis in adulthood. *Grhl3* is required for epidermal differentiation and barrier formation in a variety of species[4,14–16]. Mutation in *Grhl3* results in defective barrier formation and hyperplasia of the epidermis[4]. Mice lacking *Grhl3* die at birth due to dehydration and spina bifida. Furthermore, *Grhl3* has been shown to act as a tumor suppressor gene in squamous cell carcinoma (SCC) by directly promoting the transcription of PTEN[17,18]*. Krt14*-driven deletion of *Grhl3* in adult mouse epidermis results in wound healing defect[19] and renders the skin more susceptible to chemically-induced carcinogenesis via aberrant activation of *PI3K/AKT/mTOR* signaling[17]. GRHL3 can also indirectly regulate its potential target genes by repressing the expression of the microRNA miR-21[18], which is expressed in skin and upregulated in diseases such as psoriasis, atopic dermatitis, and squamous cell carcinoma. Keratinocytes with mutation in *Grhl3* has increased miR-21 expression and exhibit enhanced Ras-mediated tumorigenesis[18]. High-throughput gene expression/chromatin profiling has revealed that *Grhl3* transcriptionally activate genes for cell-adhesion, lipid production, cornified envelope formation and protein crosslinking in mouse skin[4,15].

Single Cell RNA-seq Rationale

Cells are fundamental units of life. Multicellular organisms are made of various cell types. Though the conventional notion of cell type is clear, a more comprehensive and rigorous definition is still lacking. Superficially, cells can be distinguished by size and shape. At the molecular level, cell types can also be characterized by distinct surface proteins.

The cells from the epidermis have been studied using microscopy and fluorescence-activated cell sorting (FACS) approaches. These methods provide high spatial and cellular resolution information but are limited in that only a few molecular markers can be visualized at the same time thus precludes high-throughput, quantitative observations. On the other hand, methods such as bulk RNA-seq or microarray allows global measurement of population-average gene expression; and are thus applied to genome-wide screening of regulated genes and the detection of compositional changes between conditions. However, bulk approaches have limitations because they mask information from rare subpopulations of cells.

In the past few years, scientists were able to integrate high-throughput genomics with microfluidics technology to allow gene expression profiling of large numbers of individual cells in single-cell RNA-seq (scRNA-seq), combining the power of quantitative, high-throughput genomics with microscopic resolution of individual cells[20]. The first single cell RNA-sequencing experiment was published in 2009[21]. A few years later, a data set of 1.3 million cells was released by 10X Genomics. A unique advantage of scRNA-seq is the ability to define cell types through unsupervised clustering based on transcriptomes. Instead of relying

strictly on a few conventionally-defined markers to define cell types, this provides unique opportunities for systematic, data-driven definition of cell identity. scRNA-seq can also provide new perspectives about temporal dynamics, spatial organization and ultimately molecular mechanisms that control differentiation[22]. It is important to be also aware that scRNA-seq can directly measure only some cells of a population and only a fraction of the RNA molecules in each cell.

Single Cell RNA-seq Analysis Overview and Clustering strategies

In a typical experiment, unreliable cells and potential doublets are removed through quality control. The data is then normalized in terms of sequencing depth and other aspects that can potentially be confounding. Because of the large number of genes assayed in scRNA-seq, the high dimensional distances between cells tend to be too small to identify cell groups in a robust manner. Therefore, the most informative genes and features (ie. genes with the highest variances) are then selected to reduce background noise[23]. Cell to cell distances are then calculated at lower dimensional space and used for downstream graphing and clustering. Euclidean distance, cosine similarity, Pearson's correlation and Spearman's correlation between cells can all be used to calculate cell to cell distances[24].

The most popular clustering algorithm is k-means clustering. This method identifies k centroids in the data and place cells to the nearest centroid. This can be applied to large data sets because it scales linearly with the number of points. But this method is prone to identify equal-sized clusters which can mask rare cell

types. Some of the new methods such as RaceID[25] and SIMLR[26] can overcome this disadvantage. The hierarchical clustering strategy sequentially groups cells into larger clusters or divide clusters into smaller subsets. This method is costly in terms of computational resource. New methods such as CIDR[27] or BackSPIN[28] can address this problem.

Chapter 2. Interfollicular epidermal differentiation is gradualistic rather than stepwise with GRHL3 controlling progression from stem to transition cell states

Introduction

During development, the epidermis emerges from the surface ectoderm as a single layer of unspecified progenitor cells, which later form the stratified interfollicular epidermis (IFE), the hair follicles, the sebaceous glands, and the tactile sensing Merkel cells. Meanwhile, mesoderm-derived immune cells and neural crest-derived melanocytes take up residence in the epidermis. Although IFE heterogeneity and differentiation have been previously studied with single cell RNA-sequencing (scRNA-seq) in adult mouse and human skin[29–32], late development of the embryonic IFE has not been studied with scRNA-seq, and key gene regulators of IFE differentiation have not been systematically identified with this method.

Conventionally, IFE differentiation is thought to be stepwise as reflected in the IFE's four distinct layers separated by clear boundaries. Cells of the basal layer, which rest on the basal lamina, are mitotically active stem cells marked by the expression of keratin (K) 5 and K14. As the cells move to the spinous layer, they exit the cell cycle and K5 and K14 expression is replaced by the expression of K1 and K10. And as the cells advance to the granular layer, they turn on the expression of barrier forming genes, including loricrin and filaggrin[33]. Eventually, the terminally differentiated cells cease transcriptional activities and die, become

flattened and de-nucleated, forming the water-tight cornified layer that is eventually shed off.

The layered structure of the epidermis suggests that molecular switches could sharply alter the expression of many genes at the boundaries of distinct cell layers. This prediction, however, has not been studied at a genome-wide scale in the postnatal day (P) 0 mouse epidermis, which—similar to the human epidermis, but in contrast to the adult mouse epidermis—contains clearly demarcated epidermal layers.

Layer-restricted expression of transcription factors is thought to contribute to the IFE differentiation program. One such transcription factor, GRHL3, is expressed throughout the developing epidermis during embryogenesis, but becomes restricted to the most differentiate layers at P0. *Grhl3*-deleted mice have a defective epidermal barrier with genome-wide expression studies indicating that GRHL3 activates genes encoding cell-cell adhesion molecules, lipid producing enzymes, and proteins required for cornified envelope formation and crosslinking[4,34]. At P0, the *Grhl3-/-* IFE contains a disorganized basal layer, thickened spinous and granular layers, and a compacted cornified layer. Based on these studies, it has been assumed that the main embryonic role of GRHL3 is to promote full differentiation of cells of the granular layer. But the nature of the epidermal hyperplasia in the P0 *Grhl3-/-* IFE remains enigmatic.

To better understand epidermal differentiation, we investigated 85,286 single cell transcriptomes from mouse skin at embryonic day (E) 14.5, E16.5, and P0 in wild type (WT) and *Grhl3*-/- mice (Figure 1A). Our findings challenge the

classical notion of a stepwise IFE differentiation, where cells within a layer are thought to be similar but undergo dramatic changes as they move to the layer above. Rather, we find a high proportion of transition cells with a character intermediate between the basal and the first spinous layer and other features suggesting that IFE differentiation is best viewed as a single-step gradualistic process. RNA velocity analysis indicates that prior to the transition-differentiation cell state boundary, cell states are plastic, whereas after this commitment point, cells states proceed strongly in a unidirectional manner toward terminal differentiation. In *Grhl3*-/- mice we find a defective activation of the terminal differentiation, as expected. But unexpectedly, we find accumulation of epidermal stem cell populations and the emergence of proliferative cell states unique to the mutated epidermis. The expanded stem cell compartment shows increased *Wnt* expression, suggesting that GRHL3-expressing differentiated cells communicate to the basal stem cell compartment to suppress Wnt signaling and expansion of IFE stem cells.

Results

GRHL3 has a major effect on IFE cell states late in embryogenesis

To gain an overview of IFE development and to understand the role of GRHL3 in this process, we generated single cell suspension from WT and *Grhl3*-/- mouse dorsal skin (E14.5, E16.5) and epidermis (P0), followed by scRNA-seq[20,35]. We used canonical markers to computationally isolate E14.5, E16.5, and P0 IFE cells, and we represented them together in an unsupervised diffusion map[36,37] (Fig. 1B). In this map, the first diffusion component (DC1) separates IFE

cells by developmental age, from E14.5 to E16.5 to P0, whereas the second diffusion component (DC2) organizes the cells by their progress in epidermal differentiation at each respective age (Fig.1C,D). Although we observed no change in cell states between WT and *Grhl3*-/- IFE cells at E14.5 and E16.5, at P0, *Grhl3*-/- IFE cells are shifted toward an earlier developmental time point compared to the WT IFE (Fig. 1D). We conclude that GRHL3 acts between E16.5 and P0 to exert a major effect on IFE cell states.

scRNA-seq captures newborn mouse epidermal cell heterogeneity

As GRHL3 has the most striking effect on cell composition in the fully developed newborn IFE (P0) (Fig. 1D), we next set out to understand normal IFE differentiation at this developmental stage. At P0, the IFE has reached its maximum thickness with morphologically well-defined layers: basal, spinous, granular, and cornified layers. We generated single cell transcriptomes from the back epidermis of a P0 mouse, capturing 5,494 cells with 38,879 mean number of reads per cell and 2,388 mean number of genes per cell. Clustering identified 16 subpopulations of epidermal cells (Fig. 2A). Each cluster was annotated by marker genes that are known to be uniquely expressed in each cell type or cell state (Fig. 2B; Supplemental Fig. 1A-C). We identified all previously defined epidermal subpopulations of the adult epidermis[30]: IFE, hair follicles, sebaceous gland cells, Langerhans cells, T cells, melanocytes, and Merkel cells (Fig. 2A, B). Four adjacent clusters of 1,779 IFE cells were identified: two basal clusters (IFE.B1 and IFE.B2; 1,002 cells), a basal-suprabasal transition cluster (IFE.T; 350 cells), and a differentiated cluster (IFE.D; 427 cells). The population of

transition cells (IFE.T) is 20% of the all IFE cells, which is a surprisingly large fraction. The Gene Ontology (GO) category enrichment[38,39] of these marker genes indeed reflects the biological functions of each population (Supplemental Fig. 1D, E, F).

The P0 IFE contains a large number of cells transitioning between the basal and spinous layers

Next, we focused on the P0 IFE cells to better understand IFE differentiation through pseudo-temporal analysis. We applied Monocle[40], an algorithm that uses changes in gene expression between cells to learn in an unbiased manner the position of each cell in a biological trajectory. Monocle introduces the concept of pseudotime as an abstract measure of how far a cell has progressed in a process—here in the differentiation of an IFE stem cell in the basal layer to a terminally differentiated keratinocyte at the top of the epidermis. Monocle ordered the IFE cells along a linear trajectory without introducing significant bifurcations (Fig. 2C). This differentiation trajectory corresponds well with our cluster annotation from Fig. 2A in that the cells progress from IFE.B through IFE.T to IFE.D. To validate the pseudotime ordering, we assessed the expression of canonical markers in the trajectory (Fig. 2D), finding that *Col17a1/K5/K14* mark cells in the early trajectory, *K1/K10* in the middle-to-end trajectory, and *Flg/Lor* in the end of the trajectory. These results suggest that Monocle ordered the IFE keratinocytes in a biologically meaningful manner from the most undifferentiated to the most differentiated cells.

We noted high number of transition cells with falling K5/K14 expression and rising K1/K10 expression—characteristics that are intermediary between the basal and spinous layers. This finding is unexpected because of the apparent sharp boundary between the basal and spinous layer[41]. To validate the existence of the transition cells, we performed RNA FISH with *Krt14* and *Krt10* probes. Consistent with the scRNA-seq data, we observed a substantial number of double positive transition cells in the basal and first spinous layers (Fig. 2E). These results indicate that 1) epidermal differentiation has already started in the basal layer[42], 2) some spinous cells still retain basal characteristics, and 3) a relatively large population of transition cells exists at the basal-spinous boundary.

The IFE differentiates gradually in a single-step

To better understand differentiation at a molecular level, we identified 4,299 genes that are dynamically expressed over the pseudotime trajectory. K-medoid clustering grouped these genes into 6 gene modules with distinct expression patterns and biological functions (Fig. 2F)[38,39,43]. Module 1, high early and falls sharply, is enriched in genes associated with cell proliferation. Module 2, also high early but falls more slowly, is enriched in genes involved in transcription and mRNA processing. Module 3, peaks in the middle of the trajectory, is enriched in genes for ribosome biogenesis and protein translation. Module 4, peaks towards the end, is enriched in lipid biosynthesis genes. Module 5, also peaks toward the end but with a slower rise, is also enriched in lipid synthesis pathways and in cell adhesion genes. Module 6, a sharp peak at the end, is enriched in crosslinking and barrier genes.

Each module contains transcription factors with known functions in IFE differentiation that matches the overall gene functional category of that cluster. At early differentiation stages, transcription factors associated with the cell cycle and DNA repair are highly expressed, including *Trp53, Smad4, Hmgb1*, and *Hmgb2*. As differentiation commences, *Mafb* and *Myc* become highly expressed whereas late differentiation transcription factors such as *Klf4* and *Grhl3* become highly expressed at late and terminal differentiation stages. In addition, we have identified many candidate transcription factors with heretofore unknown roles in the IFE (Fig. 2F).

Although these findings correspond well with known features of IFE differentiation, they also point to an unusual relationship between cell proliferation and protein synthesis which are normally tightly coupled; in the IFE, peak gene expression for protein synthesis appears after peak expression of cell cycle genes. Furthermore, these data indicate that the majority of IFE differentiation genes exhibit expression that is not layer-specific with at least four out of six gene modules straddling different layers. This result and the observed high number of transition cells challenge the traditional model of stepwise IFE differentiation, suggesting that at a transcript level, IFE differentiation is continuous and gradualistic.

The P0 IFE basal cell layer contains two distinct cell populations

Next we sought to better decipher P0 basal IFE stem cell heterogeneity by examining the unique molecular signatures expressed in the two basal subpopulations (Fig. 2A, 3A). IFE.B1 cells (446 cells) express *H19*, *Wnt*4, and

Igf2 at a higher level than IFE.B2 cells, whereas IFE.B2 cells (556 cells) exclusively express *Fst*, *Dcn*, and *Sox4* (Fig. 3A, B). We performed principle component analysis (PCA) on only these two populations and found the above markers to be the genes that explain most of the variances in the first 2 PCs. In contrast, cell cycle genes do not contribute to variances in the first few PCs (Supplemental Fig. 2A). To confirm that active cell cycling does not play a major role in defining the two basal clusters, we also regressed out the cell cycle genes and re-clustered the basal IFE cells, which did not alter the results (Supplemental Fig. 2B, C, D). Therefore, we ruled out the possibility that these clusters are due to differences in cell cycle stages among the cells. Furthermore, IFE.B1 and B2 are not distinct stages of differentiation on the pseudotime trajectory (Fig. 2C, Supplemental Fig. 2E). In addition, we used an independent clustering method, SC3[44], to recluster all basal IFE cells; SC3 clustering results were very simiar to Seurat's, confirming the robustness of this finding (Fig. 3C).

Our data, then, highlights IFE basal cell heterogeneity, suggesting there are two basal stem cell populations in the P0 IFE. To validate our findings, we carried out RNA-FISH studies with *H19* and *Wnt*4 probes. We found that *Wnt*4-high and *Wnt*4-low cells are distributed heterogeneously within the IFE basal layer (Fig. 3D) and that high *H19* expression tends to co-localize with the *Wnt*4-high cells (Supplemental Fig. 2F). These findings indicate that the IFE.B1 and IFE.B2 cells intermingle in the basal layer and are not spatially distinct groups of cells. Our findings are in agreement with previous studies showing that WNT-

secreting stem cells play a central role in IFE self-renewal during homeostasis[11,45,46].

Differentiating IFE cells go through a proliferative transition state at E14.5

To understand how the P0 IFE cell heterogeneity develops and how GRHL3 alters cell states, we next examined the E14.5 and E16.5 IFE in greater detail. For E14.5, we captured a total of 23,642 cells (WT1=4,841, WT2=5,651, *Grhl3*-/-1=6,064, *Grhl3*-/-2=7,086). As we captured the whole skin, we observed other cell types in addition to epidermal cells (Supplemental Fig. 3A-F). Focusing on the IFE cells, we identified four distinct IFE populations at this stage: two basal populations (IFE.B1 and IFE.B2), one transition population (IFE.T) and one differentiated population (IFE.D) (Fig. 4A). Among these, E14.5_IFE.B1 expresses more proliferative marker genes than E14.5_IFE.B2 (Fig. 4B). The majority of the E14.5_IFE.T cells are highly proliferative, consistent with known features of the intermediate layer at this stage, suggesting that E14.5_IFE.T cells represent the intermediate layer at E14.5. We compared the gene expression levels of late differentiation genes *Klf4*, *Tgm3*, and *Lor* between WT and *Grhl3*-/- in the E14.5_IFE.D population and observed that there is a marked reduction in expression for these genes (Fig. 4C), consistent with GRHL3's role as a transcriptional activator of differentiation genes as early as E14.5. However, in contrast to the drastic IFE cell composition change in the P0 *Grhl3*-/- IFE, no change in cell composition is observed at this stage between the WT and the *Grhl3*-/- IFEs (Figs. 1D, and 4D).

The pseudotemporal IFE differentiation trajectory at E14.5 cells is a linear trajectory with a single branch (Fig. 4E). The branch, which consists mostly of proliferating E14.5_IFE.T cells, occurs at the boundary between transition (IFE.T) and differentiated cells (IFE.D) (Fig. 4E, Supplemental Fig. 3G). This finding corroborates the idea that at E14.5, IFE keratinocytes go through a state of rapid proliferation (corresponding to the intermediate layer) before fully differentiating. The WT and *Grhl3*-/- IFE differentiation trajectories were similar at this stage, consistent with unchanged cell composition in the *Grhl3*-/- epidermis at E14.5. Together, these data show that whereas *Grhl3* loss leads to decreased expression of terminal differentiation genes at this stage, it does not alter the cell composition of the IFE at E14.5.

scRNA-seq reveals IFE cell heterogeneity at E16.5

For E16.5, we captured a total of 26,354 cells (WT1=8,535, WT2=6,482, *Grhl3*-/-1=5,168, *Grhl3*-/-2=6,169). The cell heterogeneity of the whole skin is similar to that observed at E14.5, plus a substantial number of hair follicle cells (Supplemental Fig. 4A-F). Five distinct IFE populations were identified at this stage (Fig. 5A): two basal populations (IFE.B1, IFE.B2), one transition population (IFE.T), one differentiating population (IFE.D), and one terminally differentiated population (IFE.TD). One of the basal populations (E16.5_IFE.B2) expresses more proliferative marker genes than the other basal population (E16.5_IFE.B1) (Fig. 5B). Unlike E14.5 where most IFE.T cells are highly proliferative, a smaller subset of the E16.5_IFE.T is highly proliferative (Supplemental Fig. 4G). Pseudotime analysis for E16.5 IFE cells showed a single linear trajectory with

one minor branch consisting of proliferating IFE.T cells, occuring at the IFE.T/IFE.D boundary (Fig. 5C). Also, the expression of terminal differentiation genes is reduced in the *Grhl3*-/- IFE (Fig. 5D). Similar to E14.5, we observed no change in the IFE cell composition of the *Grhl3*-/- IFE at this stage (Fig. 5E). This is consistent with previous studies showing defective barrier but no epidermal thickening in the E16.5 *Grhl3-/-* IFE[4]. Since proliferating cells are not found in IFE.T cluster at P0, we conclude that proliferation of transition cells is initially high at E14.5, becoming progressively attenuated as the epidermis reaches developmental maturity. Therefore, proliferation of transition cells prior to differentiation is more characteristic for IFE differentiation at E14.5 and E16.5 than at P0.

GRHL3 is required for the formation of the most differentiated cells of the IFE and for tempering IFE stem cell number

Previous work on GRHL3 suggested that all layers of the IFE do develop in *Grhl3*-/- embryos, although the basal layer is disorganized and the spinous and granular layers appear thicker than normal[4,34]. Bulk gene expression measurement and other studies suggested that GRHL3 transcriptionally activates the terminal differentiation gene expression program in the granular layer[4,34]. The differentiation-specific function of GRHL3 is also consistent with the restricted expression of GRHL3 in the most differentiated layers of the IFE at this developmental stage (P0).

Next we tested how GRHL3 affects P0 IFE differentiation as detected by scRNA-seq. We performed scRNA-seq on single cell suspensions from the back-

skin epidermis of two WT and two *Grhl3-/-* littermates at P0, capturing a total of 29,796 cells (WT1=6,766, WT2=6,590, *Grhl3*-/-1=9,157, *Grhl3*-/-2=7,283) (Supplemental Fig. 5A, B). As expected, we re-captured the IFE cell heterogeneity depicted in Fig. 2 (Fig. 6A, B), including the basal keratinocyte heterogeneity observed in Fig. 3. But in these experiments, we identified a third, actively proliferating basal population (IFE.B3) (Fig. 6A, Supplemental Fig. 5C); this cluster most likely emerged due to increased total number of cells captured across the 4 samples compared to the experiment in Fig. 2. We also believe that the higher number of cells included in this experiment allowed the terminally differentiated (IFE.TD) IFE cells to form a distinct cluster.

We integrated all 4 samples[47] and found that although the proportion of each of the IFE subpopulations is consistent between the two wild type epidermis, the cell composition of the mutant epidermis is clearly different from the wild type (Fig. 1D, 6B). The terminally differentiated IFE population (IFE.TD) is almost absent in the *Grhl3*-/- IFE (Fig. 6A, B), suggesting that not only does GRHL3 regulate the expression of terminal differentiation genes, but that it is also required for the formation of the most differentiated cells of the IFE. Most strikingly, however, the proportion of basal cells increased (IFE.B1, IFE.B2, IFE.B3) whereas the proportion of transitioning (IFE.T) and differentiated cells (IFE.D) decreased (Fig. 6B, Supplemental Fig. 5D) in the *Grhl3*-/- IFE. In addition, we detected a substantial number of cells representing IFE states that are exclusive to the *Grhl3*-/- epidermis. We termed these two populations aberrant IFE 1 and 2 (IFE.A1, IFE.A2). About 100 of IFE.A1 cells are found in each of the

two *Grhl3*-/- biological replicate, whereas 135 IFE.A2 cells are found almost exclusively in *Grhl3*-/-2 (Fig. 6A, B). These findings indicate that in addition to promoting terminal differentiation, a major function of GRHL3 is to suppress the abundance of epidermal stem cells and to suppress the formation of abnormal IFE cell states.

GRHL3 suppresses the formation of aberrantly differentiated IFE progenitors

The IFE.A1 cells resemble the normal transitioning cell state (IFE.T) in that IFE.A1 cells express both basal and suprabasal IFE gene signatures (Supplemental Fig. 6A). Consistently, IFE.A1 cells are found between the basal and differentiated IFE cells on PCA analysis (Supplemental Fig. 6B). RNA-FISH staining of the IFE showed an increase in the number of cells in the basal and spinous layers expressing both basal (*Krt14*) and suprabasal (*Krt10*) markers in the *Grhl3*-/- epidermis compared to WT (Fig. 6D). Immunofluorescence staining corroborates this at the protein level (Supplemental Fig. 6C), suggesting that the IFE.A1 cells are located at the basal-spinous boundary. Yet, the IFE.A1 population also expresses genes distinct from wild type IFE.T cells, including *Sprr2a3*, *Tpm2*, and *Ly6a* (Fig.6C, Supplemental Fig. 6D). Furthermore, cell cycle analysis[47,48] on all IFE subpopulations revealed that IFE.A1 cells are much more proliferative than the IFE.T cells; in fact, a subset of IFE.A1 cells express proliferation genes to a similar level as the dividing basal cluster IFE.B3 (Supplemental Fig. 6E, F). These findings indicate that GRHL3 suppresses the formation of aberrantly differentiated epidermal progenitors and drives late

epidermal developmental maturation. Given the unexpected change in basal cell composition and the emergence of aberrant transition-like cells in the *Grhl3*-/- IFE, we conclude that GRHL3 is required for maintaining the proper cell composition, acting as early as the transition state between basal and spinous cells.

Loss of GRHL3 disrupts the normal IFE differentiation trajectory at the IFE.B-IFE.T transition

To further define the differentiation stage in which GRHL3 functions during IFE differentiation, we constructed separate pseudotime trajectories for WT and *Grhl3*-/- IFE cells. In agreement with our earlier experiment (Fig. 2), the two WT IFE samples formed a linear trajectory going from basal to terminally differentiated cells (Fig. 7A, B, Supplemental Fig. 7A). In the *Grhl3*-/- IFE, the aforementioned expansion of proliferating basal cells and the presence of the IFE.A1 subpopulation both contributed to a branching of the linear differentiation trajectory (Fig. 7A). The branching occurred at the junction between basal cells and IFE.T cells. The tip of the aberrant branch in the *Grhl3*-/- IFE is composed of highly proliferating basal IFE cells (IFE.B3) and the IFE.A1 cells. Consistent with the Seurat clustering (Fig. 6A) and the RNA-FISH experiments (Fig. 6E), this pseudotime analysis also suggests that the IFE.A1 cell state is most similar to IFE.T cells. In contrast, the IFE.A2 cell cluster is more similar to cells at a later differentiated stage, albeit with higher proliferation (Supplemental Fig. 6E, F). Thus, besides being a transcriptional activator of late differentiation genes, GRHL3 performs important functions at early stages of IFE differentiation.

RNA-velocity analysis identifies a GRHL3-regulated differentiation commitment point

To further understand the differentiation abnormalities in the *Grhl3*-/- IFE, we applied RNA-velocity, a method that exploits the relative abundance of nascent (unspliced) and mature (spliced) transcripts to predict the future state of individual cells on a time scale of hours[49,50]. This method can infer directionality and dynamics of small groups of cells with respect to each other in the pseudotime differentiation trajectory. In these data, the direction of the arrows point to the fate the cells are heading toward whereas the length of the arrows reflects how fast the cells are heading in that direction.

We expected that RNA-velocity would reveal arrows with a uniform direction from IFE.B to IFE.TD in the pseudotime trajectory of the WT IFE. Surprisingly, we found that during WT IFE differentiation, the basal and transition cells exist in a dynamic equilibrium state in which no clear directionality is observed among them (Fig. 7C). Once the cells reach the IFE.T/IFE.D boundary, the arrows become longer and aligned in the direction of terminal differentiation in a uniform manner, suggesting that after cells exit the IFE.T phase, they commit to differentiate in one direction until they terminally differentiate (Fig. 7C). These results are not exclusive to this particular dimension reduction method (ie. Monocle pseudotime) as we found similar results with PCA of all IFE cells, which also happens to delineate IFE's linear differentiation trajectory from IFE.B to IFE.T to IFE.D (Supplemental Fig. 7B). In the *Grhl3*-/- IFE, much fewer cells pass the IFE.T/IFE.D commitment point where unidirectional IFE differentiation

commences. Strikingly, early transition cells display a clear propensity back toward the basal cell fate, again suggesting that *Grhl3* plays an important role in promoting the differentiation of IFE.B to IFE.T cells.

These findings suggest: 1) Normal IFE differentiation proceeds in a smooth, continuous manner rather than in punctuated stages. 2) There are crucial changes in the transcriptional regulation for IFE differentiation at the IFE.T/IFE.D boundary where cells align to differentiate. 3) Early progenitor cells transitioning to differentiation may be highly plastic in their progenitor-differentiation fate decisions and loss of GRHL3 disrupts this balance. 4) There may be aberrant signaling in the *Grhl3*-/- epidermis that actively drives the cells toward the basal fate.

GRHL3 is required to repress aberrant Wnt signaling in the IFE

In the P0 *Grhl3-/-* IFE the basal compartment is expanded and IFE.T cells exhibit a propensity to go backward toward the basal fate. These findings suggest that GRHL3 is required to suppress underlying signaling pathways that can actively drive IFE cells toward the basal fate. Previous mouse *Axin2* lineage-tracing experiments showed that *Wnt*-secreting keratinocyte stem cells are distributed in the basal IFE and that autocrine *Wnt* signaling is required for self-renewal in mouse adult epidermal homeostasis[11,46]. Hence, we hypothesized that aberrant *Wnt* signaling underpins the change in cell composition and disruption of differentiation in the *Grhl3-/-* IFE.

To determine if *Wnt* signaling is altered in the *Grhl3-/-* IFE, we compared the expression level of *Wnt* signaling pathway components between P0 WT and *Grhl3-/-* IFE populations. We observed increased expression of genes such as *Wnt4*, *Lef1*, *Ctnnb1*, *Gpc3,* and *Gsk3b* in the *Grhl3-/-* basal IFE populations and IFE.A (Fig.8A, Supplemental Fig. 8A). To validate some of this finding, we performed RNA-FISH of *Wnt4* in WT and *Grhl3-/-* IFE and observed that overall *Wnt4* is expressed at a higher level in *Grhl3-/-* IFE. Furthermore, whereas *Wnt4* expression is restricted to only a subset of basal WT IFE, it is expressed in more basal cells and in cells in the spinous layer in *Grhl3-/-* IFE (Figure 8B). This result is also consistent with the scRNA-seq results where an increase in the number of keratinocyte stem cells (IFE.B) is observed in *Grhl3-/-* IFE; and the fact that the IFE.A population exclusive to the *Grhl3-/-* IFE is a transition population expressing high level of *Wnt4*. Therefore, Grhl3 is required to ensure proper suppression of Wnt signaling upon differentiation.

Figure 1. GRHL3 is required for normal IFE cell states in late epidermal development.

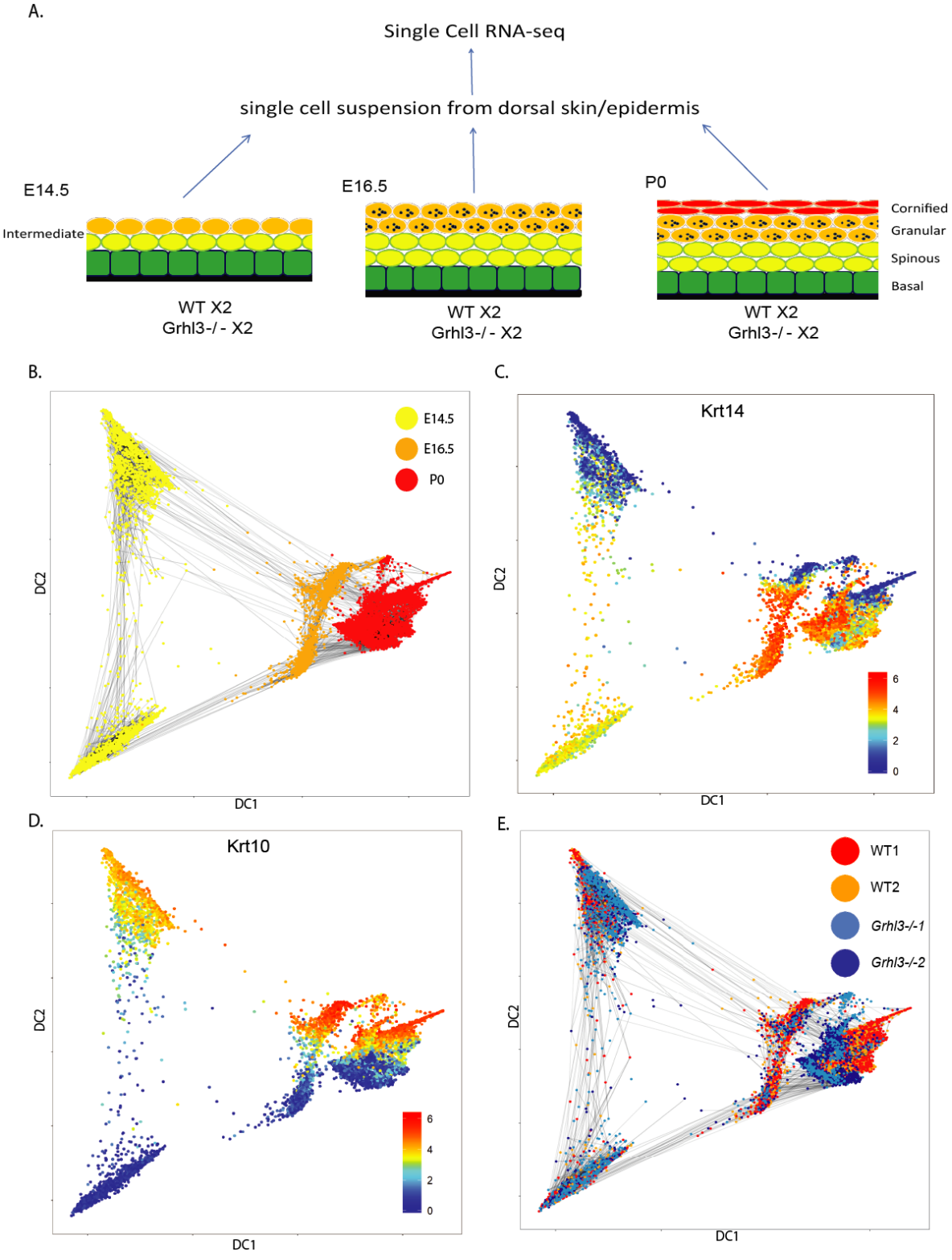


Figure 1. GRHL3 is required for normal IFE cell states in late epidermal development. A) Experimental scheme for single cell RNA-seq experiments. E14.5, E16.5 skin and P0 epidermis cells in two WT and two *Grhl3-/-* littermates at each time point were collected to generate single cell suspension for scRNA-seq using 10X Genomics Chromium. An additional WT P0 mouse epidermis was also analyzed (Figs. 2-3). B) Diffusion map of IFE cells combined from E14.5, E16.5 and P0 mouse epidermis. DC1 separates the cells by developmental age whereas DC2 separates the cells by progress in IFE differentiation. Lines connect potential similar cell states across differentiation/developmental time. C) Expression heatmap of *Krt14* and D) *Krt10* projected onto the diffusion map. E) Genotype of each cell is projected onto the diffusion map. A shift toward earlier developmental time point is observed for the *Grhl3**-/-*** IFE at P0.

Figure 2. IFE differentiation is gradualistic rather than step-wise and features a large number of transition cells.

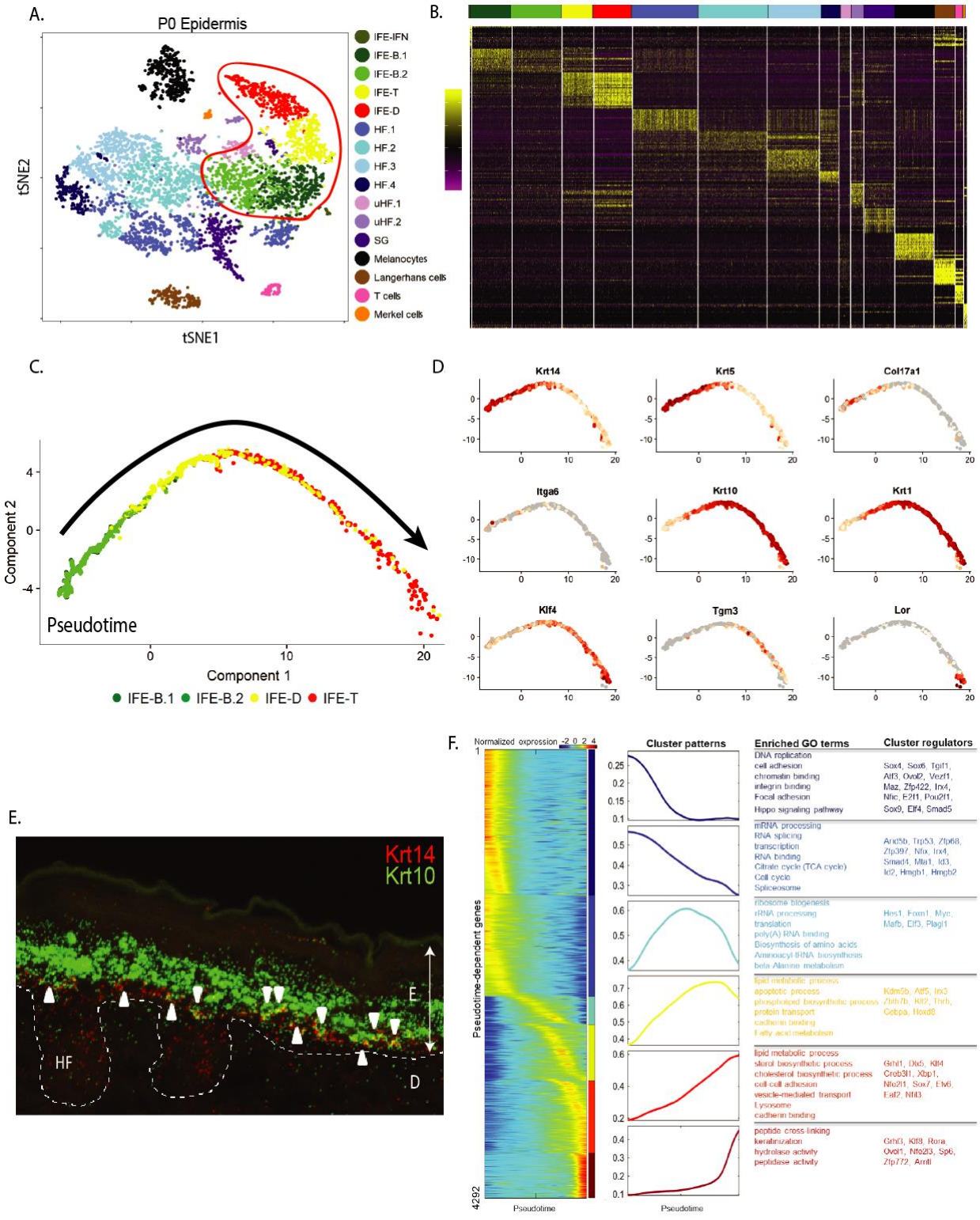


Figure 2. IFE differentiation is gradualistic rather than step-wise and features a large number of transition cells. A) tSNE plot showing all 16 epidermal subpopulations of the WT P0 mouse back epidermis; IFE cells are in the red-outlined box. B) Heatmap showing the expression of top 20 marker genes for all 16 epidermal subpopulations. C) Pseudotime analysis of the IFE with cluster identity from panel A projected on the trajectory. The trajectory goes from basal to transition to differentiated cells without major branches (arrow), consistent with unidirectional IFE differentiation. D) Expression of canonical markers for distinct stages of epidermal differentiation. E) RNA-FISH for *Krt10* and *Krt14* in WT P0 mouse epidermis; white arrowheads points to yellow *Krt10*/*Krt14* double-positive cells; E, epidermis; D, dermis; HF, hair follicle. F) Expression heatmap and expression pattern of 4,292 pseudotime-dependent genes that cluster into six gene modules. Also shown are the gene ontology and example transcription factors for each gene module.

Figure 3. The P0 IFE basal cell layer contains two distinct cell populations.

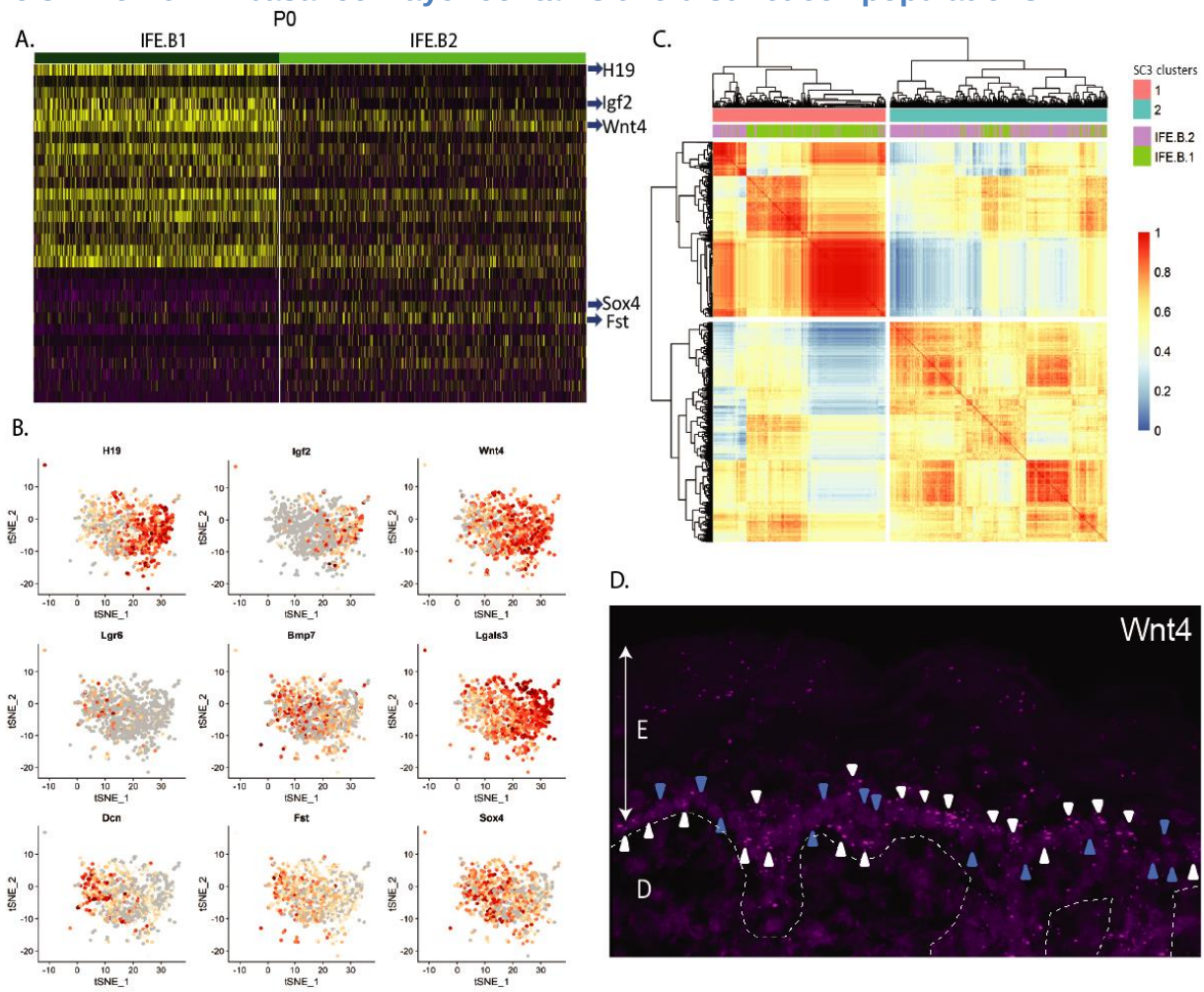


Figure 3. The P0 IFE basal cell layer contains two distinct cell populations.

A) Expression heatmap of unique gene signatures for each of the two basal IFE clusters. Blue arrows highlight select differentially expressed genes between the two basal clusters. B) Expression level of select unique gene signatures projected onto tSNE for the two basal clusters. C) Heatmap showing the correlation between cells for SC3 clustering results: 1, 2 (red, blue) compared to Seurat clustering results: IFE.B1, IFE.B2 (purple, green). D) RNA-FISH for *Wnt4* in the P0 IFE. White arrowheads point to *Wnt4*-high cells; blue arrowheads point to *Wnt4*-low cells.

Figure 4. scRNA-seq reveals a proliferating transition population in the E14.5 IFE.

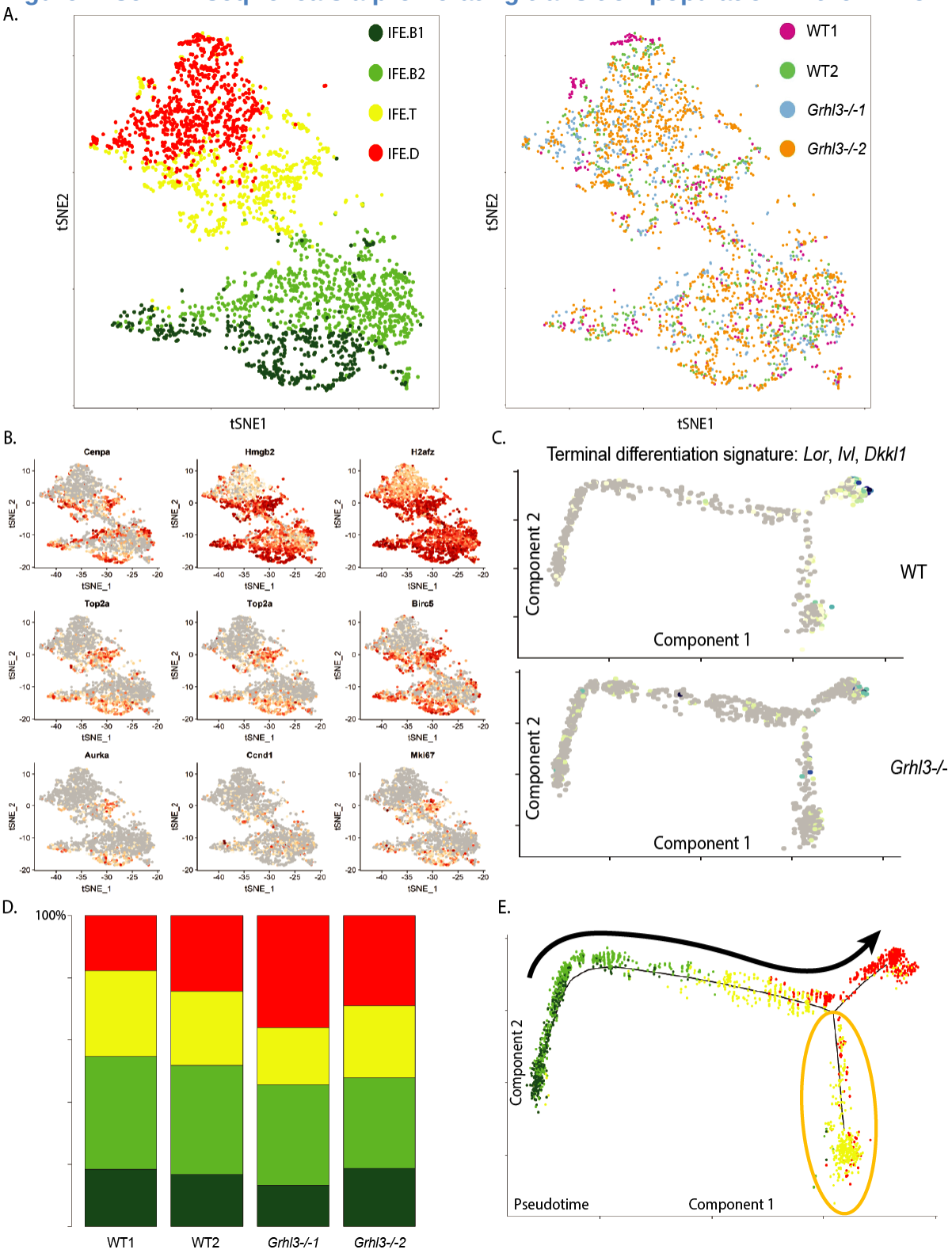


Figure 4. scRNA-seq reveals a proliferating transition population in the E14.5 IFE. A) tSNE plot showing basal (IFE.B1, IFE.B2), transition (IFE.T) and differentiated (IFE.D) cells for the WT and *Grhl3-/-* E14.5 IFE (Left); same cells colored according to genotypes (right). B) Expression level of cell proliferation genes. There is high expression of proliferation genes in one of the basal populations (IFE.B1) and in the transition cells (IFE.T). C) Terminal differentiation gene signature score based on average expression levels of *Lor*, *Ivl* and *Dkkl1* projected onto pseudotime differentiation trajectory in WT and *Grhl3-/-* E14.5 IFE. D) Bar chart showing the percentage of each subpopulation in each biological replicate. The cellular composition is similar in WT and *Grhl3*-/- E14.5 IFE. E) Pseudotime differentiation trajectory of E14.5 IFE cells; cells are colored according to cluster identity in panel A. The trajectory is consistent with IFE differentiation, going from basal to transition to differentiated cells (arrow); also, the transition population (presumed intermediate layer keratinocytes) forms a branch, apparently due to high cell proliferation. Loss of *Grhl3* does not affect the differentiation trajectory. Orange circle outlines the proliferative transition cell state.

Figure 5. scRNA-seq reveals decreased size of proliferating transition cells and progress to terminal differentiation in the E16.5 IFE.

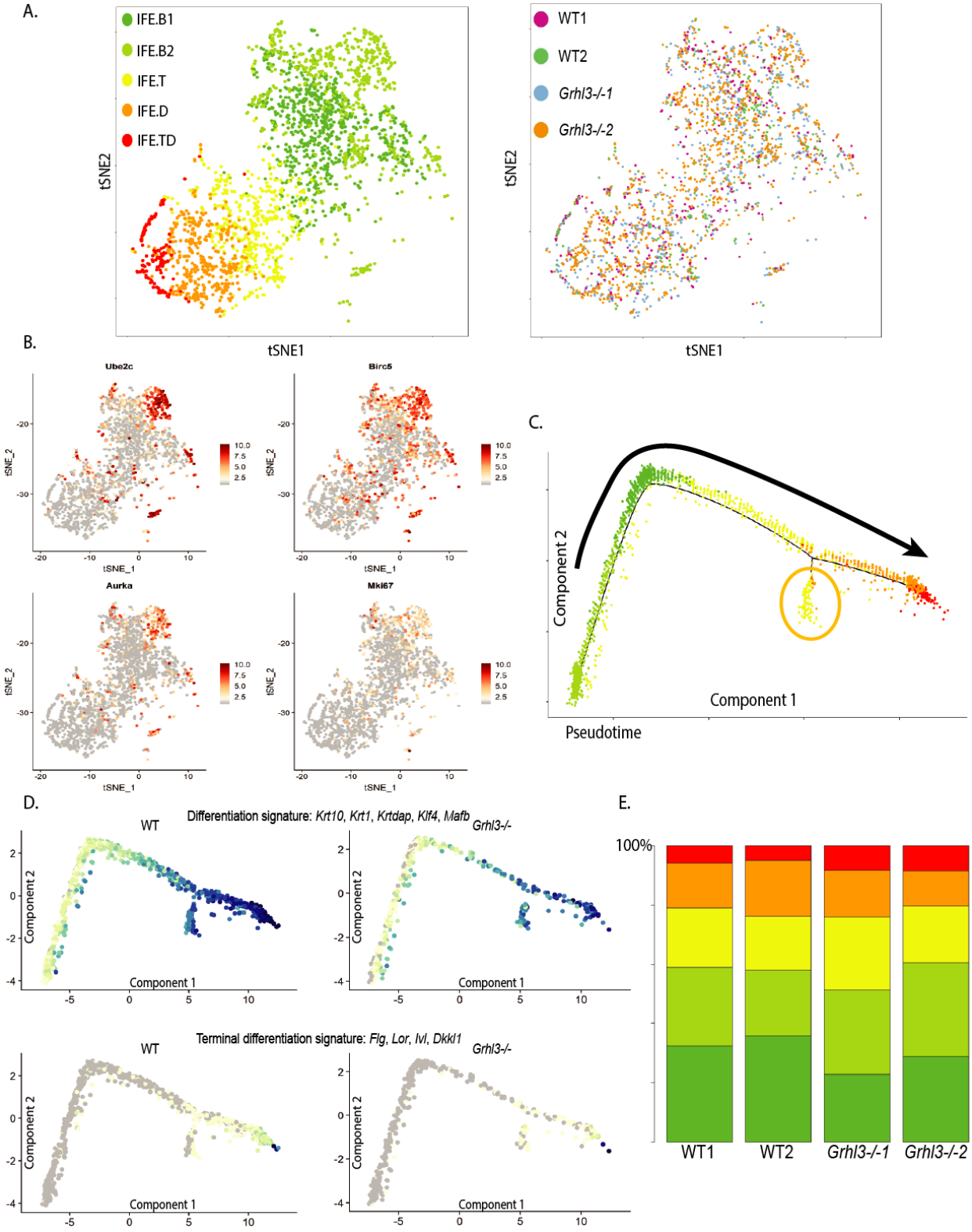


Figure 5. scRNA-seq reveals decreased size of proliferating transition cells and progress to terminal differentiation in the E16.5 IFE. A) tSNE plot showing basal (IFE.B1, IFE.B2), transition (IFE.T), differentiated (IFE.D), and terminally differentiated (IFE.TD) cells for the E16.5 WT and *Grhl3-/-* IFE (Left); same cells colored by their genotype (right). B) Expression level of proliferation genes projected onto tSNE in panel A. C) Pseudotime differentiation trajectory of E16.5 IFE cells colored according to cluster identity in panel A; IFE differentiation proceeds from basal (IFE.B1, IFE.B2) to transition (IFE.T) to differentiated populations (IFE.D, IFE.TD; arrow). Orange box highlights the proliferative transition cell state. The population of proliferating transition cells is smaller than at E14.5, creating a minor branching at the late transition population. Loss of *Grhl3* does not affect the differentiation trajectory. D) Differentiation gene signature score based on average expression levels of *Krt10*, *Krt1*, *Krtdap*, *Klf4*, and *Mafb* projected onto pseudotime in WT and *Grhl3-/-* E16.5 IFE, respectively (top); terminal differentiation score based on average expression levels of *Flg*, *Lor*, *Ivl*, *Dkkl1* (bottom). E) Bar chart showing the percentage of each IFE subpopulation in WT1, WT2, *Grhl3-/-*1, *Grhl3-/-*2; the cellular composition is similar in WT and *Grhl3-/-* E16.5 IFE.

Figure 6. Grhl3 is required for the formation of terminally differentiated cells and for suppressing aberrant expansion of basal IFE stem cells.

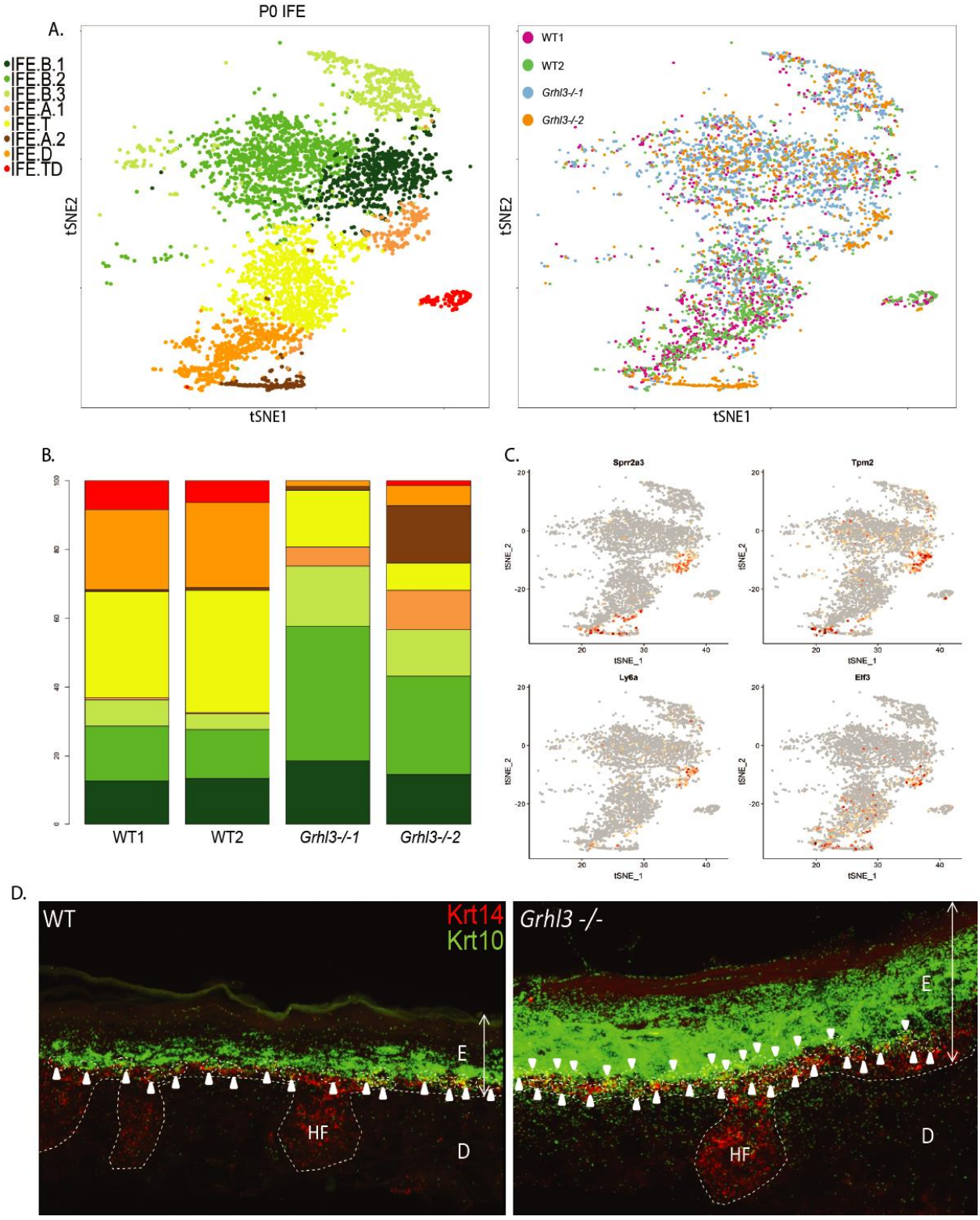


Figure 6. Grhl3 is required for the formation of terminally differentiated cells and for suppressing aberrant expansion of basal IFE stem cells. A) tSNE plot showing WT and *Grhl3-/-* P0 IFE cells colored by clusters (left) and genotype (right). B) Bar plot showing the fraction of each IFE cell type in each sample. There is expansion of basal cells and reduction of differentiated/terminally differentiated cells in *Grhl3*-/- IFE. C) Gene expression levels of markers for the IFE.A population. *Sprr2a3, Tpm2, Ly6a and Eif3* are shown on a tSNE plot of the P0 IFE cells. D) RNA-FISH of *Krt14* and *Krt10* transcripts in P0 WT and *Grhl3*-/- mouse epidermis; white arrows highlight *Krt14*/*Krt10* double-positive transition cells. Note the higher number of double-positive cells in the *Grhl3*-/- epidermis; also some transition cells are found in the spinous layer.

Figure 7. RNA velocity reveals a differentiation commitment point at the IFE.T/IFE.D transition.

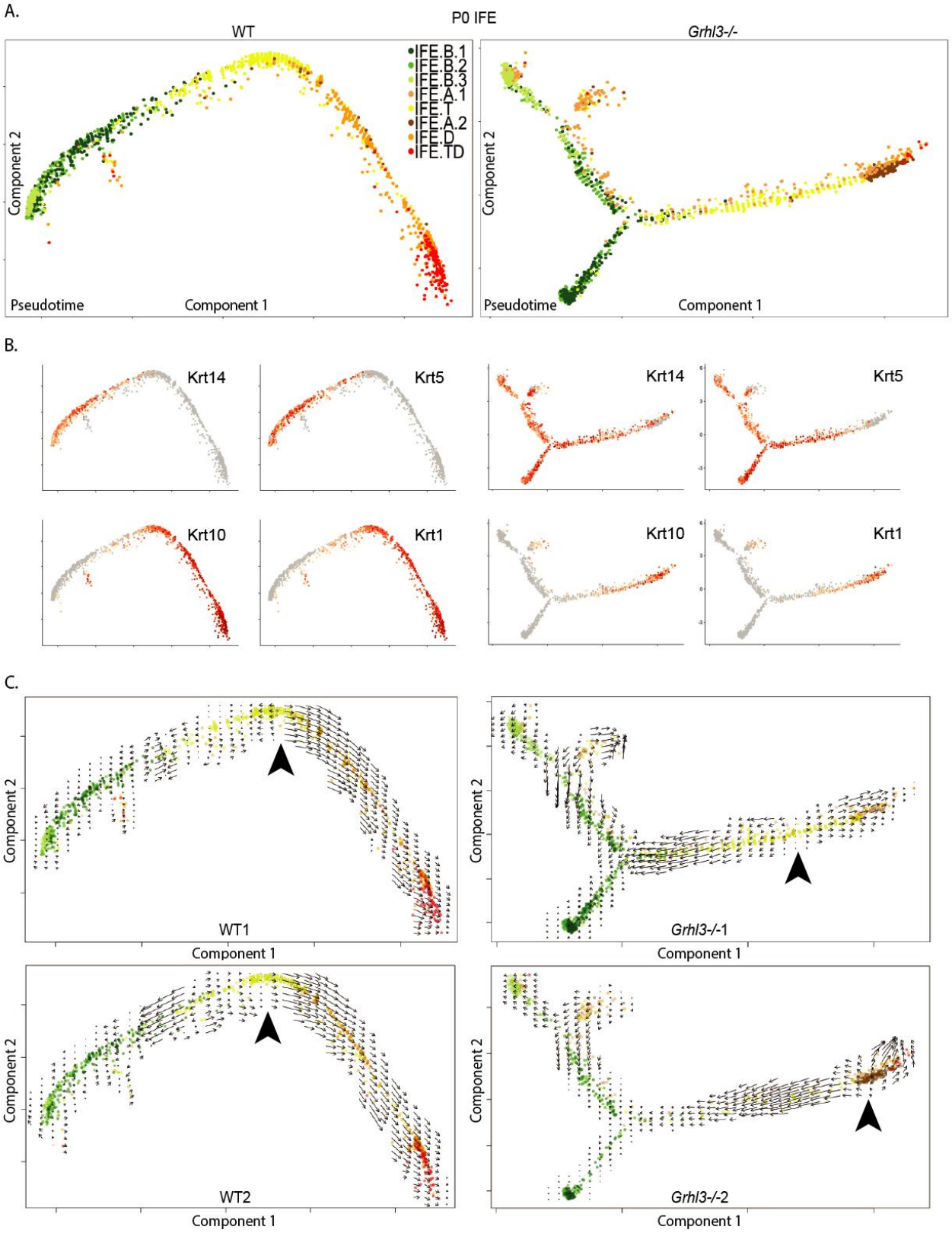


Figure 7. RNA velocity reveals a differentiation commitment point at the IFE.T/IFE.D transition. A) Pseudotime trajectories showing IFE differentiation going from IFE-B to IFE-T to IFE-D to IFE-TD for the two WT (left) and two *Grhl3*-/- (right) P0 IFEs. The *Grhl3*-/- differentiation trajectory has a branch which is attributed to the IFE-A1 population and expansion of basal, proliferating cells. B) Expression of *Krt14*, *Krt5*, *Krt10*, and *Krt1* projected on the pseudotime trajectory for the WT (left) and the *Grhl3*-/- (right) IFE. C) RNA-Velocity analysis of the pseudotime trajectory of IFE cells for WT1, WT2, *Grhl3*-/-1, *Grhl3-/-*2. Direction of the arrows points to the fate the cells are heading toward; length of the arrows reflects how fast the cells are heading toward a particular fate. Note IFE differentiation is marked by a clear commitment point (arrowhead) and is a one-step, continuous process with no intermediary stages in the WT. In contrast, most cells fail to proceed through the commitment point (arrowhead) and transition cells have a propensity to go back to basal fate in the *Grhl3-/-* IFE.

Figure 8. GRHL3 represses Wnt pathway gene expression in the basal IFE cells.

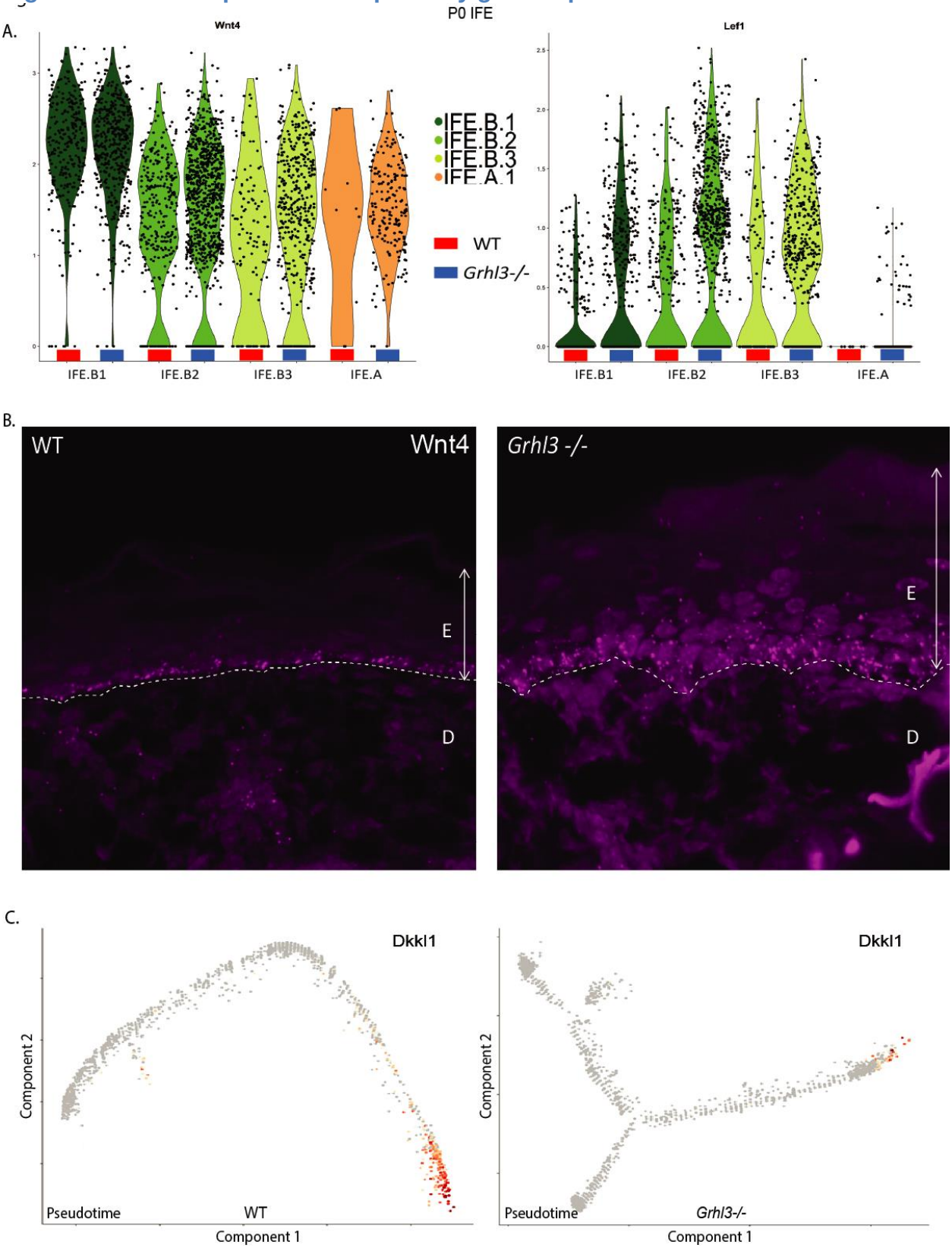
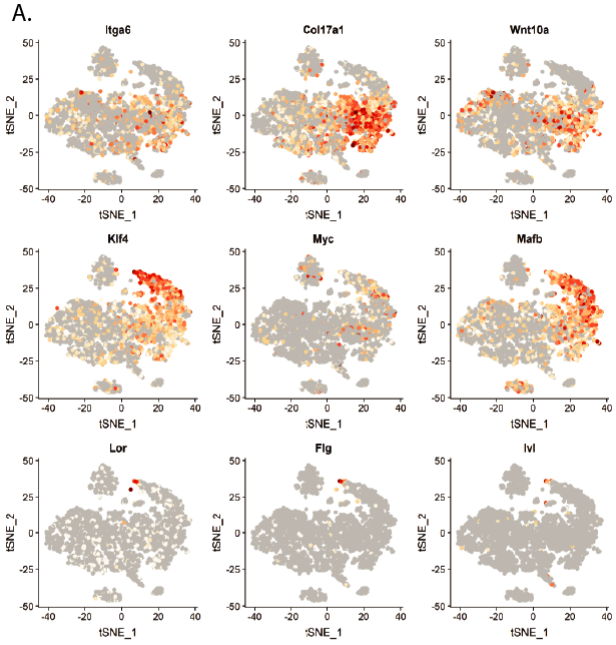


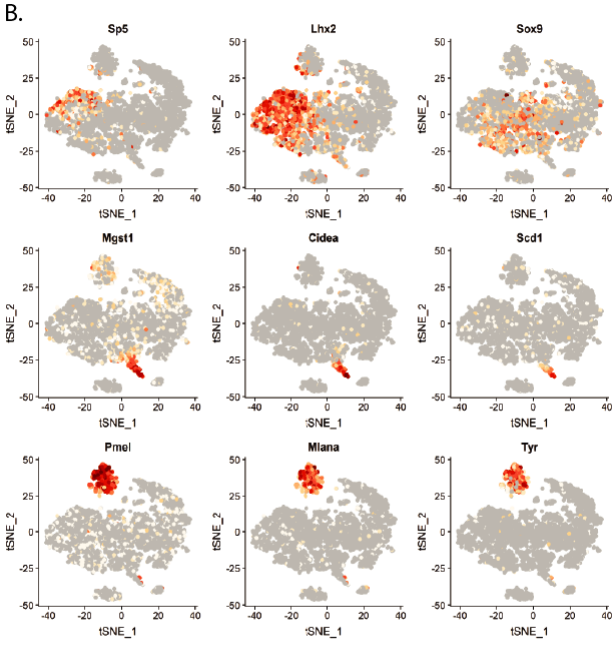
Figure 8. GRHL3 represses Wnt pathway gene expression in the basal IFE cells. A) Violin plot showing the *Wnt4* (left) and *Lef1* (right) expression for each cell among the basal cell types in the WT (both biological replicates) and the *Grhl3*-/- (both biological replicates) IFE. There is higher number of *Wnt4*-high and *Lef1*-high cells in all basal and aberrant populations in the *Grhl3*-/- than in the WT. B) Representative RNA-FISH image of *Wnt4* in WT and *Grhl3*-/- P0 epidermis. There is increased *Wnt4* expression and aberrant expression of *Wnt4* in cells above the basal layer in *Grhl3*-/- epidermis. n = 3 WT and *Grhl3-/-* littermates. C) Expression of a secreted antagonist of the Wnt signaling pathway, *Dkkl1*, projected onto the pseudotime trajectory of WT (left) and *Grhl3*-/- (right) IFE cells. Note the absence of terminally differentiated cells in *Grhl3*-/- epidermis with lower *Dkkl1* expression.

Supplemental Figure 1

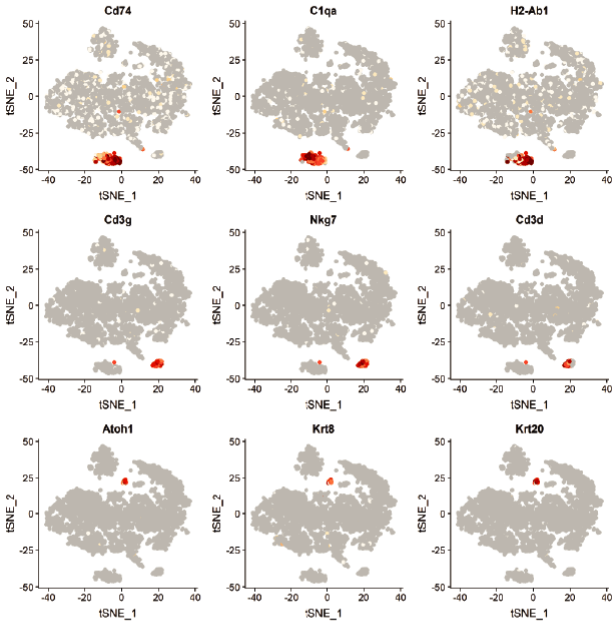
A.



B.



C.



D.

GO biological process complete	fold Enrichment	p-value
keratinocyte proliferation (GO:0043616)	48.03	1.52E-02
labyrinthine layer morphogenesis (GO:0060713)	28.91	8.81E-03
skin development (GO:0043588)	7.68	6.20E-04
cell-cell signaling by wnt (GO:0198738)	7.21	4.05E-03
Wnt signaling pathway (GO:0016055)	7.21	4.05E-03
epidermis development (GO:0008544)	6.53	1.06E-02
cell-cell adhesion (GO:0098609)	5.88	3.57E-03
regulation of epithelial cell proliferation (GO:0050678)	5.85	1.06E-02
morphogenesis of an epithelium (GO:0002009)	5.53	1.32E-04

E.

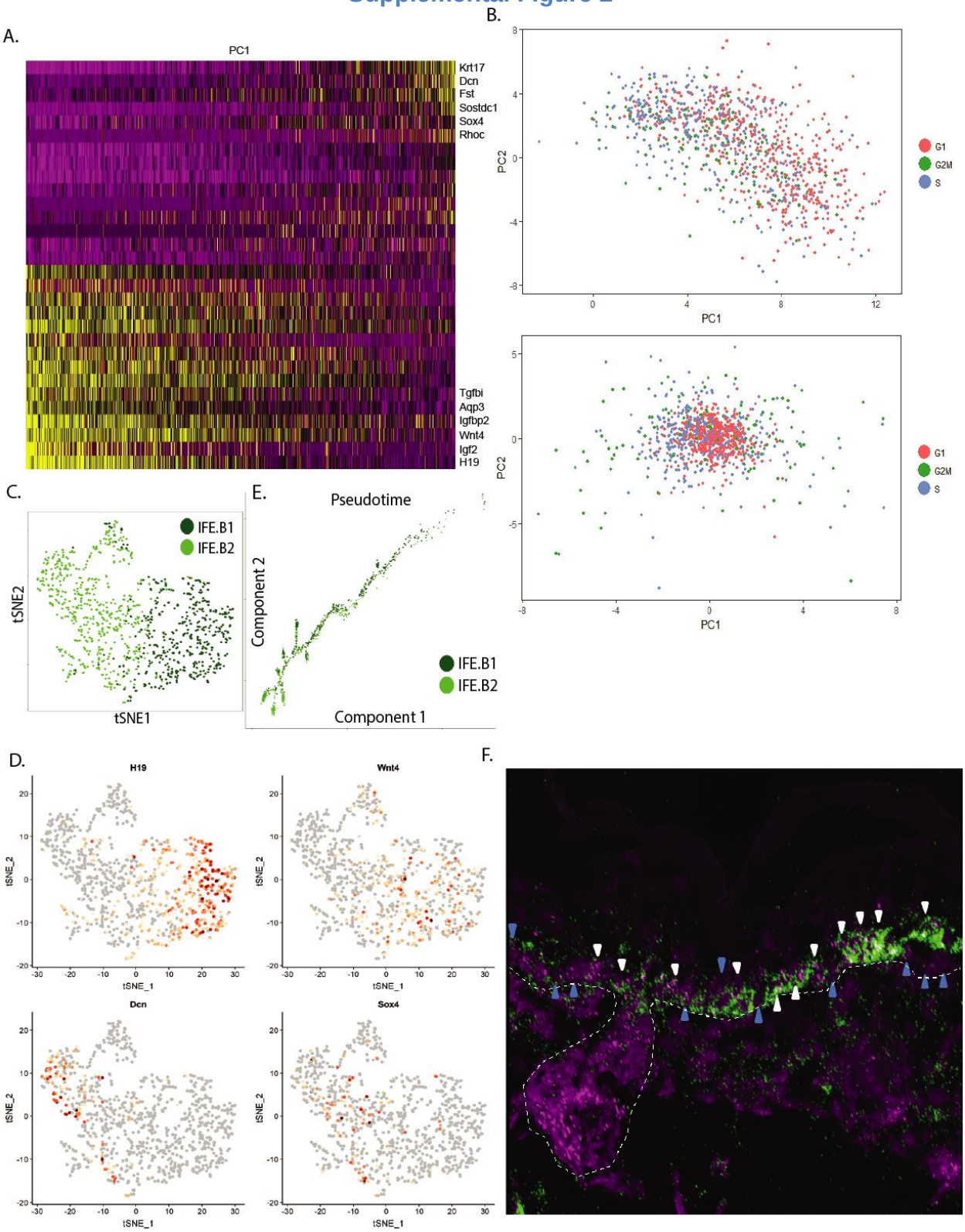
GO biological process complete	Fold Enrichment	p-value
keratinocyte differentiation (GO:0030216)	8.91	3.53E-02
epidermal cell differentiation (GO:0009913)	7.38	1.18E-02
skin development (GO:0043588)	6.78	2.55E-04
epidermis development (GO:0008544)	6.74	8.72E-05
epithelium development (GO:0060429)	2.88	3.15E-02
regulation of cell proliferation (GO:0042127)	2.45	2.02E-02
negative regulation of cellular process (GO:0048523)	1.76	2.00E-02
negative regulation of biological process (GO:0048519)	1.7	2.74E-02

F.

GO biological process complete	Fold Enrichment	p-value
desmosome organization (GO:0002934)	27.9	1.09E-02
establishment of skin barrier (GO:0061436)	18.6	1.23E-03
regulation of water loss via skin (GO:0033561)	15.63	3.91E-03
keratinization (GO:0031424)	14.77	1.44E-04
icosanoid biosynthetic process (GO:0046456)	12.6	1.61E-02
water homeostasis (GO:0030104)	10.92	1.81E-03
multicellular organismal water homeostasis (GO:0050891)	10.89	8.73E-03
sphingolipid biosynthetic process (GO:0030148)	8.37	1.59E-02
keratinocyte differentiation (GO:0030216)	6.84	8.48E-04
fatty acid biosynthetic process (GO:0006633)	6.75	9.35E-03

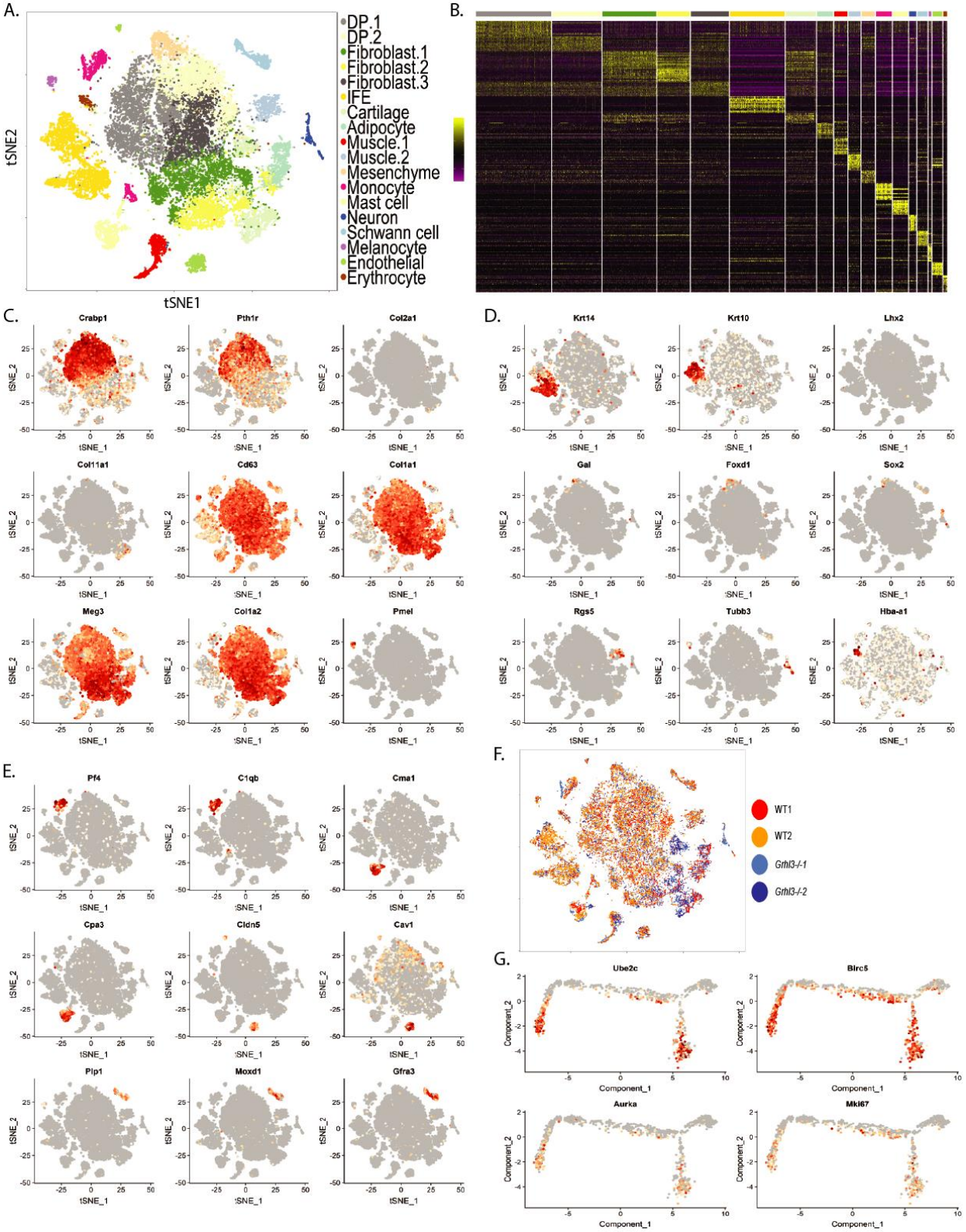
Supplemental Fig. 1. Related to Figure 2. A) Expression of canonical basal (*Itga6*, *Col17a1*, *Wnt10a*), differentiated (*Klf4*, *Myc*,*Mafb*), and terminally differentiated (*Lor*, *Flg*, *Ivl*) IFE markers. B) Expression of hair follicle (*Sp5*, *Lhx2*, *Sox9*), sebaceous gland (*Mgst1*, *Cidea*, *Scd1*), and melanocyte (*Pmel*, *Mlana*, *Tyr*) markers. C) Expression of Langerhans cell (*Cd74*, *C1qa*, *H2-Ab1*), T-cell (*Cd3g*, *Nkg7*, *Cd3d*), and Merkel cell (*Atoh1*, *Krt8*, *Krt20*) markers. D-F) GO enrichment categories with fold enrichment and p-value for the marker genes of the basal (D), transition (E), and differentiated (F) IFE populations.

Supplemental Figure 2



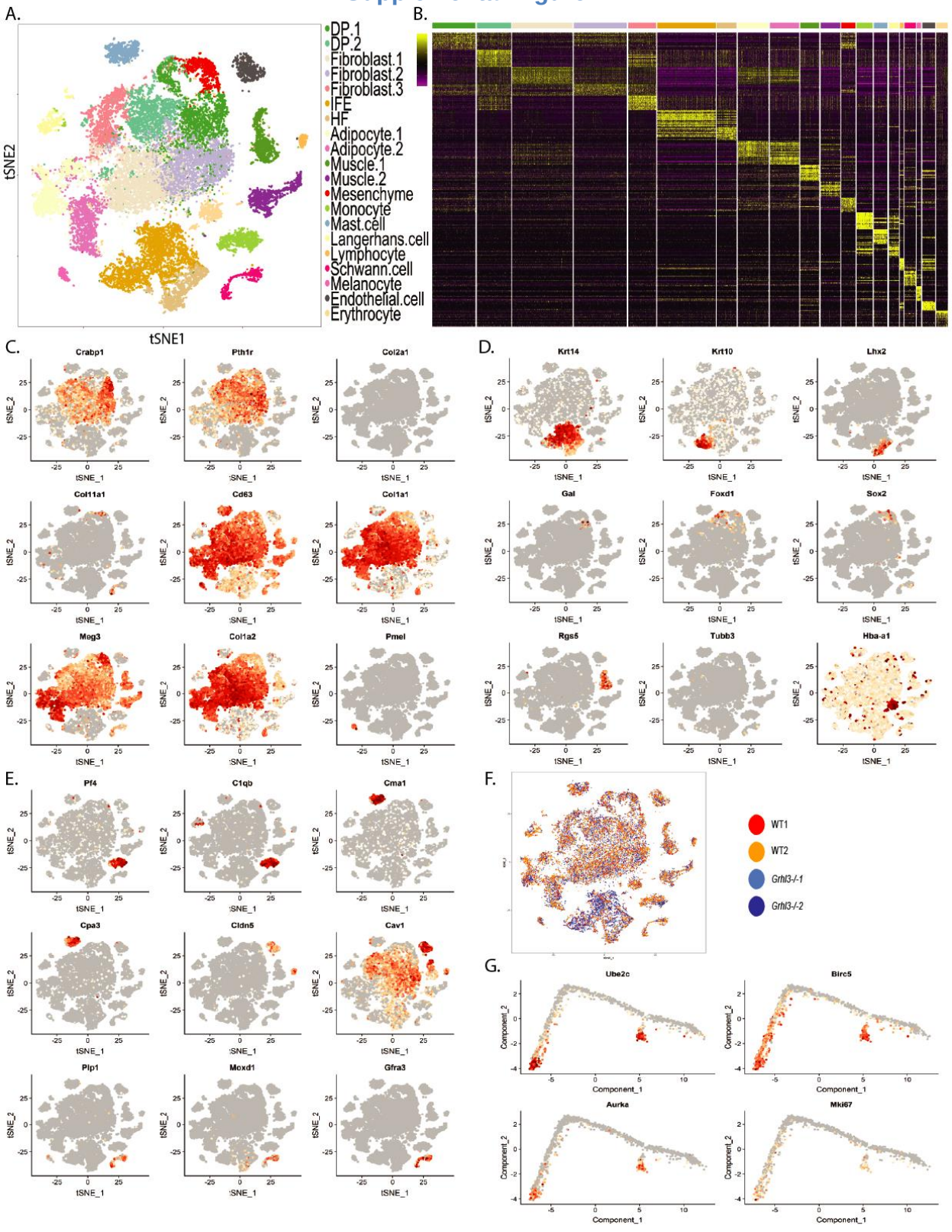
Supplemental Fig. 2. Related to Figure 3. A) Heatmap showing expression of top loading genes in PC1 for PCA analysis of all basal IFE cells. B) PCA analysis was done on all basal IFE cells using cell cycle genes. Inferred cell cycle phases are projected onto the PCA plot prior (top) and after (bottom) the effect of cell cycle genes were regressed out. C) tSNE result of all basal IFE cells after cell cycle genes were regressed out. The two basal IFE clusters identified in Fig. 2A are projected onto the plot. D) Expression level of the two basal subcluster markers *H19*, *Wnt4*, *Dcn*, and *Sox4* projected onto the tSNE plot in panel C. E) Pseudotime trajectory of all basal IFE cells with the identity of the two basal subclusters projected onto the trajectory. F) RNA-FISH of *H19* (green) and *Wnt4* (purple). White arrowheads point to *Wnt4*-high cells and blue arrows point to *Wnt4*-low cells; high *Wnt4* expression coincides with high *H19* expression.

Supplemental Figure 3



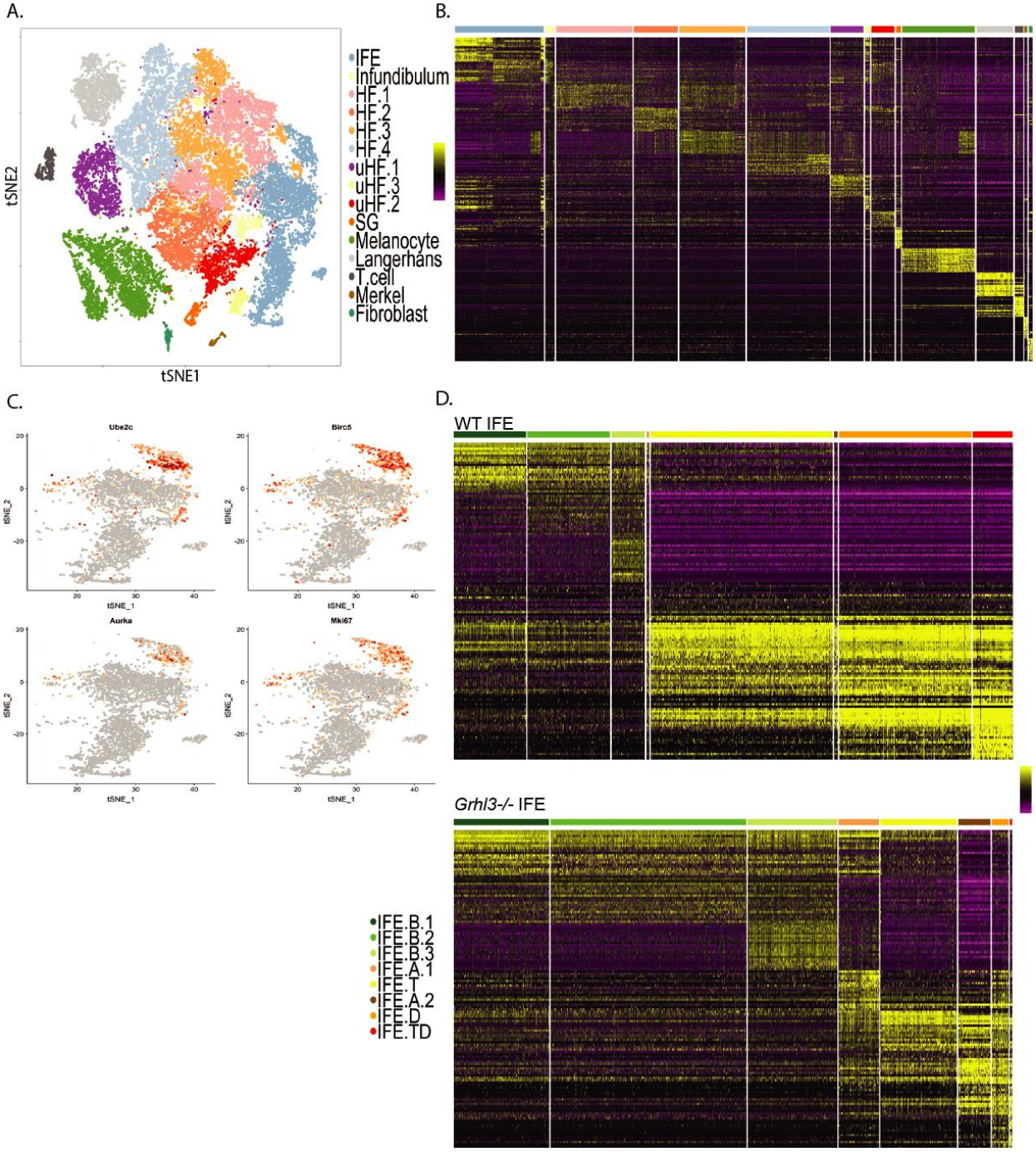
Supplemental Fig. 3. Related to Figure 4. A) tSNE plot of E14.5 (WT1, WT2, *Grhl3-/-1, Grhl3-/-2*) skin cells. B) Heatmap showing expression of marker genes for each population in panel A. C) Expression of level of dermal papillae (*Crabp1*, *Pth1r*), cartilage (*Col2a1*, *Col11a1*), dermal fibroblast (*Cd63*, *Col1a1*, *Meg3*, *Col1a2*) and melanocytes (*Pmel*) markers projected onto tSNE. D) Expression of basal IFE (*Krt14*), suprabasal IFE (*Krt10*), hair follicle (*Lhx2*), mesenchyme (*Gal*, *Foxd1*), astrocytes (*Sox2*), muscle (*Rgs5*) neuron (*Tubb3*) and erythrocyte (*Hba-a1*) markers projected onto tSNE. E) Expression level of monocytes (*Pf4*, *C1qb*), mast cells (*Cma1*, *Cpa3*), endothelial cells (*Cldn5*, *Cav1*), and Schwann cell (*Plp1*, *Moxd1*, *Gfra3*) markers projected onto tSNE. F) Genotype of each cell projected onto the tSNE plot shown in panel A. G) Expression levels of proliferation markers *Ube2c*, *Birc5*, *Aurka*, and *Mki67* projected onto pseudotime trajectory in Fig. 4B.

Supplemental Figure 4



Supplemental Fig. 4. Related to Figure 5. A) tSNE plot of E16.5 (WT1, WT2, *Grhl3-/-1, Grhl3-/-2*) skin cell populations. B) Heatmap showing expression of marker genes for each population in panel A. C) Expression level of dermal papillae (*Crabp1*, *Pth1r*), cartilage (*Col2a1*, *Col11a1*), dermal fibroblast (*Cd63*, *Col1a1*, *Meg3*, *Col1a2*) and melanocytes (*Pmel*) markers projected onto tSNE. D) Expression of basal IFE (*Krt14*), suprabasal IFE (*Krt10*), hair follicle (*Lhx2*), mesenchyme (*Gal*, *Foxd1*), astrocytes (*Sox2*), muscle (*Rgs5*) neuron (*Tubb3*) and erythrocyte (*Hba-a1*) markers projected onto tSNE. E) Expression level of monocytes (*Pf4*, *C1qb*), mast cells (*Cma1*, *Cpa3*), endothelial cells (*Cldn5*, *Cav1*), and Schwann cell (*Plp1*, *Moxd1*, *Gfra3*) markers projected onto tSNE. F) Genotype of each cell projected onto the tSNE plot shown in panel A. G) Expression level of proliferation markers *Ube2c*, *Birc5*, *Aurka*, and *Mki67* projected onto pseudotime trajectory in Fig. 5B.

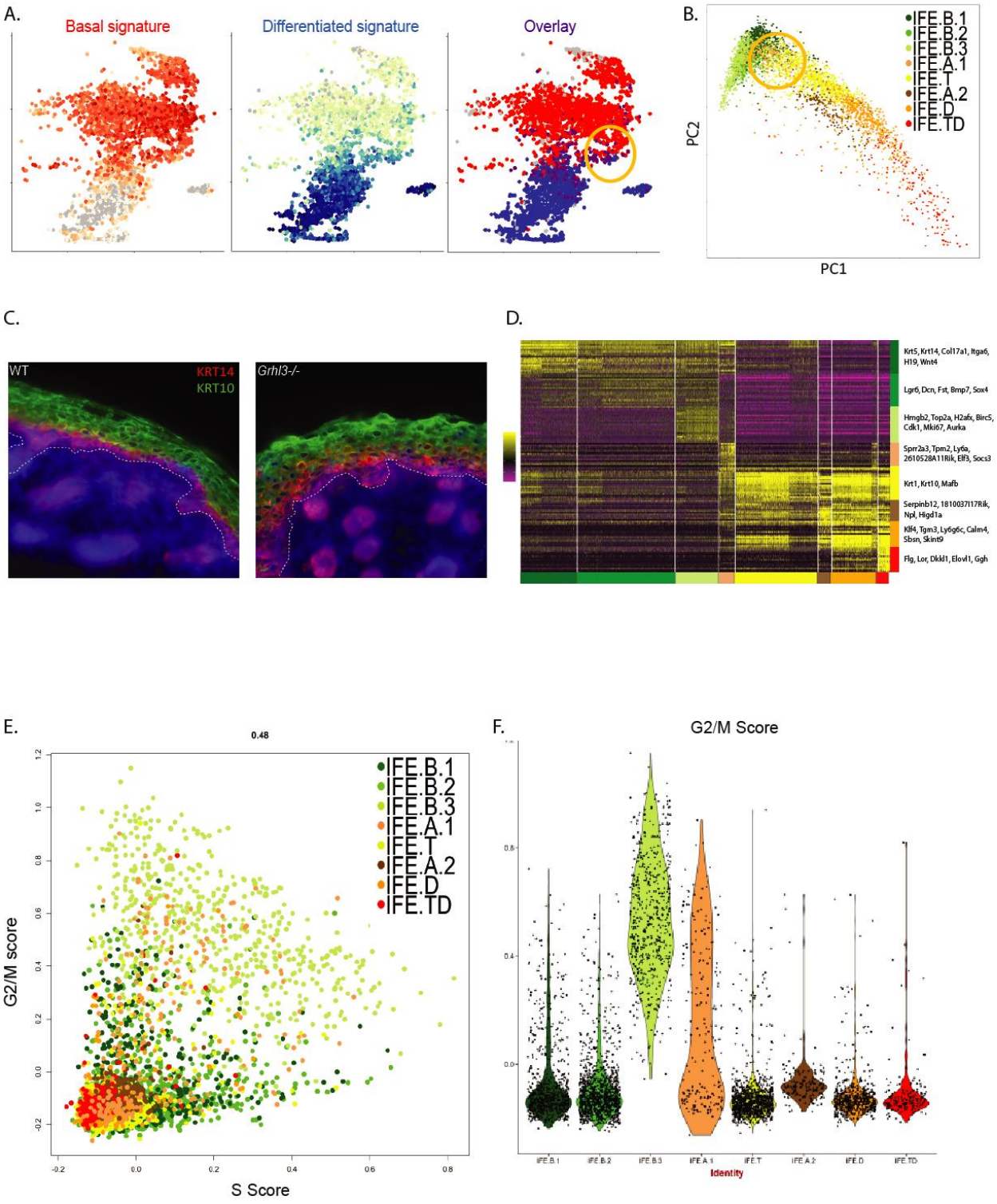
Supplemental Figure 5



Supplemental Fig. 5. Related to Figure 6. A) tSNE plot of P0 (WT1, WT2, *Grhl3-/-1, Grhl3-/-2*) epidermis cell populations. B) Heatmap showing expression of

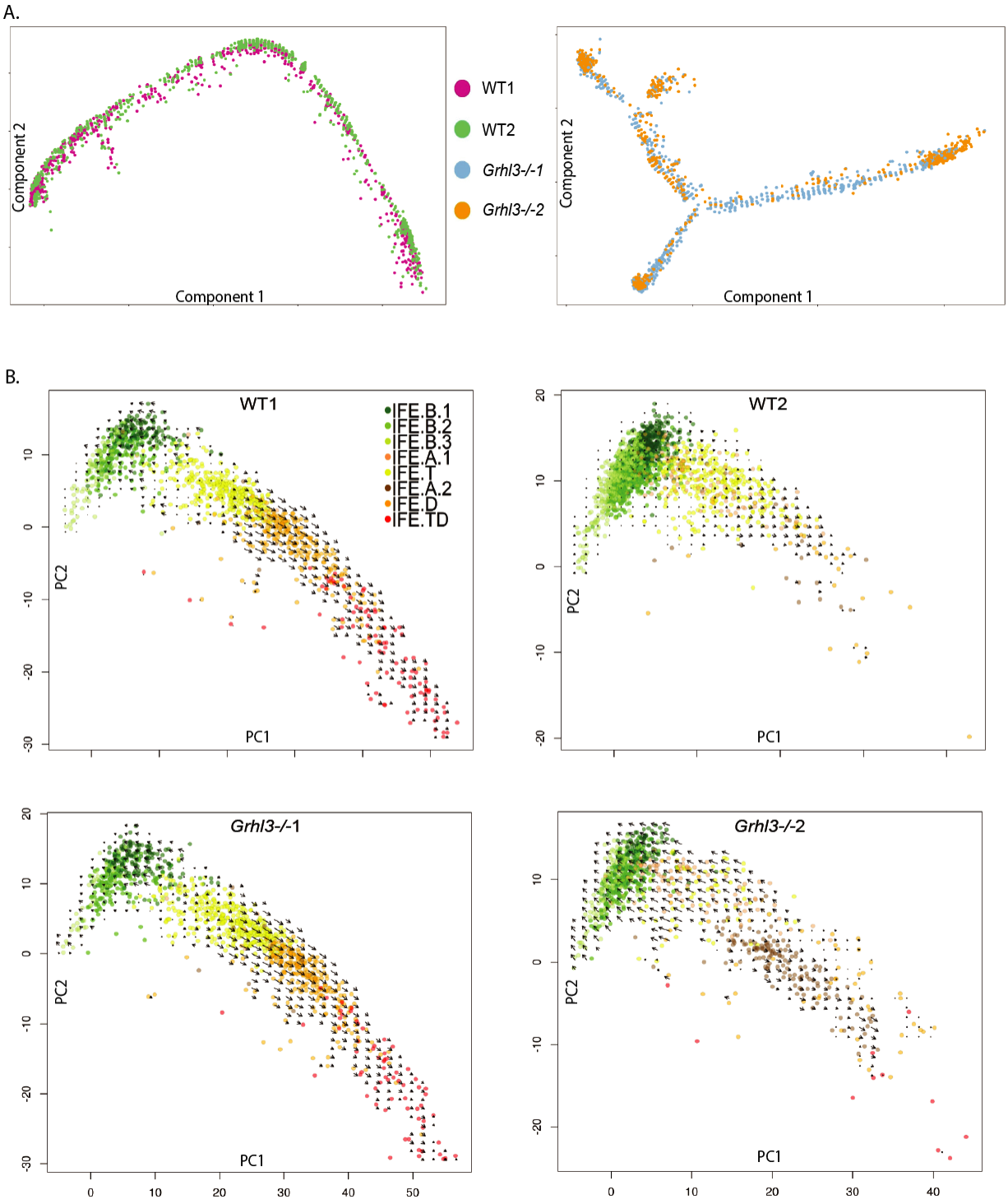
marker genes for each population in panel A. C) Expression level of cell proliferation genes *Ube2c, Birc5, Aurka,* and Mki67 projected onto IFE's tSNE plot in Fig. 6A. D) Heatmap showing the expression of IFE subpopulation markers in WT (top) and *Grhl3-/-* IFE cells.

Supplemental Figure 6



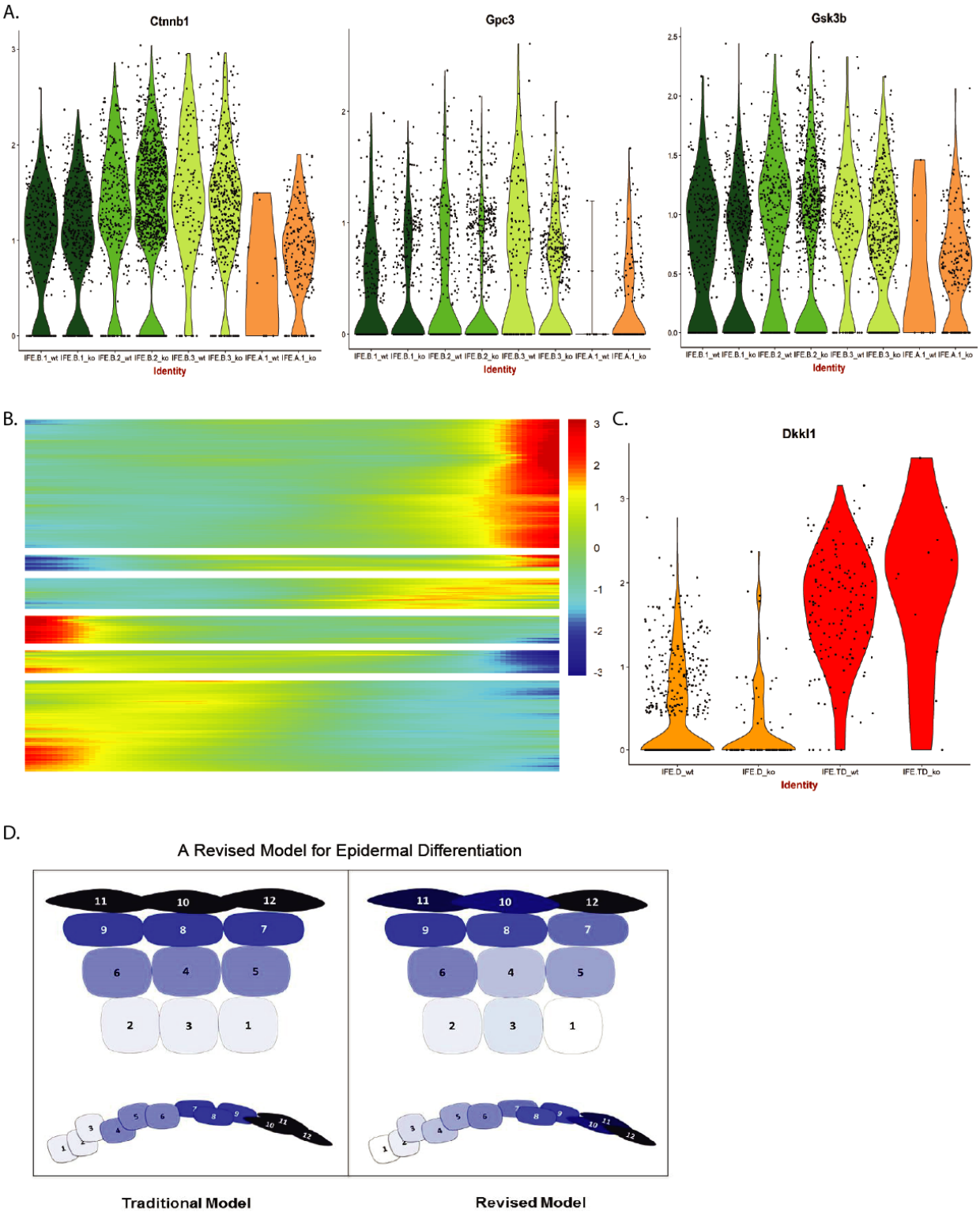
Supplemental Fig. 6. Related to Figure 6. A) A basal signature score as reflected by expression levels of *Krt14, Krt5, Col17a1, Itga6,* and *Wnt10a*; and (left), suprabasal signature score as reflected by the expression levels of *Krt10, Krt1, Krtdap, Klf4,* and *Mafb* (middle) and overlay of basal and suprabasal signature score (right) projected onto tSNE plot of P0 IFE. Orange circle highlights the IFE.A1 population where both basal and suprabasal signature scores are high as indicated in purple. B) PCA plot of P0 IFE populations. Orange circle highlight the position of IFE.A1 population, which is in between the basal and the transition population. C) Immunofluorescence of WT and *Grhl3-/-* P0 mouse epidermis against KRT14 and KRT10. A thicker epidermis and the presence of more orange KRT14/KRT10 double positive cells is observed. D) Heatmap showing the expression of markers for each IFE cluster in WT and *Grhl3-/-* IFE cells. Sprr2a3, Tpm2, Ly6a etc. are specifically expressed in IFE.A population. E) P0 IFE cells are scored based on the expression of distinct cell cycle genes. G2/M score vs S Score for all P0 IFE cells are shown. A subset of IFE.A1 population show high G2/M score, indicative of higher proliferation. F) Violin plot of G2/M scores for P0 IFE populations. A subset of IFE.A population shows high G2/M score.

Supplemental Figure 7



Supplemental Fig. 7. Related to Figure 7. A) Genotype of WT1, WT2, *Grhl3-/-1*, *Grhl3-/-2* are projected onto P0 IFE pseudotime trajectories shown in Fig. 7A. B) RNA-velocity analysis of WT1, WT2, *Grhl3-/-1*, *Grhl3-/-2* using PCA embedding. Note the PCA happens to delineate the differentiation trajectory from basal to transition to differentiated. The continuous differentiation shown in WT samples is abolished in *Grhl3-/-* and a propensity toward basal fate is observed.

Supplemental Figure 8



Supplemental Fig. 8. Related to Figure 8. A) Violin plots showing the expression of Wnt pathway components *Ctnnb1, Gpc3,* and *Gsk3b* in WT and *Grhl3-/-* basal and IFE.A1 populations. B) Heatmap showing the expression of genes that are expressed in a dynamic manner throughout E14.5, E16.5 and P0 IFE populations. C) Violin plots showing the expression of *Dkkl1* in WT and *Grhl3-/-* IFE.D and IFE.TD populations. D) Cartoon depicting a model for a gradualistic IFE differentiation process as opposed to the traditional model where differentiation occurs in punctuated stages.

Chapter 3. Methods

Mouse work

The *Grhl3-/-* mice and their genotyping were described previously[4]. Timed pregnancies were used to obtain embryos; the developmental stage of embryos was verified using external features of the embryos[51]. All experiments were performed on WT and *Grhl3-/-* littermates from the same pregnancy. Mice were fed food and water ad libitum and maintained on a regular 12h day/night cycle. All animal experiments were performed in accordance with Institutional Animal Care and Use Committee at University of California, Irvine (Protocol No. AUP-19-012).

Tissue isolation

For E14.5 and E16.5, back skin was micro-dissected from the embryos and incubated in 2.5U/mL Dispase (STEMCELL Technologies) in EpiLife medium (Thermo Fisher Scientific) for 2 hours at room temperature (RT). The tissue were then washed with media and incubated in Accutase (STEMCELL Technologies) for 30 min at RT followed by dissociation into single cells. For P0, back skin was dissected from the mice and incubated in 2.5U/mL Dispase in EpiLife medium overnight. The epidermis was then manually separated from the dermis and incubated in Accutase for 30 min at RT followed by dissociation into single cells. For all time points, the dissociated cell suspension was strained with 40-uM filter and washed with media. Dead Cell Removal kit (Miltenyi Biotec) was used to remove dead cells prior to resuspension in 0.04% BSA (Thermo Fisher Scientific).

Chromium Single Cell 3' v1/v2 (10X Genomics) library preparation was then performed by the Universtiy of California, Irvine, Genomic High Throughput Facility (UCI-GHTF) according to manufacturer's protocol.

Immunofluorescence localization of markers

Fresh frozen OCT 10 µM sections were incubated in acetone at -20^{o}C for 10 min, washed with TBS, fixed in 4% PFA for 10 min, and then washed with TBS 3 times. Tissues were then permeablized using TBS with 0.3% TritonX-100 for 10 min and blocked in TBS with 0.5% BSA for 1 h. Primary antibodies, Krt14 (Abcam) and Krt10 (Covance), were diluted 1:1000 and incubated overnight at 4°C. Secondary antibodies (Abcam, Life Technologies) were diluted 1:1000 and incubated at RT for 1 h. Images were acquired using a Keyence BZ-X700 fluorescent microscope. Fiji was used for image analysis.

RNA Fluorescent *in situ* hybridization (RNA-FISH)

RNA-FISH was performed using the RNAscope Multiplex Fluorescent Detection Kit v1 according to manufacturer's instructions on fresh frozen 10 µM thick OCT sections. All sections were counterstained with ProLong Gold antifade reagent with DAPI. Images were acquired on a Leica SP8 confocal microscope. To ensure that images were comparable, they were all processed the same maximum intensity projection and brightness. Three biological replicates were analyzed for each FISH staining experiment.

Sequencing

Chromium Single Cell 3' v1/v2 libraries were sequenced with either a Illumina HiSeq 2500 or a HiSeq4000 following the manufacturer's protocol.

Primary computational analysis

Raw sequencing data were demultiplexed and processed using Cellranger (10X Genomics version 2.01) using MM10 reference provided by 10X Genomics. Cells were filtered and clustered using Seurat version 2.3.4[47]. For E14.5 and E16.5, cells with <600 and >4800 genes detected were removed. For P0, cells with <900 and >7700 genes detected were removed. For all time points, genes that are expressed in <6 cells were removed, and cells with >10% mitochondrial genes detected were removed. Raw gene-cell matrices were normalized and scaled. Percentage of mitochondrial genes, and the number of unique molecular identifiers (UMI) were regressed out using the RegressOut function. High variable genes were identified using a x.low.cutoff of 0.0125 and y.cutoff of 0.5 on mean variance dispersion plot. Data for each developmental age were integrated with canonical correlation analysis (CCA) on each biological replicate followed by alignment of subspace of each sample. Louvain clustering was then performed on the integrated samples. Single-cell consensus clustering (SC3) was carried out using k=2 for basal IFE subclustering. Cell cycle analysis was carried out in Seurat using a list of cell cycle genes from Regev laboratory[48]. Diffusion maps were implemented using URD version 1.0.2[36].

Pseudotemporal trajectory and RNA-velocity analysis

Monocle version 2.10.1 was used to construct differentiation trajectories using the highly variable genes identified from Seurat package as an ordering filter. DDRTree was used for dimension reduction. No root state is specified and ordering was done in an unsupervised manner. The p-value cutoff for identifying pseudotime-dependent genes was $p<0.01$. To model gene expression changes in pseudotime, we used scEpath package[43], which first divides the pseudotime into 10 equally spaced bins. Then the expression of each gene in each bin is estimated by the trimean of the expressions of this gene across all the cells located in this bin. Furthermore, smoothens the average expression of each gene using cubic regression splines. To study the temporal patterns and functional signatures of pseudotime-dependent genes, we performed k-medoids clustering based on the smoothed gene expression profiles, using Matlab function "kmedoids" with six clusters. We then created a heatmap to show the normalized-smoothed expression pattern of pseudotime-dependent genes. Genes within this heatmap were ordered such that nearest neighbors have similar expression profiles and genes within each cluster were ordered according to expression peak. The average expression pattern of each cluster is calculated by the trimean of smoothed expressions of all the genes in that cluster.

RNA-velocity analysis was performed using the velocyto pipeline (https://github.com/velocyto-team/velocyto.R)[49]. The "run10x" function was used on Cellranger ouputs to generate the loom files using default parameters, mm10 gtf file provided by 10X Genomics and repeat mask gtf file from the UCSC

Genome Browser. For the RNA velocity estimation, we used the standard R implementation of velocyto and only considered cells that were part of the pseudotime. The cell-to-cell distance matrix was calculated based on the PCA embedding. RNA velocity was estimated using gene-relative model with k-nearest neighbor cell pooling (k = 20). Embeddings from PCA and Monocle pseudotime analysis were used for velocity field projections.

Chapter 4. Summary and Conclusion

Here we report the first single cell transcriptome profiling study of the developing mouse epidermis at E14.5, E16.5, and P0 in WT and *Grhl3-/-* mice. The data includes a total of 85,286 cells, a valuable resource available for analysis beyond the scope of this paper. We identified all known cell types of the epidermis and defined unique gene signatures for each cell type. We have also uncovered numerous hitherto unknown genes important for epidermal development from E14.5 to P0 (Supplemental Fig. 8B).

Our study, which focused on IFE differentiation and its regulation by GRHL3, has revealed a number of new insights: 1) At a transcript level, IFE differentiation is continuous and gradualistic rather than stepwise; 2) There is a large number of basal-spinous transition cells; 3) An important regulatory change occurs at the IFE.T/IFE.D boundary where IFE cells commit from a plastic state to an irreversible fate toward terminal differentiation; 4) There are two main populations of basal stem cells and a third state linked to cell proliferation; 5) In addition to a role in promoting terminal differentiation, GRHL3 is required at an earlier differentiation stage to suppress Wnt signaling and to suppress the expansion of stem cells.

Studies on stem cells in the basal layer of the adult epidermis have suggested cell heterogeneity within this layer, although the nature of this heterogeneity remains controversial. The hierarchical model of stem cell differentiation suggests that the basal layer comprises rare slow-cycling long-term stem cells and their fast-cycling committed progenitors[52,53]. Another related

model suggests that slow-cycling and fast-cycling stem cells occupy distinct regions of the basal layer and renew within their respective regions[8,9]. In contrast, the stochastic model of stem cell differentiation suggests that all basal cells are equivalent in terms of their stemness and that they differentiate in a stochastic manner[11,12,54,55]. We found two populations of stem cells in the P0 IFE with one major difference between these being the higher expression of *Wnt4* in IFE.B1. Wnt4 was previously shown to be important for IFE stem cell self-renewal[11] and our RNA-FISH result also supports a model in which *Wnt4*-high stem cells are distributed randomly throughout the basal IFE. We propose that *Wnt4*-high basal IFE cells are workhorses of IFE maintenance as increased *Wnt4* due to *Grhl3* mutation leads to considerable change in cell composition of the IFE.

Previous work on GRHL3 has focused on suprabasal IFE cells where *Grhl3* is expressed to its highest level. scRNA-seq allowed us to quantitatively characterize the cell compositional change in *Grhl3-/-* epidermis and uncover another important function of GRHL3, which is to suppress aberrant Wnt signaling in the basal layer. Our findings suggest that there is a single intermediate cell state (IFE.T) during IFE differentiation that is in a plastic state with the basal stem cells (IFE.B). GRHL3 promotes the progression of the IFE.B state to the IFE.T state by repressing aberrant Wnt signaling.

Our scRNA-seq data suggests paracrine mechanisms can be at play. The secreted Wnt pathway antagonist DKKL1, which is normally expressed in the terminally differentiated cells of the IFE, is lost in the *Grhl3-/-* IFE, thereby potentially decreasing the suppression of Wnt signaling in the basal layer (Fig.

8C, Supplemental Fig. 8C). Alternatively, data from human keratinocytes suggest that GRHL3 may act earlier in maintaining proper proliferation/differentiation balance[56,57].

Clearly the IFE contains morphologically distinct layers. Yet, our scRNA-seq findings challenge the conventional notion of stepwise IFE differentiation. At a transcript level, differentiation occurs in a continuous manner from the most primitive stem cell to the most terminally differentiated IFE cell at the top of the epidermis. Consistent with previous findings, our findings suggest that basal stem cells (IFE.B) exist in a reversible equilibrium with a large population of transition cells (IFE.T) and that is not until after passing the transition cell stage that IFE cells fully commit to differentiation (Supplemental Fig. 8D). We also uncovered a prominent accumulation of epidermal stem cells in the *Grhl3-/-* epidermis, which points to a previously unknown function of GRHL3 in the progression of basal cells to transition cells and thereby suppression of the basal stem cell population; and that GRHL3 achieves this effect by suppressing Wnt signaling in stem cells.

References

1. Rangarajan, A. *et al.* Notch signaling is a direct determinant of keratinocyte growth arrest and entry into differentiation. *The EMBO Journal* **20**, 3427–3436 (2001).
2. Koster, M. I. *et al.* p63 induces key target genes required for epidermal morphogenesis. *PNAS* **104**, 3255–3260 (2007).
3. Segre, J. A., Bauer, C. & Fuchs, E. Klf4 is a transcription factor required for establishing the barrier function of the skin. *Nat Genet* **22**, 356–360 (1999).
4. Yu, Z. *et al.* The Grainyhead-like epithelial transactivator Get-1/Grhl3 regulates epidermal terminal differentiation and interacts functionally with LMO4. *Dev. Biol.* **299**, 122–136 (2006).
5. Stern, C. D. Neural induction: old problem, new findings, yet more questions. *Development* **132**, 2007–2021 (2005).
6. Fuchs, E. Scratching the surface of skin development. *Nature* **445**, 834–842 (2007).
7. Mascré, G. *et al.* Distinct contribution of stem and progenitor cells to epidermal maintenance. *Nature* **489**, 257–262 (2012).
8. Gomez, C. *et al.* The Interfollicular Epidermis of Adult Mouse Tail Comprises Two Distinct Cell Lineages that Are Differentially Regulated by Wnt, Edaradd, and Lrig1. *Stem Cell Reports* **1**, 19–27 (2013).
9. Sada, A. *et al.* Defining the cellular lineage hierarchy in the interfollicular epidermis of adult skin. *Nature Cell Biology* **18**, 619–631 (2016).
10. Clayton, E. *et al.* A single type of progenitor cell maintains normal epidermis. *Nature* **446**, 185–189 (2007).
11. Lim, X. *et al.* Interfollicular Epidermal Stem Cells Self-Renew via Autocrine Wnt Signaling. *Science* **342**, 1226–1230 (2013).

12. Doupé, D. P., Klein, A. M., Simons, B. D. & Jones, P. H. The Ordered Architecture of Murine Ear Epidermis Is Maintained by Progenitor Cells with Random Fate. *Developmental Cell* **18**, 317–323 (2010).

13. Lopez-Pajares, V. *et al.* A LncRNA-MAF:MAFB Transcription Factor Network Regulates Epidermal Differentiation. *Developmental Cell* **32**, 693–706 (2015).

14. Hopkin, A. S. *et al.* GRHL3/GET1 and Trithorax Group Members Collaborate to Activate the Epidermal Progenitor Differentiation Program. *PLOS Genetics* **8**, e1002829 (2012).

15. Klein, R. H. *et al.* GRHL3 binding and enhancers rearrange as epidermal keratinocytes transition between functional states. *PLoS Genet.* **13**, e1006745 (2017).

16. Venkatesan, K., McManus, H. R., Mello, C. C., Smith, T. F. & Hansen, U. Functional conservation between members of an ancient duplicated transcription factor family, LSF/Grainyhead. *Nucleic Acids Res* **31**, 4304–4316 (2003).

17. Darido, C. *et al.* Targeting of the Tumor Suppressor GRHL3 by a miR-21-Dependent Proto-Oncogenic Network Results in PTEN Loss and Tumorigenesis. *Cancer Cell* **20**, 635–648 (2011).

18. Bhandari, A. *et al.* The Grainyhead transcription factor Grhl3/Get1 suppresses miR-21 expression and tumorigenesis in skin: modulation of the miR-21 target MSH2 by RNA-binding protein DND1. *Oncogene* **32**, 1497–1507 (2013).

19. Gordon, W. M. *et al.* A GRHL3-regulated repair pathway suppresses immune-mediated epidermal hyperplasia. *J Clin Invest* **124**, 5205–5218 (2014).

20. Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).

21. Tang, F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* **6**, 377–382 (2009).

22. Qiu, X. *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nature Methods* **14**, 979–982 (2017).

23. Brennecke, P. *et al.* Accounting for technical noise in single-cell RNA-seq experiments. *Nat Methods* **10**, 1093–1095 (2013).

24. Kiselev, V. Y., Andrews, T. S. & Hemberg, M. Challenges in unsupervised clustering of single-cell RNA-seq data. *Nat Rev Genet* **20**, 273–282 (2019).

25. Grün, D. *et al.* Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* **525**, 251–255 (2015).

26. Wang, B., Zhu, J., Pierson, E., Ramazzotti, D. & Batzoglou, S. Visualization and analysis of single-cell RNA-seq data by kernel-based similarity learning. *Nat Methods* **14**, 414–416 (2017).

27. Lin, P., Troup, M. & Ho, J. W. K. CIDR: Ultrafast and accurate clustering through imputation for single-cell RNA-seq data. *Genome Biology* **18**, 59 (2017).

28. Zeisel, A. *et al.* Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142 (2015).

29. Cheng, J. B. *et al.* Transcriptional Programming of Normal and Inflamed Human Epidermis at Single-Cell Resolution. *Cell Reports* **25**, 871–883 (2018).

30. Joost, S. *et al.* Single-Cell Transcriptomics Reveals that Differentiation and Spatial Signatures Shape Epidermal and Hair Follicle Heterogeneity. *cels* **3**, 221-237.e9 (2016).

31. Fan, X. *et al.* Single Cell and Open Chromatin Analysis Reveals Molecular Origin of Epidermal Cells of the Skin. *Developmental Cell* **47**, 21-37.e5 (2018).

32. Joost, S. *et al.* Single-Cell Transcriptomics of Traced Epidermal and Hair Follicle Stem Cells Reveals Rapid Adaptations during Wound Healing. *Cell Reports* **25**, 585-597.e7 (2018).

33. Watt, F. M. Involucrin and Other Markers of Keratinocyte Terminal Differentiation. *Journal of Investigative Dermatology* **81**, S100–S103 (1983).

34. Ting, S. B. *et al.* A Homolog of Drosophila grainy head Is Essential for Epidermal Integrity in Mice. *Science* **308**, 411–413 (2005).

35. Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat Commun* **8**, 1–12 (2017).

36. Farrell, J. A. *et al.* Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* **360**, eaar3131 (2018).

37. Haghverdi, L., Buettner, F. & Theis, F. J. Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics* **31**, 2989–2998 (2015).

38. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat Genet* **25**, 25–29 (2000).

39. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res* **47**, D330–D338 (2019).

40. Qiu, X. *et al.* Reversed graph embedding resolves complex single-cell developmental trajectories. *bioRxiv* (2017) doi:10.1101/110668.

41. Simpson, C. L., Patel, D. M. & Green, K. J. Deconstructing the skin: cytoarchitectural determinants of epidermal morphogenesis. *Nat Rev Mol Cell Biol* **12**, 565–580 (2011).

42. Schweizer, J., Kinjo, M., Fürstenberger, G. & Winter, H. Sequential expression of mRNA-encoded keratin sets in neonatal mouse epidermis: Basal cells with properties of terminally differentiating cells. *Cell* **37**, 159–170 (1984).

43. Jin, S., MacLean, A. L., Peng, T. & Nie, Q. scEpath: energy landscape-based inference of transition probabilities and cellular trajectories from single-cell transcriptomic data. *Bioinformatics* **34**, 2077–2086 (2018).

44. Kiselev, V. Y. *et al.* SC3: consensus clustering of single-cell RNA-seq data. *Nature Methods* **14**, 483–486 (2017).

45. Choi, Y. S. *et al.* Distinct Functions for Wnt/β-Catenin in Hair Follicle Stem Cell Proliferation and Survival and Interfollicular Epidermal Homeostasis. *Cell Stem Cell* **13**, 720–733 (2013).

46. Nusse, R. & Clevers, H. Wnt/β-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* **169**, 985–999 (2017).

47. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology* **36**, 411–420 (2018).

48. Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).

49. Manno, G. L. *et al.* RNA velocity of single cells. *Nature* **560**, 494–498 (2018).

50. Svensson, V. & Pachter, L. RNA Velocity: Molecular Kinetics from Single-Cell RNA-Seq. *Molecular Cell* **72**, 7–9 (2018).

51. Richardson, L. *et al.* EMAGE mouse embryo spatial gene expression database: 2014 update. *Nucleic Acids Res* **42**, D835–D844 (2014).

52. Wuidart, A. *et al.* Quantitative lineage tracing strategies to resolve multipotency in tissue-specific stem cells. *Genes Dev.* **30**, 1261–1277 (2016).

53. Tumbar, T. *et al.* Defining the Epithelial Stem Cell Niche in Skin. *Science* **303**, 359–363 (2004).

54. Doupé, D. P. & Jones, P. H. Interfollicular epidermal homeostasis: dicing with differentiation. *Experimental Dermatology* **21**, 249–253 (2012).

55. Rompolas, P. *et al.* Spatiotemporal coordination of stem cell commitment during epidermal homeostasis. *Science* **352**, 1471–1474 (2016).

56. Li, J. *et al.* HNRNPK maintains epidermal progenitor function through transcription of proliferation genes and degrading differentiation promoting mRNAs. *Nat Commun* **10**, 1–14 (2019).

57. Noutsou, M. *et al.* The Cohesin Complex Is Necessary for Epidermal Progenitor Cell Function through Maintenance of Self-Renewal Genes. *Cell Reports* **20**, 3005–3013 (2017).